Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Journal of Hydrology 320 (2006) 3–17
Journal of Hydrology
www.elsevier.com/locate/jhydrol

Model Parameter Estimation Experiment (MOPEX): An overview of science strategy and major results from the second and third workshops

Q. Duan[a,*], J. Schaake[b], V. Andréassian[c], S. Franks[d], G. Goteti[e], H.V. Gupta[f], Y.M. Gusev[g], F. Habets[h], A. Hall[i], L. Hay[j], T. Hogue[k], M. Huang[l], G. Leavesley[j], X. Liang[l], O.N. Nasonova[g], J. Noilhan[h], L. Oudin[c], S. Sorooshian[m], T. Wagener[n], E.F. Wood[o]

[a]*Energy and Environment Directorate, Lawrence Livermore National Laboratory, 7000 East Avenue, Livermore, CA 94550, USA*
[b]*Office of Hydrologic Development, National Weather Service, Silver Spring, MD, USA*
[c]*Groupe Hydrologie, Cemagref, Antony, France*
[d]*Department of Civil, Surveying & Environmental Engineering, The University of Newcastle, Callaghan, Australia*
[e]*Department of Earth System Sciences, University of California, Irvine, CA, USA*
[f]*Department of Hydrology and Water Resources, University of Arizona, Tucson, AZ, USA*
[g]*Institute of Water Problems, Russian Academy of Sciences, Moscow, Russia*
[h]*Météo France, Toulouse, France*
[i]*Water Resources Application Project/GEWEX, Cooma, Australia*
[j]*US Geological Survey, Denver, CO, USA*
[k]*Department of Civil & Environmental Engineering, University of California, Los Angeles, CA, USA*
[l]*Department of Civil & Environmental Engineering, University of California, Berkeley, CA, USA*
[m]*Department of Civil & Environmental Engineering, University of California, Irvine, CA, USA*
[n]*Department of Civil & Environmental Engineering, Pennsylvania State University, University Park, PA, USA*
[o]*Department of Civil Engineering, Princeton University, Princeton, NJ, USA*

Abstract

The Model Parameter Estimation Experiment (MOPEX) is an international project aimed at developing enhanced techniques for the a priori estimation of parameters in hydrologic models and in land surface parameterization schemes of atmospheric models. The MOPEX science strategy involves three major steps: data preparation, a priori parameter estimation methodology development, and demonstration of parameter transferability. A comprehensive MOPEX database has been developed that contains historical hydrometeorological data and land surface characteristics data for many hydrologic basins in the United States (US) and in other countries. This database is being continuously expanded to include more basins in all parts of the world. A number of international MOPEX workshops have been convened to bring together interested hydrologists and land surface modelers from all over world to exchange knowledge and experience in developing a priori parameter estimation techniques.

* Corresponding author.
E-mail address: qduan@llnl.gov (Q. Duan).

doi:10.1016/j.jhydrol.2005.07.031

This paper describes the results from the second and third MOPEX workshops. The specific objective of these workshops is to examine the state of a priori parameter estimation techniques and how they can be potentially improved with observations from well-monitored hydrologic basins. Participants of the second and third MOPEX workshops were provided with data from 12 basins in the southeastern US and were asked to carry out a series of numerical experiments using a priori parameters as well as calibrated parameters developed for their respective hydrologic models. Different modeling groups carried out all the required experiments independently using eight different models, and the results from these models have been assembled for analysis in this paper. This paper presents an overview of the MOPEX experiment and its design. The main experimental results are analyzed. A key finding is that existing a priori parameter estimation procedures are problematic and need improvement. Significant improvement of these procedures may be achieved through model calibration of well-monitored hydrologic basins. This paper concludes with a discussion of the lessons learned, and points out further work and future strategy.

Keywords: MOPEX; A priori parameter estimation; Model calibration; Rainfall–runoff modeling; Regionalization; Uncertainty analysis

1. Introduction

A critical step in applying a hydrologic model to a watershed or a land surface parameterization scheme (LSPS) of an atmospheric model to a specific grid element is to estimate the coefficients or constants in the model or LSPS known as parameters. These parameters are inherent in all models. While certain parameters may take on universally accepted values (e.g. gas constant, acceleration of gravity), the values of many parameters are not universally constant and may be highly uncertain. In general, parameters vary spatially so they are unique to each watershed or to a grid point, and some may even vary seasonally. Moreover, some parameters may be space-time scale dependent (Koren et al., 1999; Finnerty et al., 1997). The question of how to estimate model parameters has been receiving increasing attention from the hydrology and land surface modeling community (Franks and Beven, 1997; Bastidas et al., 1999; Gupta et al., 1999; Duan et al., 2001; Duan et al., 2003; Jackson et al., 2003; Wagener et al., 2003).

A common approach within the hydrologic modeling community to parameter estimation is to calibrate hydrologic models to historical observations by tuning model parameters. A plethora of model calibration techniques have been reported in the literature. For a detailed review of model calibration techniques, readers are referred to Duan et al. (2003); Duan (2003). To conduct model calibration, a sufficient amount of historical hydrologic data is typically required. Hydrologists have the advantage of working with watersheds, many of which are well monitored with climate stations and stream gauges. For ungauged basins and for LSPS applications, it is difficult to obtain adequate data that are needed for model calibration. A further complication is that LSPSs are typically applied to large spatial scales and involve many grid elements. To estimate model parameters in these cases, it is necessary to assign model parameter values a priori.

A priori parameter estimation procedures are available for many hydrologic models and LSPSs. But these procedures have not been fully validated through rigorous testing using retrospective hydrometeorological data and corresponding land surface characteristics data. This is partly because the necessary database for such testing has not been available until recently. Moreover, there is a gap in our understanding of the links between model parameters and the land surface characteristics. Generally, available information about soils (e.g. texture) and vegetation (e.g. type or vegetation index) only indirectly relates to model parameters such as hydraulic properties of soils and rooting depths of vegetation. Some models which are built using a top-down approach are by nature empirical, and no direct link has yet been established between measurable watershed characteristics and model parameters. Also it is not clear how heterogeneity associated with spatial land surface characteristics data affects those characteristics at the scale of a basin or a grid cell. Consequently, there is a considerable degree of uncertainty associated with the parameters given by existing a priori procedures.

The Project for Intercomparison of Land-surface Parameterization Schemes (PILPS) has revealed widely discrepant simulation results by different

LSPSs (see Chen et al., 1997; Wood et al., 1998; Pitman et al., 1999; Schlosser et al., 2000; Slater et al., 2001). Interestingly, the LSPSs included in the PILPS experiments were driven by the same meteorological forcing data and were required to use the same values for commonly named parameters (such as soil hydraulic properties and vegetation phenology parameters). The large scattering of model results can be partially explained by the uncertainty in the values of the parameters used in each scheme.

The improper choice of model parameters results in poor model performance (Liston et al., 1994; Duan et al., 1995). It is necessary to develop enhanced a priori parameter estimation methodologies for hydrologic models and LSPSs. Toward this goal, a project known as the Model Parameter Estimation Experiment (MOPEX) was initiated in 1996. The MOPEX project has been an international collaborative endeavor, with the involvement of international scientists and hydrologic data assembled from different countries. MOPEX has the endorsement of several international organizations and projects including: the World Meteorological Organization (WMO) Commission on Hydrology, International Association of Hydrological Sciences (IAHS) Prediction for Ungauged Basins (PUB) Initiative (Sivapalan, 2003) and the Global Energy and Water Cycle Experiment (GEWEX). The Office of Global Programs in the National Oceanic and Atmospheric Administration (NOAA) and funding agencies in different countries have all provided financial support for scientists to participate in MOPEX activities. A series of international workshops on MOPEX have been convened over the last few years. The first one was held in July 1999, as a part of International Union of Geodesy and Geophysics (IUGG) 21st General Assembly in Birmingham, England. The second MOPEX workshop, co-sponsored by the National Weather Service Hydrology Laboratory (NWS/HL) and National Science Foundation Center for Sustainability of semi-Arid Hydrology and Riparian Areas (SAHRA) at the University of Arizona, was held in Tucson, Arizona, in April 2002. The third MOPEX workshop was held in Sapporo, Japan, in July 2003 as a part of the 22nd IUGG General Assembly. The fourth MOPEX workshop was held in Paris, France in July 2004, co-sponsored by Cemagref of France and the NWS/HL. The fifth MOPEX workshop was held in Foz do Iguaçu, Brazil, in April 2005.

The MOPEX workshops were designed to bring together interested international hydrologists and land surface modelers to share experience in estimation of hydrologic model parameters. Each workshop has a special focus, either in terms of hydroclimatology (i.e. humid or semi-arid) or in terms of special applications (i.e. flood forecasting).

This paper presents an overview of the results from the second and third MOPEX workshops. For these workshops, a set of numerical experiments was constructed. The MOPEX participants were given data for 12 basins located in the southeastern quadrant of the US. Numerical test results from different modeling groups were assembled for the workshops. The paper is organized as follows. First the MOPEX rationale and science strategy are presented. Then a discussion of the objectives and numerical experiment design is given. The data sets assembled for the workshop are described, and a comprehensive analysis of the results is conducted to understand the differences in the results from the different models. Finally, further work and future strategy are discussed.

2. Model Parameter Estimation Experiment strategy

The MOPEX science strategy involves three major steps (Fig. 1). The first step is to develop the necessary data sets. The second step is to use these data to develop a priori parameter estimation methodology. The third step is to demonstrate that new a priori techniques produce better model results than existing a priori techniques for basins which were not used to develop the new a priori techniques.

Step two is accomplished using a three-path strategy illustrated in Fig. 1. The first path makes reference model runs with parameters estimated using existing a priori parameter estimation procedures. The second path makes model runs using calibrated values of selected model parameters. Within the second path, the calibrated parameters are analyzed to improve relations between model parameters and basin characteristics (i.e. climate, soils, vegetation and topographic features). These improved relations are

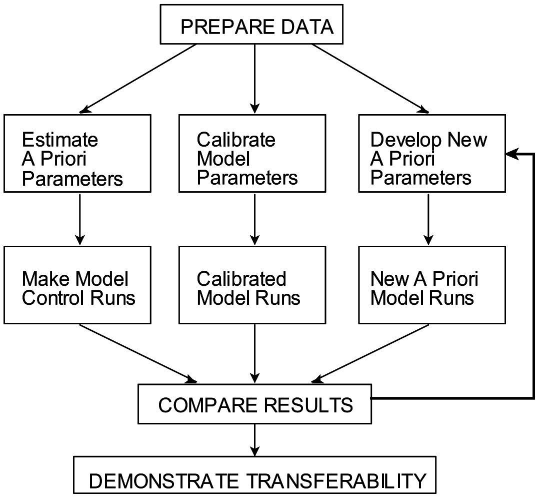

Fig. 1. MOPEX implementation strategy.

used to estimate new a priori parameters which are then used to make model runs in the third path. The success of step two is measured by how much improvement in model performance is achieved in the third path compared with results from the reference runs in the first path.

The MOPEX Project has assembled hydrometeorological data as well as land surface characteristics data that are needed to implement its parameter estimation strategy. Data from many basins in the US and other parts of the world are being assembled which cover a wide variety of climates. In Sections 3.2 and 3.3, the MOPEX data requirements and the data set used for the second and third MOPEX workshops are described in details.

A key in implementing the MOPEX strategy is to develop systematic procedures for calibration of selected model parameters and to apply these procedures to a large number of basins in different hydroclimatic regimes. Then, empirical relations will be sought between the parameters and various characteristics of soils, vegetation and climate. Much progress has been made in the area of model calibration (Duan et al., 2003). Duan et al. (1992, 1994) developed a robust optimization method known as Shuffled Complex Evolution (SCE-UA) method for optimal estimation of model parameters. Yapo et al. (1998); Gupta et al. (1999) have extended Duan's approach in the context of multi-objective theory. Recently, there is a surge of interest toward producing multiple sets of model parameters, as a means to account for uncertainty related to model structure, calibration data and model parameters. These methods use Monte Carlo sampling techniques to produce a set of solutions, all of which are regarded as 'equifinal' (i.e. all of the solutions are equally valid). Examples include the Generalized Likelihood Uncertainty Estimation (GLUE) by Beven and Binley (1992), the Markov Chain Monte Carlo (MCMC) Metropolis scheme by Kuczera and Parent (1998) and the Shuffled Complex Evolution Method Metropolis (SCEM) scheme by Vrugt et al. (2003). For more on the state-of-the-art on model calibration methods, readers are referred to Duan et al. (2003).

Numerous studies have been directed at developing improved a priori parameter estimation procedures for hydrologic models and LSPSs. Earlier examples of a priori parameter estimation procedures are from the field of soil physics, in which soil hydraulic properties (as appear in many hydrologic models and LSPSs) are related to soil texture classes in a tabular format (see e.g. Clapp and Hornberger, 1978; Cosby et al., 1984; Rawls et al., 1991; Carsel and Parrish, 1988). Many land surface modelers have directly adopted the a priori parameter estimation schemes developed by soil physicists to assign values to parameters in LSPSs (Dickinson et al., 1986; Sellers et al., 1986). Duan et al. (2001) pointed out that this practice is questionable because the tabular relations between soil hydraulic properties and soil texture classes are based on analysis of soil samples tested at laboratories. These relations may not hold up in the real world, especially over grid elements of several hundred to several thousand square kilometres. For typical hydrologic models and LSPSs, it is often the case that the relations between many of the model parameters and land surface characteristics are not obvious. One approach to solving this dilemma is to develop a priori relations between land surface characteristics and model parameters for basins where the model is appropriately calibrated (Abdulla et al., 1996; Duan et al., 1996; Merz and Bloschl, 2004; Lamb and Kay, 2004; McIntyre et al., 2004; Wagener et al., 2004). With the advent of Geographic Information Systems (GIS), many more a priori parameter estimation procedures have appeared. These schemes are model specific and are still being evolved. A number of these GIS-based schemes are being tested in the second and third MOPEX workshops and are part of the analysis presented in this paper.

3. Design and database of the second and third MOPEX workshops

3.1. Workshop objectives

The second and third MOPEX workshops focused on the second step of the MOPEX strategy: data preparation and development of parameter estimation procedures. The emphasis of the workshops was on validating existing a priori procedures and on evaluating potential improvement from model calibration. Because all hydrologic models are formulated differently, parameter estimation procedures are model-specific. A challenge facing hydrologic modelers is how the knowledge gained from one model can be transferred to another model. This is also the principal reason to convene these MOPEX workshops. A specific objective of these workshops is to examine the state of a priori parameter estimation techniques and how they can be potentially improved with observations from well-monitored hydrologic basins. Particularly, we seek to answer the following questions:

(1) How do we define the relations between model parameters and basin characteristics?
(2) How can model calibration be used to refine the a priori parameters?
(3) How do we evaluate the uncertainty due to model structure, calibration data and model parameters?

3.2. Design of MOPEX numerical experiment

To answer the above questions, a set of numerical experiments was designed. Data for 12 basins located in the southeastern quadrant of the United States were prepared. The data sets include hydrometeorological data as well as basin land surface characteristics data. More discussion on these data sets is given in Section 3.3. The data were distributed to MOPEX participants via ftp and CD-ROMs. The MOPEX participants were asked to make two sets of model runs. In the first set of model runs, the participating modelers were asked to run their respective models on all 12 basins using existing a priori parameters developed for their models. The second set of model runs involved model calibration for pre-selected common data periods. After model calibration, the participants were asked to run their models using calibrated parameters for the calibration and verification data periods. All results were collected for analysis by the MOPEX workshop organizers.

3.3. Description of the data set

3.3.1. MOPEX data requirements

The initial step in the MOPEX strategy is to assemble a large number of high quality data sets for a wide range of Intermediate Scale Area (ISA) river basins (500–10,000 km^2) throughout the world. There are strict requirements for MOPEX data sets in terms of data type, quantity and quality. The two basic data types gathered for MOPEX basins are hydrometeorological data and land surface characteristics data. The MOPEX basins should be unregulated basins and cover a variety of climate regimes. The basic hydrometeorological data required for MOPEX include daily precipitation, daily maximum and minimum temperature, daily streamflow data and climatic potential evaporation data. More desirable hydrometeorological data include hourly surface meteorological data, including precipitation, incoming long-wave and short-wave radiation, air temperature, air humidity, atmospheric pressure, and wind speed, etc. The quality of precipitation data is critically important to parameter estimation. MOPEX has established a minimum density requirement for raingauges based on basin size (Schaake et al., 2000). To ensure various hydrologic events are represented in the hydrometeorological data, MOPEX requires that the data length exceed 10 years. A desirable data length is 20 years or more.

The basic land surface characteristics data include basin boundary, soil texture and vegetation type data. More desirable land surface data sets include high resolution (1 km or finer) Digital Elevation Model (DEM) data, seasonal land cover/land use data such as Normalized Deviation of Vegetation Index (NDVI),

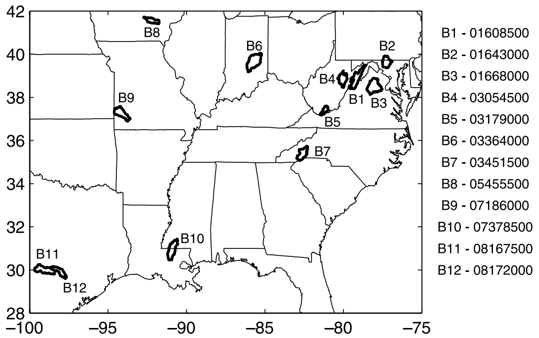

Fig. 2. Location of 12 basins for second MOPEX workshop in Tucson.

greenness fraction, snow cover and soil moisture climatology, etc.

3.3.2. MOPEX data for the second and third MOPEX workshops

For the second and third international MOPEX workshops, hydrometeorological data as well as basin land surface characteristics data for 12 basins in the Southeastern quadrant of the United States were assembled. Fig. 2 shows the location of the 12 basins. These basins represent a wide range of different climate, as indicated by the ratios of annual precipitation (P) and potential evapotranspiration (PE) in Fig. 3. A high value for P/PE indicates wet climate and a low value indicates dry climate (Dooge, 1997). The climatic seasonal precipitation and streamflow distributions are shown in Fig. 4.

The hydrometeorological data sets prepared for the workshops included hourly mean areal precipitation, daily streamflow, and climatic daily potential evapotranspiration. The hourly precipitation data sets were developed by the NWS Hydrology Laboratory (HL) based on hourly and daily raingauge data from the National Climate Data Center (NCDC). The daily streamflow data were obtained from the US Geological Survey (USGS). The climatic potential evaporation data was derived from the NOAA Freewater Evaporation Atlas (Farnsworth and Peck, 1982). Also included are basin averaged hourly meteorolological forcing data, including precipitation, air temperature, wind speed, surface pressure, short-wave and long-wave radiation and specific humidity. All meteorological forcing data except precipitation were processed from the 1/8° meteorological forcing data for the conterminous US developed by the University of Washington (UW) (Maurer et al., 2001). The UW hourly meteorological data set is derived from NCDC daily precipitation, daily minimum and maximum temperature and wind speed data obtained from National Center for Environmental Predictions/National Center for Atmospheric Research (NCEP/NCAR) Global Reanalysis data (Kistler et al., 2001). The historical data from different sources span over different data periods. For this study, a common period, 1960–1998, is chosen so data from all sources are available.

The land surface characteristics data sets assembled for this study include 1 km soil type data from the STATSGO data set (Miller and White, 1999), the 1 km vegetation type, and 5-min greenness fraction data (Loveland et al., 2000; Hansen and Reed, 2000; Gutman and Ignatov, 1998). Table 1 lists the major land surface properties of each basin including the area, elevation, and dominant soil and vegetation

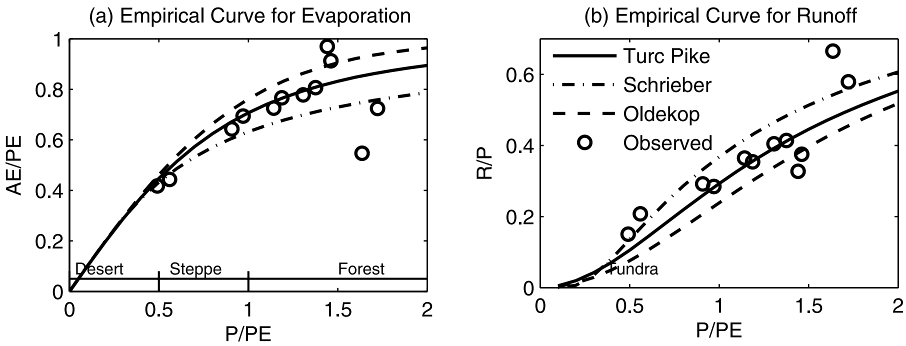

Fig. 3. Ratios of average annual hydrological variables.

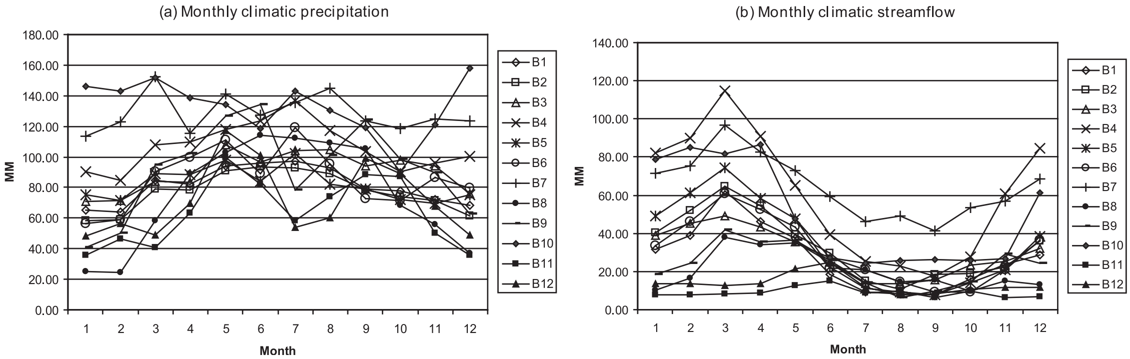

Fig. 4. Climatic monthly precipitation and streamflow.

Table 1
The basin land surface properties and average annual hydrologic variables

USGS ID	Lat.	Lon.	Area (km^2)	Elev. (m)	Soil type	Veg. type
01608500	39.4469	−78.6544	3810	171	Loam	Dec. broad leaf
01643000	39.3880	−77.3800	2116	71	Silt loam	Dec. broad leaf
01668000	38.3222	−77.5181	4134	17	Clay loam	Mixed forest
03054000	39.1500	−80.0400	2372	390	Loam	Dec. broad leaf
03179000	37.5439	−81.0106	1020	465	Si cl loam/loam	Dec. broad leaf
03364000	39.2000	−85.9256	4421	184	Si loam/cl loam	Croplands
03451500	35.6092	−82.5786	2448	594	Loam	Mixed forest
05455500	41.4664	−91.7156	1484	193	Clay loam	Cropland
07186000	37.2456	−94.5661	3015	254	Si loam/cl loam	Dec. broad leaf
07378500	30.4639	−90.9903	3315	0	Silt loam	Ever. Needleaf
08167500	29.8606	−98.3828	3406	289	Clay	Crop/nat. veg.
08172000	29.6650	−97.6497	2170	98	Clay	Crop/nat. veg.

type. Other land surface data made available to MOPEX participants include basin boundary, elevation, monthly surface albedo and roughness length. Basin climatologic data such as monthly long-term average precipitation, streamflow and potential evapotranspiration have also been made available.

4. Results and analysis

Eight hydrologic models and LSPSs have completed all of the required numerical experiments as described in Section 3.2. A few additional groups submitted incomplete numerical experiment results which have not been included in the analysis. Table 2 lists the eight participating models. Of the eights models, the first four models (SWB, SAC, GR4J and PRMS) are watershed rainfall–runoff models, while the last three (ISBA, SWAP, and Noah models) are LSPSs. The VIC model has been used both as a watershed model and a LSPS in atmospheric models. The analysis presented below is based on the comparison of the simulated streamflow from the eight models and the corresponding observations at daily or monthly time steps. It should be emphasized that the purpose of the intercomparison study is not intended to rank the models as being ‘better’ or ‘worse’ with respect to each other. Instead, the intercomparison study was conducted to understand the differences between approaches and use this knowledge to develop new a priori parameter estimation procedures. For this reason, this paper lists all participating models in Table 2, but the analysis does not refer to individual model names directly in all subsequent text or figures.

Table 2
Participating models and modeling agencies

Model names	Model agencies
Simple Water Balance (SWB)	NWS, USA
Sacramento (SAC)	NWS, USA
GR4J	Cemagref, France
PRMS	USGS, USA
VIC-3L	University of California at Berkeley/ Princeton University, USA
ISBA	Météo France, France
SWAP	Russian Academy of Sciences, Russia
Noah LSM	NWS, USA

4.1. Simulation results using existing a priori parameters

The purpose of simulations using existing a priori parameters is to establish benchmarks for the current a priori parameter estimation procedures used by the participating models. Any new a priori parameter procedures developed in the future for these models should at least perform better than the benchmarks. It should be noted that among the eight models under study, some models already have established a priori parameter estimation procedures, while others have no such systematic procedures. This discrepancy is reflected in the results discussed below. Fig. 5 displays the comparison of the simulated average annual streamflow totals from the a priori runs and the corresponding observed values. The spread of simulated streamflow annual totals is quite large between the models. None of the models were able to generate simulated streamflow values that match the observed values for all basins. The maximum over-bias exceeds 400 mm/year and the maximum under-bias is about 340 mm/year.

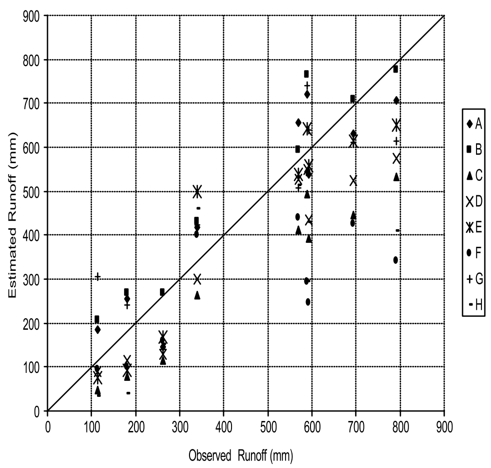

Fig. 5. Comparison of simulated and observed streamflow when a priori parameters are used.

The Nash-Sutcliffe (NS) efficiency is a commonly used goodness-of-fit measure between the simulated time series and observed time series. It is expressed as:

$$\mathrm{NS} = 1 - \frac{\sum_{i=1}^{n}(Q_i - Q_i^*)^2}{\sum_{i=1}^{n}(Q_i - \bar{Q}_i)^2} \quad (1)$$

where Q_i* and Q_i are the simulated and observed values at time i, and n is the number of data points. $\bar{Q}$ is the average of observed values. A value of 1 indicates perfect fit between $Q*_i$ and Q_i, while a value of <0 implies that simulated value is (on average) a poorer predictor than the long-term mean of the observations. Fig. 6 shows the NS efficiency of the daily streamflow simulations by the eight models. The NS values have been sorted from the lowest to the highest for each model. Fig. 7a and b shows the means and the standard deviations of the NS values, respectively. These figures reveal some interesting findings. Even though some models have some of the higher ranked NS values for most basins, they do not rank high for all basins. On the other hand, some models are shown to be consistent in all basins. This consistency is reflected in the low standard deviations for those models. (e.g. Models C,E,F and H). It should

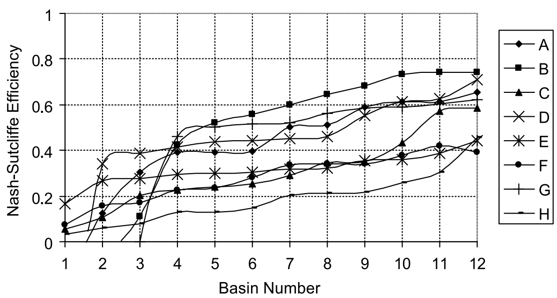

Fig. 6. Daily Nash-Sutcliffe efficiency of each model, sorted in increasing order, from a priori results.

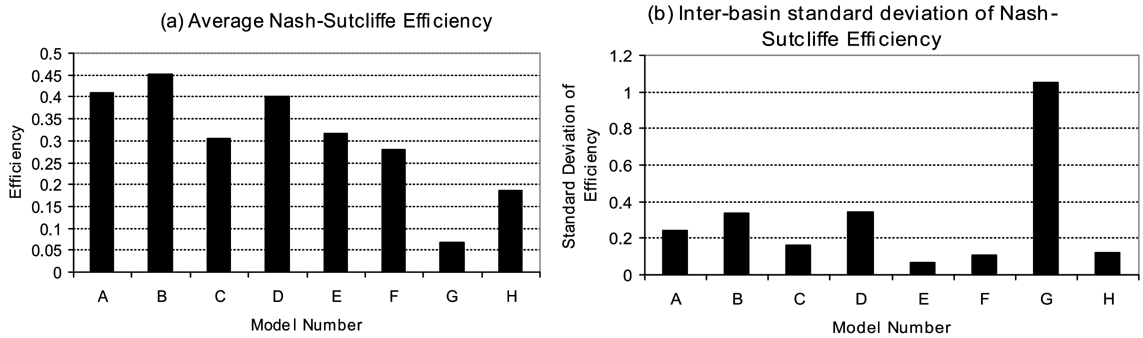

Fig. 7. Average daily Nash-Sutcliffe efficiency of each model and the standard deviations from a priori results.

also be noted that some models perform worse than long-term average for some basins, indicating a definitive need to improve a priori parameter estimates under those circumstances.

Figs. 8 and 9 show the same information for all models as in Figs. 6 and 7, but are evaluated on a monthly time step. The NS statistics for watershed models on a monthly time step generally show an improvement over those on a daily time step, while three LSPSs display a degraded average performance. For some models, the model simulations produce worse statistics than the long-term average of observations for some basins. Model G has good NS statistics for most basins compared to other models. But the large negative NS statistics for two basins have dragged down the average NS statistic to 0. The fact that a model does well for most basins, but poorly for only a few, tells us that the modeler should probably focus attention on the basins with poor results when looking for enhanced a priori parameter estimates.

4.2. Simulation results using calibrated parameters

There are several objectives in this exercise. First, we hope to quantify the potential improvement in model performance when the models are calibrated using observations, as compared to those using a priori parameters. Second, we want to make sure that there is consistency in streamflow simulations between calibration and validation data periods when the calibrated parameters are used. The ultimate objective of this exercise is to use the calibrated parameters to establish new a priori parameter estimates.

All model groups were asked to calibrate and validate their models for all 12 basins using historical hydrologic data. Originally, a split sample approach was to be used. Years 1980–1990 were to be used for calibration, while the first 19 years (1960–1979) were to be used for validation. Because different groups used different 19-year periods for calibration, it is not possible to make a direct comparison of all eight models using the split-sample approach. However, since the differences in the calibrated model performance between the different 19-year periods were much smaller than the differences in model performance between the a priori and calibrated runs, it seemed the best way to achieve the study objectives was to use the entire 1960–1998 period to evaluate model results for both the a priori and calibrated runs.

Fig. 10 shows the simulated average annual streamflow totals using calibrated parameters versus the observed average annual streamflow totals.

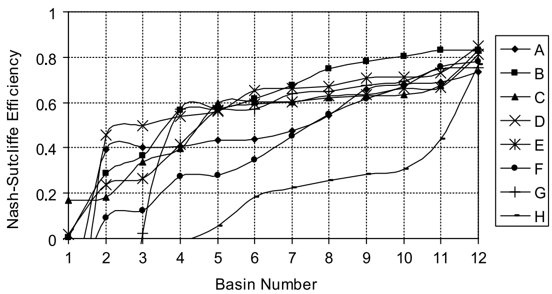

Fig. 8. Monthly Nash-Sutcliffe efficiency of each model, sorted in increasing order, from a priori results.

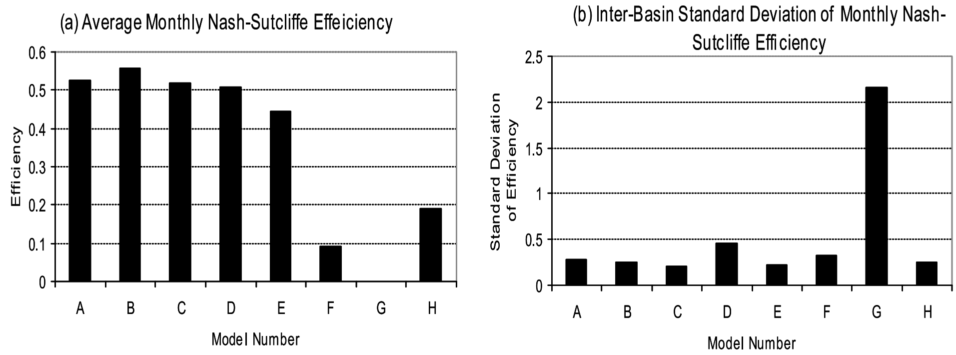

Fig. 9. Average monthly Nash-Sutcliffe efficiency and the standard deviation of each model from a priori simulation results for the entire data period 1960–1998.

Compared to Fig. 5, the scatter around the diagonal line is much smaller, indicating the agreement between observed and simulated streamflow is better when using calibrated parameters versus a priori parameters. Fig. 11 displays the sorted NS values for all models for the calibration period 1980–1998, while Fig. 12 shows the average NS values and standard deviations at the daily time step. All of these figures confirm that the NS values have been improved compared to the results. All NS values are now positive when calibrated parameters are used.

Figs. 13 and 14 show the same information as in Figs. 10 and 11 for all models, but the NS values are computed using monthly aggregated values.

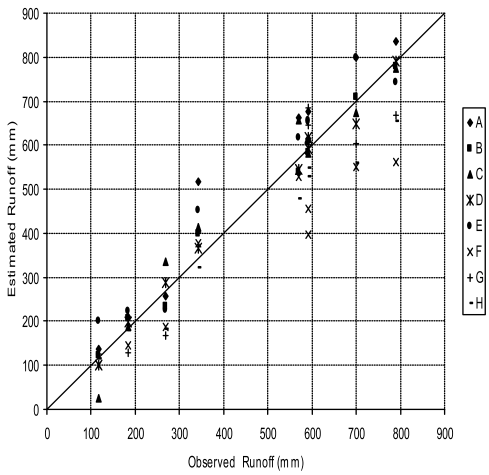

Fig. 10. Comparison of simulated and observed streamflow when calibrated parameters are used.

4.3. Calibration versus a priori results

Fig. 15a and b compares the daily and monthly NS values, respectively, for the entire data period where a priori and calibrated parameters are used. Both figures show that almost all of the points are on the left side of the diagonal line, indicating improvement resulting from the calibration exercise. The improvement is more apparent when examining monthly NS statistics. There are certain cases when the NS values from the calibration runs do not improve over those from the a priori runs. This is due to the fact that different modeling groups performed model calibration using different approaches. Particularly for one model (Model F), the modeler did not calibrate its model parameters to fit observed streamflow data during calibration. For another model (Model G), the modeler manually calibrated only one parameter (soil hydraulic conductivity at saturation) to get a

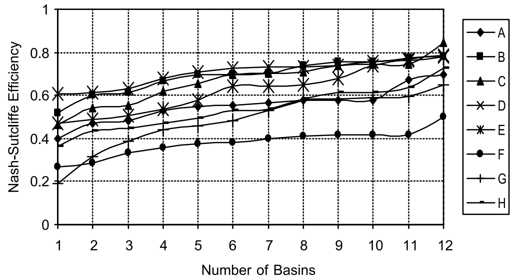

Fig. 11. Daily Nash-Sutcliffe efficiency of each model, sorted in increasing order, from calibrated results.

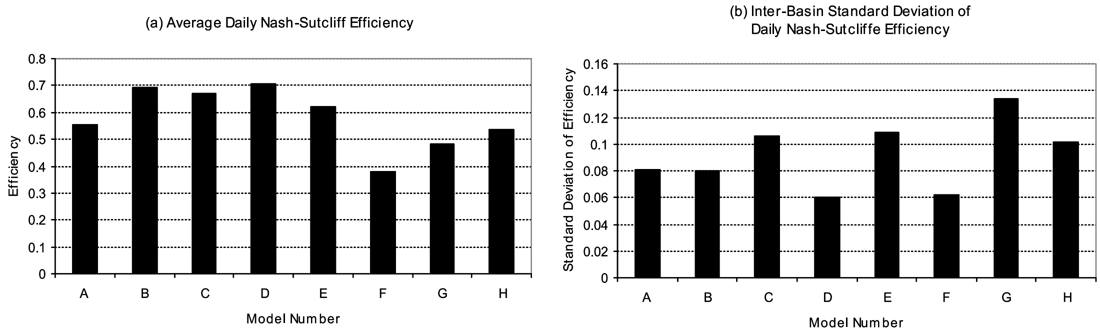

Fig. 12. Average daily Nash-Sutcliffe efficiency of each model and the standard deviations from calibrated results.

good agreement between observed and simulated annual streamflow.

4.4. Joint correlation between simulated streamflow from multiple models and observations

It is recognized that each of the models participating in this study is an imperfect representation of the hydrologic process that occur in the real world. It seemed interesting to ask how much total information about each basin is contained in the set of all models. Accordingly, the simulated streamflow time series from all eight models are used together as independent variables to construct a multiple regression model to predict the observed streamflow. The joint correlation coefficient from this regression analysis is a measure of the total information content of all of the models, jointly. By comparing the joint correlation coefficient from the regression analysis with the simple correlation coefficients for each model we can get an idea not only of the total information content but also which models contribute most of the information. Fig. 16 shows the scatter plot of the joint correlation coefficients and individual correlation coefficients at the daily time step. All of the points lie to the left of the diagonal line, which delineates the limiting value of the regression coefficient for any individual model. The relative position of points along the abscissa indicates the contribution of individual models to the joint correlation. In Fig. 16a, it is clear that Model B contributes most to the joint correlation because most of the points associated with this model are closest to the diagonal line. In Fig. 16b, a number of models make significant contribution to the joint correlation. These figures point to the potential that the multi-model approach (e.g. Georgakakos et al., 2004) is a plausible approach to obtain improved prediction.

5. Lessons, conclusions and future directions

A summary and analysis of the numerical experiment results of eight different models submitted to the second and third MOPEX workshops was presented. A number of lessons can be drawn from these results. First, the results confirm earlier statements that the existing a priori parameter estimation procedures are problematic and need improvement.

Second, calibration results clearly demonstrate the huge potential for improvement in a priori parameter estimation. Third, different models seem to represent hydrologic processes differently and all of them are imperfect. This suggests it may be possible to improve some of the models. It also suggests that improved

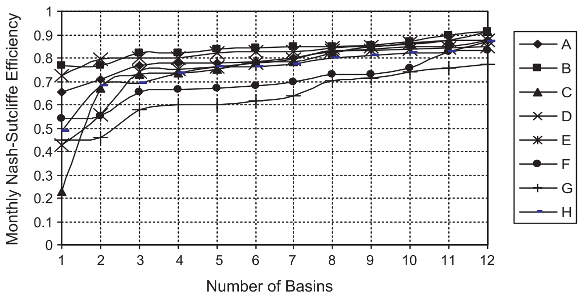

Fig. 13. Monthly Nash-Sutcliffe efficiency of each model, sorted in increasing order, from calibrated results.

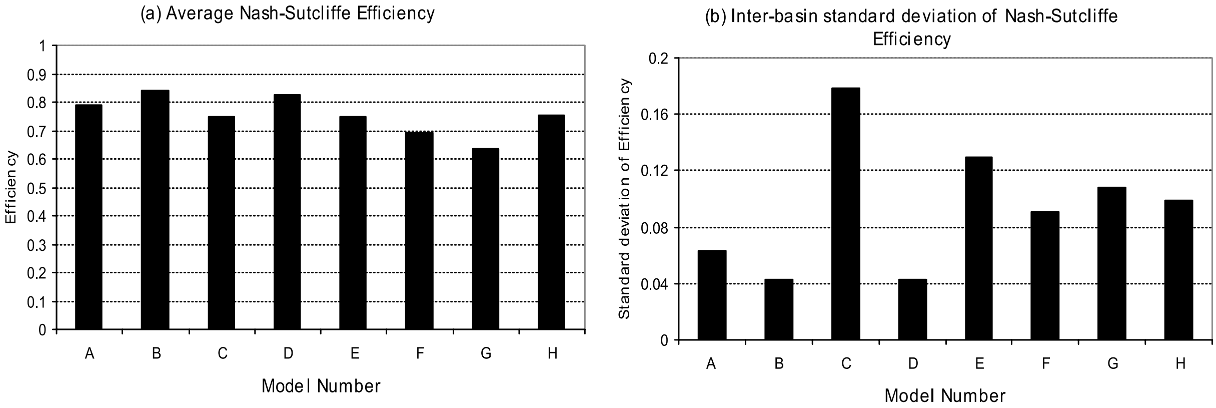

Fig. 14. Average monthly Nash-Sutcliffe efficiency and the standard deviation of each model from calibrated results for the calibration data period 1980–1998.

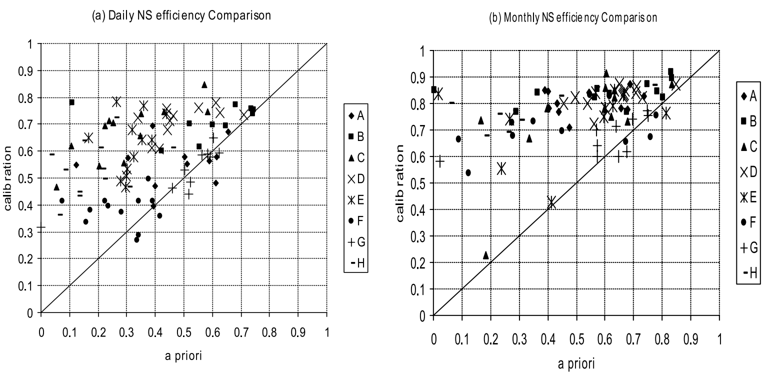

Fig. 15. Comparison of the NS values when calibrated and a priori parameters are used. (a) Evaluated at daily time step; (b) evaluated at monthly time step.

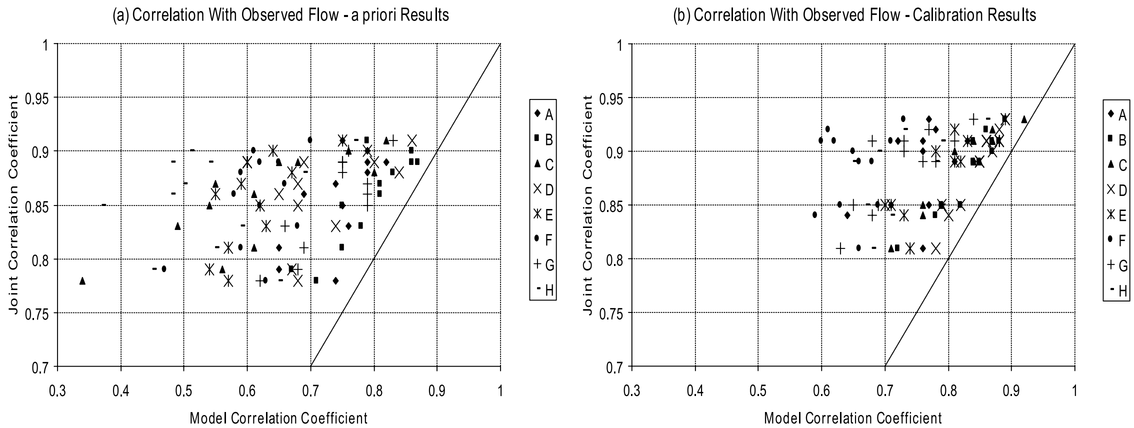

Fig. 16. Comparison of joint correlation coefficients and individual correlation coefficients.

prediction may be possible through an ensemble of different models or, possibly, an ensemble of a given model using different parameter sets.

Much research needs to be done to understand how model parameters are related to basin characteristics, especially considering that modelers are not sure that the presently available *observable* characteristics (mostly land surface characteristics) are the most relevant descriptors of the factors that control basin hydrological behavior. Further, how to use the calibrated results for improving a priori parameters is still not clear and this issue also needs addressing. Different modeling groups can learn from each other because many model parameters have similar physical interpretations and should have similarity in space-time patterns.

One issue that has not been examined in the workshops is the parameter transferability issue. This issue is very important for Predictions for Ungaged Basins (PUBs) and for application in land surface parameterization schemes. To study the transferability issue, data from a wide range of climatic conditions should be used. The MOPEX project is continuing to assemble data from many different countries. These data should be used to test enhanced a priori parameters.

One of the driving forces behind the progress in parameter estimation research is the increasing array of data sources, including satellite and other advanced observational technologies. With the wealth of new data sources, it is important to investigate the ways to maximize the use of high-resolution spatio-temporal information. Meanwhile, the issue of uncertainty attributed to data errors should be addressed.

Any improvement in parameter estimation procedures must be tied to how we represent the hydrological processes. As our knowledge of these processes advance and as the availability of distributed forcing inputs increase, improved hydrologic models are likely to emerge. This will bring new challenges in terms of parameter estimation and model calibration. Much of the work cited above has already been reported, or is in the process of being reported, by MOPEX project participants. At this moment, there is no consensus about what type of parameter estimation approach is likely to lead to success. One of the advantages of MOPEX is that we keep it open for all different approaches. With a true collaborative spirit by international scientists, enhanced a priori parameter estimation should be available. This in turn should result in improved skill in hydrologic predictions.

Acknowledgements

The authors would like to express our appreciation to Drs Neil McIntyre of London Imperial College and Doug Boyle of Desert Research Institute for their thoughtful review comments that have resulted in tremendous improvement to this paper. We would also like to acknowledge the support by our respective institutions.

References

Abdulla, F.A., Lettenmaier, D.P., Wood, E.F., Smith, J.A., 1996. Application of a macroscale hydrologic model to estimate the water balance of the Arkansas-red river basin. J. Geophys. Res. 101 (D3), 7449–7459.

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.

Beven, K., Binley, A., 1992. The future of distributed hydrological modeling. Hydrol. Process. 6, 189–206.

Carsel, R.T., Parrish, R.S., 1988. Developing joint probability distributions of soil water retention characteristics. Water Resour. Res. 24 (5), 755–769.

Chen, T.H., Henderson-Sellers, A., Milly, P.C.D., Pitman, A.J., Beljaars, A.C.M., Polcher, J., Abramopoulos, F., Boone, A., Chang, S., Chen, F., Dai, Y., Desborough, C.E., Dickinson, R.E., Dumenil, L., Ek, M., Garratt, J.R., Gedney, N., Gusev, Y.M., Kim, J., Koster, R., Kowalczyk, E.A., Laval, K., Lean, J., Lettenmaier, D., Liang, X., Mahfouf, J.-F., Mengelkamp, H.-T., Mitchell, K., Nasonova, O.N., Noilhan, J., Robock, A., Rosenzweig, C., Schaake, J., Schlosser, C.A., Schulz, J.-P., Shmakin, A.B., Verseghy, D.L., Wetzel, P., Wood, E.F., Xue, Y., Yang, Z.-L., Zeng, Q., 1997. Cabauw experimental results from the project for intercomparison of land surface parameterisation schemes. J. Climate 10, 1144–1215.

Clapp, R.B., Hornberger, G.M., 1978. Empirical equations for some soil hydraulic properties. Water Resour. Res. 14 (4), 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.

Dickinson, R.E., Henderson-Sellers, A., Kennedy, P.J., Wilson, M.F., 1986. Biosphere-atmosphere transfer scheme (BATS) for

NCAR Community Climate Model, NCAR/TN-275+STR National Center for Atmospheric Research 1986. 69p.

Dooge, J.C., 1997. Scale problems in hydrology. In: Buras, N. (Ed.), Reflections on Hydrology: Science and Practice. American Geophysical Union, Washington, DC, pp. 84–143.

Duan, Q., 2003. Global optimization for watershed model calibration. In: Duan, Q. et al. (Ed.), Calibration of Watershed Models. AGU Water Sci. & App. 6, Washington, DC, pp. 89–104.

Duan, Q., Sorooshian, S., Gupta, V.K., 1992. Effective and efficient global optimization for conceptual rainfall-runoff models. Water Resour. Res. 28 (4), 1015–1031.

Duan, Q., Sorooshian, S., Gupta, V.K., 1994. Optimal use of the SCE-UA global optimization method for calibrating watershed models. J. Hydrol. 158, 265–284.

Duan, Q., Schaake, J., Koren, V., Chen, F., Mitchell, K., 1995. Testing of a physically-based soil hydrology model: a priori parameters vs tuned parameters. EOS Trans. AGU 76 (46), 92 (Fall Meeting Suppl.).

Duan, Q., Koren, V., Koch, P., Schaake, J., 1996. Use of NDVI and soil characteristics data for regional parameter estimation of hydrologic models. EOS. AGU 77 (17), 138 (Spring Meeting Suppl.).

Duan, Q., Schaake, J., Koren, V., 2001. *A priori* estimation of land surface model parameters. In: Lakshmi, V. et al. (Ed.), Land Surface Hydrology, Meteorology, and Climate: Observations and Modeling Water Science and Application 3. American Geophysical Union, Washington, DC, pp. 77–94.

Duan, Q., Gupta, H., Sorooshian, S., Rousseau, A., Turcotte, R. (Eds.), 2003. Advances in Calibration of Watershed Models, Water Science and Application, Series 6. American Geophysical Union, Washington, DC, p. 345.

Farnsworth, R.K., Peck, E.L., 1982. Evaporation atlas for the contiguous 48 United States, NOAA Tech. Rep. NWS 33, US Dept. of Commerce, Washington DC.

Finnerty, B.D., Smith, M.B., Seo, D.-J., Koren, V., Moglen, G.E., 1997. Space-time scale sensitivity of the Sacramento model to radar-gage precipitation inputs. J. Hydrol. 203, 21–38.

Franks, S.W., Beven, K.J., 1997. Bayesian estimation of uncertainty in land surface-atmosphere flux prediction. J. Geophys. Res. 102 (D23), 23991–23999.

Georgakakos, K.P., Seo, D., Gupta, H.V., Schaake, J., Butts, M.B., 2004. Characterizing streamflow simulation uncertainty through multimodel ensembles. J. Hydrol. 298 (1–4), 222–241.

Gupta, H.V., Bastidas, L.A., Sorooshian, S., Shuttleworth, W.J., Yang, Z.L., 1999. Parameter estimation of a land surface scheme using multi-criteria methods. J. Geophys. Res. 104 (D16), 19491–19504.

Gutman, G., Ignatov, A., 1998. The derivation of the green fraction from NOAA/AVHRR data for use in numerical weather prediction models. Int. J. Remote Sens. 19 (8), 1533–1543.

Hansen, M., Reed, B.C., 2000. A comparison of the IGBP DISCover and university of Maryland global land cover products. Int. J. Remote Sens. 26 (6/7), 1365–1374.

Jackson, C., Xia, Y., Sen, M., Stoffa, P.L., 2003. Optimal parameter and uncertainty estimation of a land surface model: a case study using data from Cabauw, Netherlands. J. Geophys. Res. 108 (D18).

Kistler, R., Kalnay, E., Collins, W., Saha, S., White, G., Woollen, J., Chelliah, M., Ebisuzaki, W., Kanamitsu, M., Kousky, V., van den Dool, H., Jenne, R., Fiorino, M., 2001. The NCEP-NCAR 50-year reanalysis: monthly means CD-ROM and documentation. Bull. Am. Meteorol. Soc. 82 (2), 247–267.

Koren, V.I., Finnerty, B.D., Schaake, J.C., Smith, M.B., Seo, D.-J., Duan, Q.-Y., 1999. Scale dependencies of hydrologic models to spatial variability of rainfall. J. Hydrol. 217, 285–302.

Kuczera, G., Parent, E., 1998. Monte Carlo assessment of parameter uncertainty in conceptual catchment models: th metropolis algorithm. J. Hydrol. 211, 69–85.

Lamb, R., Kay, A.L., 2004. Confidence intervals for a spatially generalized, continuous flood frequency model for Great Britain. Water Resour. Res. 40, W07501.

Liston, G.E., Sud, Y.C., Wood, E.F., 1994. Evaluating GCM land surface hydrology parameterizations by computing river discharge using a runoff routing model: application to the Mississippi River basin. J. Appl. Meteorol. 33 (3), 394–404.

Loveland, T.R., Reed, B.C., Brown, J.F., Ohlen, D.O., Zhu, J., Yang, L., Merchant, J.W., 2000. Development of a global land cover characteristics database and IGBP DISCover from 1-km AVHRR data. Int. J. Remote Sens. 21 (6/7), 1303–1330.

Maurer, E.P., O'Donnell, M.G., Lettenmaier, D.P., Roads, J.O., 2001. Evaluation of the land surface water budget in NCEP/NCAR and NCEP/DOE reanalyses using an off-line hydrologic model. J. Geophys. Res. 106 (D16), 17841–17862.

McIntyre, N.R., Lee, H., Wheater, H.S., Young, A.R., 2004. Tools and approaches for evaluating uncertainty in streamflow predictions in ungauged UK catchments. In: Pahl-Wostl, C., Schmidt, S., Jakeman, T. (Eds.), iEMSs 2004 International Congress: 'Complexity and Integrated Resources Management', June. Int. Env. Modelling and Software Soc., Osnabrueck, Germany.

Merz, R., Bloschl, G., 2004. Regionalisation of catchment model parameters. J. Hydrol. 287, 95–123.

Miller, D.A., White, R.A., 1999. A conterminous united states multi-layer soil characteristics data set for regional climate and hydrology modeling. Earth Interactions, 2, (available at http://www.EarthInteractions.org).

Pitman, A.J., Henderson-Sellers, A., Yang, Z.-L., Abramopoulos, F., Boone, A., Desborough, C.E., Dickinson, R.E., Gedney, N., Koster, R., Kowalczyk, E., Lettenmaier, D., Liang, X., Mahfouf, J.-F., Noilhan, J., Polcher, J., Qu, W., Robock, A., Rosenzweig, C., Schlosser, C., Shmakin, A.B., Smith, J., Suarez, M., Verseghy, D., Wetzel, P., Wood, E., Xue, Y., 1999. Key results and implications from phase 1(c) of the project for inter-comparison of landsurface parameterization schemes. Clim. Dyn. 15, 673–684.

Rawls, W.J., Gish, T.J., Brakesiek, D.L., 1991. Estimating soil water retention from soil physical properties and characteristics. Adv. Soil Sci. 16, 213–234.

Schaake, J.C., Duan, Q., Smith, M., Koren, V., 2000. Criteria to Select Basins for Hydrologic Model Development and Testing.

Preprints, 15thConference on Hydrology, 10–14 January 2000, Amer. Meteor. Soc., Long Beach, CA, Paper P1.8.

Schlosser, C.A., Slater, A., Robock, A., Pitman, A.J., Vinnikov, K.Y., Henderson-Sellers, A., Speranskaya, N.A., Mitchell, K., Boone, A., Braden, H., Cox, P., DeRosney, P., Desborough, C.E., Dai, Y.J., Duan, Q., Entin, J., Etchevers, P., Gedney, N., Gusev, Y.M., Habets, F., Kim, J., Kowalczyk, E., Nasanova, O., Noilhan, J., Polcher, J., Shmakin, A., Smirnova, T., Verseghy, D., Wetzel, P., Xue, Y., Yang, Z.-L., 2000. Simulations of a boreal grassland hydrology at Valdai, Russia: PILPS Phase 2(d). Mon. Weather Rev. 128 (2), 301–321.

Sellers, P.J., Mintz, Y., Sud, Y.C., Dalcher, A., 1986. A simple biosphere model (SiB) for use within general circulation models. J. Atmos. Sci. 43, 505–531.

Sivapalan, M., 2003. Prediction in ungauged basins: a grand challenge for theoretical hydrology (invited commentary). Hydrol. Process. 17 (15), 3163–3170.

Slater, A., Schlosser, C.A., Desborough, C.E., Pitman, A.J., Henderson-Sellers, A., Robock, A., Vinnikov, K.Ya., Speranskaya, N.A., Mitchell, K., Boone, A., Braden, H., Chen, F., Cox, P., de Rosney, P., Dickinson, R.E., Dai, Y.-J., Duan, Q., Entin, J., Etchevers, P., Gedney, N., Gusev, Ye.M., Habets, F., Kim, J., Koren, V., Kowalczyk, E., Nasanova, O.N., Noilhan, J., Shaake, J., Shmakin, A.B., Smirnova, T., Verseghy, D., Wetzel, P., Xue, Y., Yang, Z.-L., 2001. The representation of snow in land-surface schemes: results from PILPS 2(d). J. Hydrometeorol. 2, 7–25.

Vrugt, J.A., Gupta, H.V., Bastidas, L.A., Bouten, W., Sorooshian, S., 2003. Effective and efficient algorithm for multiobjective optimization of hydrologic models. Water Resour. Res. 39 (8), 5.1–5.19.

Wagener, T., McIntyre, N., Lees, M.J., Wheater, H.S., Gupta, H.V., 2003. Towards reduced uncertainty in conceptual rainfall-runoff modelling: dynamic identifiability analysis. Hydrol. Proc. 17 (2), 455–476.

Wagener, T., Wheater, H.S., Gupta, H.V., 2004. Rainfall-Runoff Modelling in Gauged and Ungauged Catchments. Imperial College Press, London, UK. 300 pp.

Wood, E.F., Lettenmaier, D.P., Liang, X., Lohmann, D., 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.-C., 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 Planetary Change 19, 115–135.

Yapo, P., Gupta, H.V., Sorooshian, S., 1998. Multi-objective global optimization for hydrologic models. J. Hydrol. 204, 83–97.