Journal of Hydrology (2008) **353**, 109–128

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jhydrol

Fast and efficient optimization of hydrologic model parameters using a priori estimates and stepwise line search

Vadim Kuzmin [a,*], Dong-Jun Seo [b,c], Victor Koren [b]

[a] *Department of Civil and Environmental Engineering, University of Melbourne, Victoria 3010, Australia*
[b] *Hydrology Laboratory, Office of Hydrologic Development, National Weather Service, 1325 East-West Highway, Silver Spring, MD 20910, USA*
[c] *University Corporation for Atmospheric Research, Visiting Scientist Programs, 3300 Mitchell Lane, Boulder, CO 80301, USA*

Received 22 March 2007; received in revised form 4 October 2007; accepted 4 February 2008

KEYWORDS
Automatic calibration;
Hydrologic model;
A priori estimates;
Parametric uncertainty;
Stepwise line search;
Shuffled complex evolution

Summary Routine availability of hydrological, geological, and other physiographic data today allows us to obtain a priori estimates of hydrologic model parameters prior to explicit model calibration. When informative a priori estimates of model parameters are available, the problem of hydrologic model calibration becomes one of filtering, i.e. improving the a priori estimates based on observations of input and output to and from the hydrologic system, respectively, rather than one of bounded global optimization based solely on the input and output data as in traditional model calibration. Given that global optimization is computationally very expensive and does not, in general, transfer the spatial patterns of soil and land surface characteristics to the model parameters, the filtering approach is particularly appealing for automatic calibration of distributed hydrologic models. Toward that ultimate goal, we explore in this work calibration of a lumped hydrologic model via limited optimization of a priori estimates of the model parameters. The technique developed for the purpose is a simple yet effective and efficient pattern search algorithm called the Stepwise Line Search (SLS). To evaluate the methodology, calibration and validation experiments were performed for 20 basins in the US National Weather Service West Gulf River Forecast Center's (NWS/WGRFC) service area in Texas. We show that SLS locates the posterior parameter estimates very efficiently in the vicinity of the a priori estimates that are comparable, in terms of reducing the objective function value, to those from global minimization. A cross validation experiment indicates that, when parametric

* Corresponding author. Tel.: +61 4 8344 5628; fax: +61 4 8344 4616.
E-mail address: v.kuzmin@civenv.unimelb.edu.au (V. Kuzmin).

doi:10.1016/j.jhydrol.2008.02.001

uncertainty due to lack of calibration data is considered, limited optimization of a priori parameters using SLS may be preferred to global optimization.

Introduction

With the widespread availability of spatial data sets of soil, land cover, land use, and others, it is now possible in many parts of the world to obtain a priori estimates of hydrologic model parameters following some form of re-parameterization of them in terms of pedologic and physiographic data (Koren et al., 2000, 2003; Leavesley et al., 2003; Mertens et al., 2004; Vieux and Moreda, 2003). While such a priori estimates are undoubtedly subject to significant uncertainties due to various sources of error and scale differences in observations and modeling, they reflect, to a varying degree of accuracy, the spatial variability of the parameters, and can provide an informative first-guess estimate for the hydrologic model parameters prior to or in the absence of explicit model calibration. When such a priori estimates are available, the problem of automatic calibration becomes one of filtering, i.e., improving the a priori estimates based on observed data (typically precipitation and streamflow), rather than one of bounded global optimization as in traditional model calibration. While tremendous advances have been made in recent years in estimation of parameters in lumped hydrologic models and assessment of their uncertainty (Beven and Kirkby, 1979; Duan, 2003; Gupta et al., 2003; Koren et al., 2003; Kuczera and Parent, 1998; Schaake et al., 2001), the currently available automatic calibration techniques are generally based on global optimization which requires a very large number of function evaluations. As such, they are not very amenable to estimation of distributed parameters in fine-scale hydrologic modeling. Also, while Monte Carlo-type automatic calibration techniques may be reasonable for models that operate at daily or 6-hourly time steps, they may be computationally too expensive to be operationally viable at 1-hour time step even for lumped models. In addition to computational considerations, automatic calibration based on global optimization as practiced today in lumped modeling does not transfer the spatial patterns of pedologic and physiographic characteristics observed in the spatial data to the hydrologic model parameters very well. For example, automatically-calibrated parameters from global optimization for two adjacent basins in the same area with similar pedologic and physiographic properties may not share similarities that are duly expected from physical considerations. As such, automatic calibration via global optimization is not very conducive to hydrologic modeling over a large area where interdependence of hydrologic model parameters among adjacent basins may be important (Koren et al., 2003). Given the above observations, we argue that some combination of estimation of a priori parameters and 'limited' optimization of them offers a more effective and practical path to estimation of distributed parameters than simply extending global optimization to a distributed parameter setting (Refsgaard, 1997). The overarching objective of this work is to explore such a filtering approach toward development of an operationally viable methodology for routine automatic calibration of distributed hydrologic models. As a first step toward that ultimate goal, we focus in this work on calibration of a lumped hydrologic model. Specifically, we seek a local (as opposed to global) parameter optimization technique that offers the kind of performance and computational efficiency for calibration of lumped models to bring optimization of distributed parameters within the realm of possibility for operational hydrologic forecasting. We do recognize that successful application of such a technique to lumped models may not necessarily translate to that to distributed models, for which different strategies may be pursued (see, e.g., Eckhardt and Arnold, 2001; Madsen, 2003; Vieux et al., 2004; Heuvelmans et al., 2006; Francés et al., 2007). We note here that application of the local optimization technique described in this work to distributed hydrologic models is ongoing and the results will be reported in the near future.

In classical estimation theory, a priori information about the model parameters is treated typically as a penalty term added to the (typically quadratic) objective function (Jazwinski, 1970; Schweppe, 1973):

$$\min \quad J = [Q_o - Q_s(X)]^T P_Q^{-1} [Q_o - Q_s(X)] + [X - X_a]^T R_X^{-1} [X - X_a], \tag{1}$$

$$\text{subject to} \quad X_L \leqslant X \leqslant X_U. \tag{2}$$

In the above, denotes the vector of the observed streamflow, X denotes the vector of the model parameters being estimated, $X = (x_1, \ldots, x_N)$, $Q_s(X)$ denotes the vector of the model-simulated flow corresponding to Q_o, P_Q denotes the error covariance matrix associated with the model-simulated flow, X_a denotes the vector of the a priori estimates of the model parameters, R_X denotes the error covariance matrix associated with X_a, and X_L and X_U denote the vectors of the lower and upper bounds of X, respectively. Throughout this paper, the upper- and lowercase letters denote vector and scalar variables, respectively. Though cast as a least squares problem for simplicity and intuitiveness, the above formulation, or its variant, is essentially equivalent to Bayesian estimation (Misirli et al., 2003) or Kalman filtering (Jazwinski, 1970) under a varying set of assumptions and interpretations. The details of their relationships are not central to the development of this paper and are not given here. The above type of formulation has been widely used in groundwater modeling with gradient-based optimization techniques.

Because $Q_s(X)$ is generally a highly nonlinear function of X and the a priori estimates of X are subject to potentially significant biases due to errors and scale differences in observations and modeling, the above formulation faces at least two serious difficulties in estimation of hydrologic model parameters. The first is that, because of large nonlinearity of $Q_s(X)$, even a small systematic bias or random error in the a priori estimates of X can lead to a very poor solution.

The second is that, because of large nonlinearity, algorithms that are computationally efficient for smooth functions, such as gradient-based minimization, work very poorly even if the a priori estimates are in the right ballpark in the parameter space. For these reasons, the current practice of automatic calibration in hydrologic modeling is dominated by bounded global optimization (Archetti and Schoen, 1984; Brazil, 1988; Duan et al., 1992, 2003; Evtushenko, 1973; Yapo et al., 1997) despite heavy computational burden. In this work, we explore an alternative approach based on informative a priori estimates of the parameters and quasi-local optimization. The qualifier 'quasi' signifies that, due to a fixed discrete step size used (see below for comments), the numerical optimum found may not be collocated with a local minimum closest to a priori parameter set.

This paper is organized as follows. In Section 'Proposed approach', we describe its critical components, estimation of biases in input and unit hydrograph, and local optimization of lumped hydrologic model parameters. In Section 'Study area and data', we describe the used study area and the data. In Section 'Results and evaluation', we describe the experiment design and present the results. The final section summarizes and concludes the paper.

Proposed approach

In light of the above observations, the approach taken here is to start from the a priori parameter estimates, and locate the nearest quasi-local minimum via pattern search. Fig. 1 illustrates a hypothetical search sequence where only two parameters are being optimized. It may be seen that, for the suggested approach to offer advantage over bounded global optimization (e.g. Duan et al., 1992; Vrugt et al., 2003a,b), not only the a priori estimates have to be 'sufficiently informative' but also the search technique has to be very efficient. These points will be made clearer later in this section when the search algorithm is described. It is important to point out in Fig. 1 that neither strategy guarantees the 'true' optimum. Our experience from extensive calibration experiments indicates that the objective function surface usually has numerous small pits, which renders 'global' optimum practically unattainable. A local search algorithm, on the other hand, may fast locate a physically realistic quasi-minimum whose objective function value is competitive with those of other local minima in the larger parameter space.

The hydrologic model used in this work is the Sacramento soil moisture accounting model (SAC, Burnash et al., 1973) used operationally in the US National Weather Service (NWS). The a priori estimates of the SAC parameters are obtained from the soil-based parameterization of Koren et al. (2000, 2003). The methodology relates the SAC parameters with soil's moisture-holding capacities estimated from the STATSGO data (Miller and White, 1999). For further details, the reader is referred to Koren et al. (2000, 2003). The soil-based estimates of the SAC parameters have been used extensively with generally favorable results in various applications: the Distributed Model Intercomparison Project (DMIP, Smith et al., 2004; Reed et al., 2004), the North-American Land Data Assimilation System (NLDAS, Cosgrove et al., 2003; Lohmann et al., 2004; Mitchell et al., 2004), the experimental implementation of variational assimilation of hydrologic and hydrometeorological data at the NWS West Gulf River Forecast Center (WGRFC; Seo et al., 2003a,b), and the field evaluation of the NWS Hydrology Laboratory Research Modeling System (HLRMS; Koren et al., 2004; also known as the Distributed Hydrologic Model, DHM). In Section 'Results and evaluation' of this paper, we assess the quality of the soil-based a priori estimates of the SAC parameters by comparing model simulations based on the soil-based estimates with those generated from operational parameters.

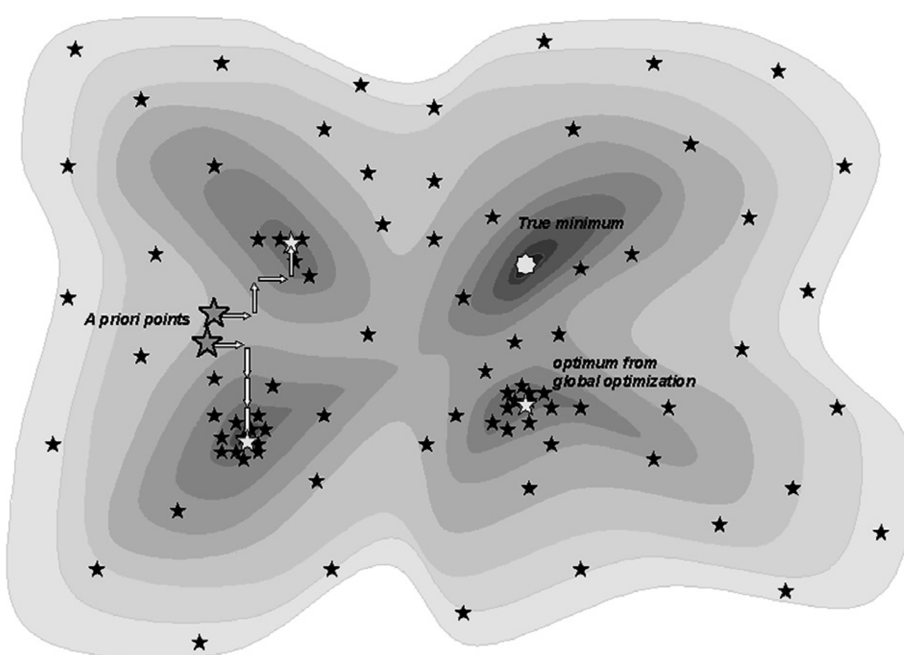

Figure 1 Schematic in 2-D space of global and quasi-local search.

Overview of overall calibration process

Because estimation of hydrologic model parameters depends significantly on the availability and quality of the precipitation and streamflow data and the accuracy of the routing model used, it is necessary to consider bias correction of forcing data and estimation of unit hydrograph to be an integral part of automatic calibration. Below, we describe the three-step process employed in this work of (1) estimating the long-term biases in the precipitation data and potential evaporation climatology, (2) estimating the empirical unit hydrographs (UH) used for routing, and (3) locally optimizing the SAC parameters for soil moisture accounting. The above approach amounts to decomposing the very large problem of simultaneously estimating input biases and parameters for routing and hydrologic models into three separate, smaller estimation problems and solve them sequentially.

With the availability of high-resolution radar-based precipitation data (Hudlow, 1988; Young et al., 2000; Seo and Breidenbach, 2002), the time step for the operational hydrologic models used in NWS is being reduced from 6- to 1-hour for many fast-responding basins. While the radar-based precipitation estimates offer a higher space–time resolution, they are subject to various systematic and random errors that depend significantly on the type of the event and the location of the basin relative to the radar (Wilson and Brandes, 1979; Smith et al., 1996). As such, the magnitude and direction (i.e. over- or underestimation) of the errors associated with radar precipitation estimates vary significantly in space and time. Accordingly, due care must be taken in hydrologic model calibration when radar-based precipitation estimates are used. Also, the unit hydrographs (UH) used operationally at the River Forecast Centers (RFC) typically have a time step of 6 hours. While techniques exist to infer 1-hour UHs from the 6-hour (Chow et al., 1988; Maidment et al., 1996), the resulting 1-hour UHs reflect only those features that can be resolved at the 6-hour time step. Hence, it is much preferable to obtain 1-hour UHs directly from the hourly observations of precipitation and streamflow, particularly for fast-responding basins. Below, we describe in some detail how long-term biases in precipitation data and climatological potential evaporation, and 1-hour empirical UHs were estimated prior to calibration of the soil moisture accounting model.

Bias correction of input data

As noted above, radar precipitation estimates are subject to various sources of systematic and random errors at various scales (Wilson and Brandes, 1979; Smith et al., 1996; Chaubey et al., 1999; Seo et al., 1999; Mazi et al., 2004; McCabe et al., 2005). Because such errors are event- and sampling geometry-specific, they should ideally be corrected for each event and for each basin. Such correction, however, requires water budget analysis on an event-by-event basis, which is not very practical in an operational setting. Here, we correct only for long-term biases in MAPX data and MAPE climatology, where MAPX denotes the mean areal precipitation estimates based on radar and rain gauge data and MAPE denotes the mean areal potential evaporation estimate based on monthly climatology. The resulting adjustment factors to MAPX and MAPE are referred to herein as PXADJ and PEADJ, and denoted by β_P and β_E, respectively.

The adjustment factors to precipitation and potential evaporation were estimated by solving the following inverse problem via the variational assimilation method (Jazwinski, 1970; Li and Navon, 2001; Seo et al., 2003a,b):

$$\min_B \qquad J = \left[\sum_k q_{o,k} - \sum_k q_{s,k}(B, U_a, X_a)\right]^2, \tag{3}$$

$$\text{subject to} \quad S_{k+1} = F(S_k, B, X_a). \tag{4}$$

In Eq. (3), $q_{o,k}$ denotes the observed flow at the kth hour, $q_{s,k}(B, U_a, X_a)$ denotes the simulated flow from SAC and UH at the kth hour, B denotes the multiplicative biases in MAPX and MAPE, β_P and β_E, respectively, U_a and X_a denote the a priori estimates of the UH (i.e. the vector of the a priori UH ordinates) and SAC parameters, respectively, and the summation is over the entire duration of record except the warm-up period of about a year at the beginning of the record. In Eq. (4), S_k denotes the SAC states at hour k, and $F(S_k, B, X_a)$ denotes the soil moisture dynamics of SAC. Note in Eq. (3) that, because we are minimizing the difference in the long-term aggregated flow, only the soil moisture accounting matters and the shape of the UH does not come into play.

Estimation of 1-hour empirical unit hydrograph

To estimate empirical UHs at 1-hour time step from 1-hour MAPX and streamflow data, we solved the following constrained minimization problem via variational assimilation (see also Appendix of Seo et al., 2003a,b):

$$\min_U \qquad J = \sum \left[q_{o,k} - q_{s,k}(B^*, U, X_a)\right]^2, \tag{5}$$

$$\text{subject to} \quad S_{k+1} = F(S_k, B^*, X_a). \tag{6}$$

In the above, the adjustment factor B^* is from the minimization in Eqs. (3) and (4), and U denotes the vector of the ordinates of the UH being estimated. It is worth pointing out that, because routing via UH is a linear operation, the above minimization is linear in U, and as such the quadratic objective function in Eq. (5) yields global optimum. Our experience indicates that solutions from the above minimization are subject to numerical oscillations near the origin and over the right tail end of the estimated UH. As such, some minor adjustment, or trimming, may be necessary to the 'raw' UH in order to produce physically realistic hydrographs. Such adjustment may be made automatically by following a set of simply rules to ensure monotonicity in the tail ends of the UH.

The final solution of U, or U^*, from Eqs. (5) and (6) depends on the choice of the a priori estimates of the SAC parameters, X_a, and, to a lesser extent, on the bias correction factors, B^*. Accordingly, the three-step process of estimating B, U and X described above should ideally be iterated until all solutions converge. Our experience from this and other works indicates that the adjustment factors and UHs are not very sensitive to the choice of the a priori SAC parameters between the soil-based and the RFC-operational, and that for practical purposes additional iteration of the three-step process is generally not necessary.

Local optimization of SAC parameters

Once B^* and U^* obtained, the a priori SAC parameters, X_a, are locally optimized. Below, we describe the optimization procedure and the objective function used.

Stepwise line search

The local search technique developed in this work, referred to herein as the Stepwise Line Search (SLS), is essentially a successive minimization procedure along coordinate directions (Press et al., 1986), but with a fixed step size along each coordinate (Fig. 1). In classification of calibration procedures, SLS may be considered a modification of the pattern search method (Whitehead et al., 1986; Teleb and Azadivar, 1994; Torczon, 1997). Algorithmically, SLS is made of the following steps: (1) start with the a priori estimates of the hydrologic model parameters, (2) with the rest of the parameter estimates fixed to the a priori, increase or decrease the value of the first parameter by one step to the direction of decreasing objective function value, (3) with the first parameter now set to the new (or old, if the objective function value did not decrease) value, decrease or increase the value of the second parameter by one step to the direction of decreasing objective function value, (4) repeat Step 3 until the objective function is minimized with respect to each of all remaining parameters, (5) repeat Steps 2 through 4 until no further reduction in the objective function is realized. Based on extensive testing and analysis of the SAC response function surface for the work reported in this paper, a simplification has been added to the above basic algorithm; if a given parameter value remains the same in three consecutive loops, where a loop represents an iteration of the stepwise line search of all parameters being optimized, the parameter is eliminated from further consideration. This simplification usually results in as much as a fourfold reduction in computational amount. Depending on the basin and the period of record, the set of such non-sensitive parameters varies, a reflection that interdependence among parameters varies in space and time. The elimination of certain parameters in the SLS procedure is based on the local sensitivity of the parameters, which may vary considerably within the larger search space for highly nonlinear models. As such, it is possible that the parameters that are eliminated at certain stages in the iteration may become sensitive as the search progresses, an aspect that is not accounted for in SLS currently.

If the step size used is too large, SLS may miss a minimum. Also, because the shape of the objective function within the one-step neighborhood of the parameter space may not be monotonic, the order in which the parameters are searched may affect the outcome. These and other issues are examined via sensitivity analyses and summarized in Section 'Results and evaluation'. The use of the SLS procedure, considered a ''dumb'' method for minimization of smooth functions (Press et al., 1986), reflects the highly nonlinear and irregular nature of the objective function surface with respect to the hydrologic model parameters. Its main benefits are physically realistic posterior estimates, algorithmic simplicity, and computational efficiency.

Multi-scale objective function (MSOF)

One of the most important aspects of manual calibration of hydrologic models is that human beings are very good at assessing the quality of model simulation at many different temporal scales of aggregation simultaneously (Seibert, 2000; Smith et al., 2003; Brazil, 1988). To emulate this multi-scale nature of manual calibration, we employ in this work an objective function composed of contributions from a wide range of time scales of aggregation (Madsen, 2000). The rationale behind the approach is similar to Parada et al. (2003) and is not elaborated here. The particular objective function used in this work has the following form:

$$J = \sqrt{\sum_{k=1}^{n} \left(\frac{\sigma_1}{\sigma_k}\right)^2 \sum_{i=1}^{m_k} (q_{o,k,i} - q_{s,k,i}(X))^2}, \tag{7}$$

where $q_{o,k,i}$ and $q_{s,k,i}$ denote the observed and simulated flows averaged over time interval (i.e. the kth aggregation scale), σ_k denotes the standard deviation of discharge at that scale, n denotes the total number of scales used, and m_k is the number of ordinates at the scale k. In this work, we used hourly, daily, weekly and monthly scales corresponding to $k = 1$, 2, 3 and 4, respectively. Note in Eq. (7) that the weight associated with each term is given by the inverse of the standard deviation of the flow at the respective scales. This weighting scheme assumes that the uncertainty in modeled streamflow at each scale is proportional to the variability of the observed flow at that scale. Another important motivation for using the multi-scale objective function (MSOF) is that it smoothes the objective function surface, and hence reduces the likelihood of the search getting stuck in tiny 'pits.' These aspects of MSOF are further described in Section 'Results and evaluation'.

Study area and data

To evaluate the automatic calibration procedure described above, a series of experiments were carried out for 20 headwater basins in the NWS West Gulf River Forecast Center's (NWS/WGRFC) service area (see Fig. 2). The basins, initially chosen for experimental implementation and evaluation of the variational data assimilation procedure (Seo et al., 2003a,b), are scattered over a large area that covers diverse hydrology and hydroclimatology. Table 1 shows the basin identifiers, time to peak, area, mean hourly flow, mean annual precipitation, and estimates of the impervious fractional area (Seann Reed, personal communication).

The hydrologic model used is the Sacramento model driven by 1-hourly radar-based mean areal precipitation (MAPX) and the climatological mean areal potential evaporation (MAPE) as adjusted by the Normalized Vegetation Difference Index (NDVI) climatology (Koren et al., 1998). The period of record for MAPX data is from January 1995 through December 2002, of which the first year were used for model warm-up. Observed hourly discharges at the basin outlets were used in calibration and evaluation.

Results and evaluation

Estimation of long-term biases in input data using Eqs. (3) and (4) assumes that SAC closes water balance in the long-term completely (i.e. all sources and sinks are accounted for). It is known, however, that the study area

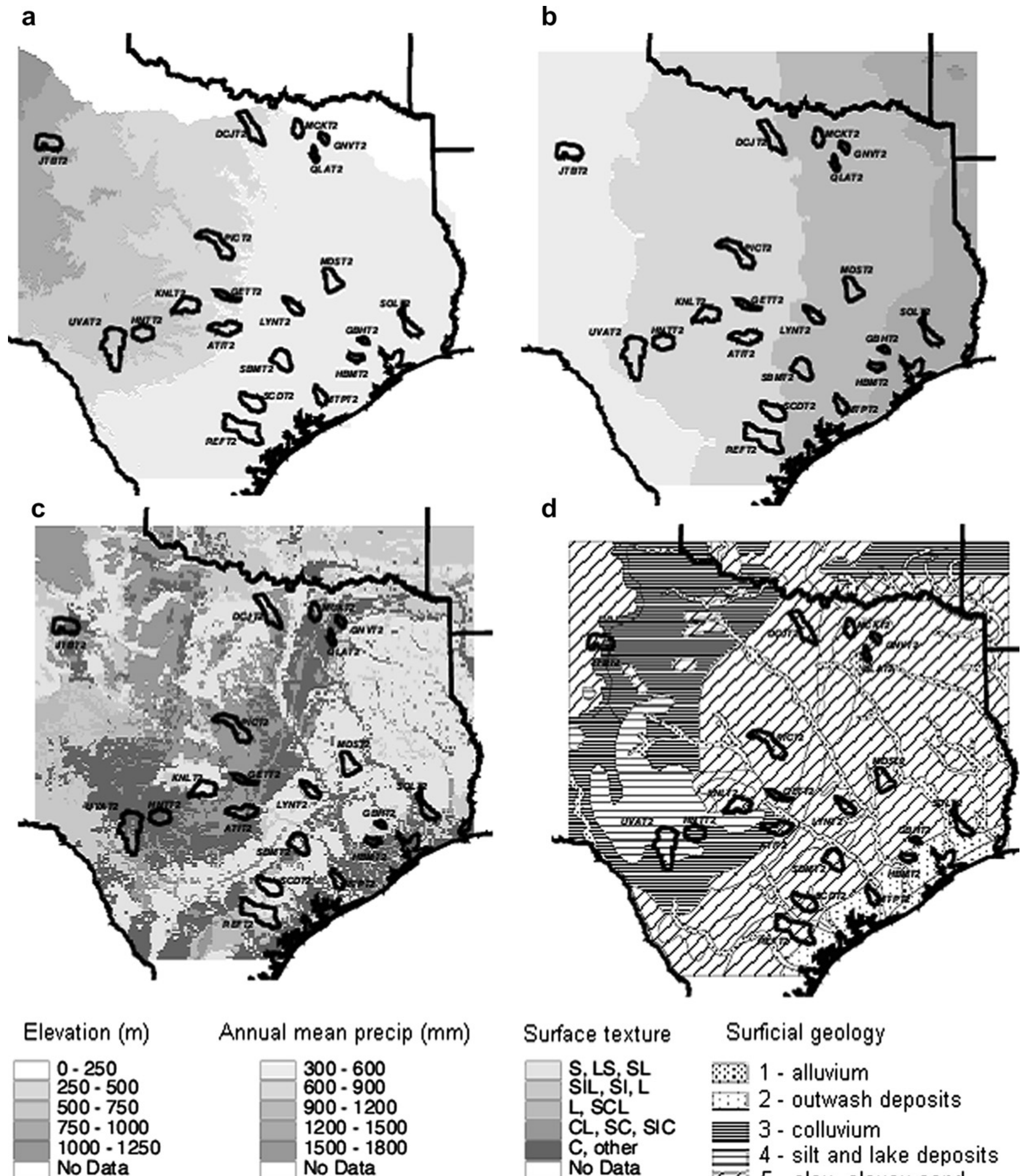

Figure 2 (a) Elevation; (b) precipitation; (c) soil surface texture; (d) surficial geology of the study area in the WGRFC's service area. Also shown are the boundaries of the study basins.

is geologically diverse (see Fig. 2d) and includes karst formations, which are not modeled by SAC. Table 2 shows the adjustment factors, β_P and β_E, for all basins studied in this work as estimated from the bias estimation procedure described above. An adjustment factor of β_P greater or less than unity indicates an under- or overestimation in precipitation amounts in MAPX, respectively. Note in Table 2 that the adjustment factors for MAPE are very close to unity, suggesting that MAPE is relatively bias-free. Those for MAPX, on the other hand, show a larger basin-to-basin variability around unity. Given that soil moisture balance is probably not completely closed by SAC for some basins, significant deviations from unity in β_P may not be completely attributable to biases in MAPX alone. We are currently comparing the MAPX data with estimates of mean areal precipitation based on rain gauge data alone (referred to as MAP) so that the magnitude and direction of the biases in MAPX may be ascertained.

Table 1 Geomorphological and hydroclimatological characteristics of the study basins in WGRFC

No.	Basin ID	USGS ID	Time-to -peak, h	Area, km^2	Average annual discharge, m^3/s	Average annual precipitation, mm	Impervious area, %
1	ATIT2	08159000	9	844	2.45	684	5.8
2	DCJT2	08053500	6	1039	2.38	684	–
3	GBHT2	08076000	7	137	2.58	640	28.8
4	GETT2	08104900	10	334	3.59	1175	1.3
5	GNVT2	08017200	16	212	1.47	605	7.6
6	HBMT2	08075000	4	246	2.28	982	44.1
7	HNTT2	08165500	3	769	8.69	999	–
8	JTBT2	08079600	8	945	1.93	640	–
9	KNLT2	08152000	7	904	1.28	491	0.3
10	LYNT2	08110100	19	508	2.08	868	2.1
11	MCKT2	08058900	13	427	1.63	806	1.6
12	MDST2	08065800	24	870	4.05	859	10.7
13	MTPT2	08162600	25	435	4.90	877	5.4
14	PICT2	08101000	6	1178	4.69	763	0.5
15	QLAT2	08017300	10	197	4.15	701	5.7
16	REFT2	08189500	36	1787	2.41	929	1.3
17	SBMT2	08164300	14	896	4.79	763	1.4
18	SCDT2	08176900	18	932	5.08	763	0.9
19	SOLT2	08041700	82	1746	2.82	763	–
20	UVAT2	08190000	2	1981	13.1	1315	4.8

Table 2 Multiplicative biases for MAPX and MAPE

No.	Basin ID	Basin name	β_P^{SOIL}	β_E^{SOIL}	β_P^{OPER}	β_E^{OPER}
1	ATIT2	Austin – Onion C	0.91	1.03	0.98	1.01
2	DCJT2	Justin – Denton C	0.82	1.07	0.84	1.07
3	GBHT2	Houston – Greens Bayou	1.06	0.97	1.04	0.99
4	GETT2	Georgetown – S Fk San Gabriel	1.02	0.99	1.14	0.93
5	GNVT2	Greenville – Cowleech C	1.08	0.97	1.09	0.96
6	HBMT2	Houston – Brays Bayou	1.53	0.67	1.51	0.64
7	HNTT2	Hunt – Guadalupe R	0.73	1.03	n/a	n/a
8	JTBT2	Justiceburg – Double Mt Fork	1.17	0.95	1.24	0.94
9	KNLT2	Kingsland – Sandy C	0.90	1.05	n/a	n/a
10	LYNT2	Lyons – Davidson C	0.90	1.04	0.98	1.01
11	MCKT2	McKinney – East Fork Trinity	1.07	0.94	1.10	0.92
12	MDST2	Madisonville – Bedias C	1.04	0.98	1.08	0.96
13	MTPT2	Midfield – Tres Palacios	1.18	0.84	1.13	0.88
14	PICT2	Pidcoke – Cowhouse C	0.80	1.09	0.88	1.07
15	QLAT2	Quinlan – South Fork Sabine	1.19	0.91	1.20	0.89
16	REFT2	Refugio – Mission R	0.85	1.06	0.93	1.04
17	SBMT2	Sublime – Navidad R	1.02	0.99	1.07	0.94
18	SCDT2	Schroeder – Coleto C	0.88	1.02	0.96	1.01
19	SOLT2	Sour Lake – Pine Island B	0.99	1.27	0.96	1.12
20	UVAT2	Uvalde – Nueces	0.71	1.06	0.82	1.05
Average			1.01	0.99	1.05[a]	0.97[a]

[a] Excludes HNTT2 and KNLT2.

In the semi-arid western part of the study area, significant events occur rather infrequently. Hence, the warm-up period may not be of sufficient length for some basins. Due to the shortness of the radar-based precipitation data, independent validation over a long period was not possible. Instead, here we carried out a cross validation experiment using five different combinations of 4- and 1-year dependent and independent data sets, respectively. The calibration and validation experiments were designed to help answer the following questions: (1) How do the a priori soil-based estimates of the SAC parameters compare with the parameter estimates used operationally? (2) How much can the a priori estimates be improved by optimization? (3) How does SLS compare with the Shuffled Complex Evolution algorithm (SCE, Duan et al., 1992; Duan, 2003) in terms of the quality of optimized parameters, the rate of convergence, and the number of function evaluations needed? (4) How much do uncertainties in the a priori parameter estimates affect the quality of the posterior estimates? (5) What is the optimum step size for SLS?

Because estimation of PXADJ and PEADJ depends on the SAC parameters and estimation of UH depends on the adjustment factors as well as the SAC parameters, the final solution for the SAC parameters from the 3-step estimation process described in Section 'Proposed approach' depends on the choice of the starting (i.e. first-guess) SAC parameter estimates. To assess the sensitivity to the initial parameter estimates and to evaluate the goodness of the soil-based a priori estimates, we used two different sets of a priori SAC parameter estimates in this work, the soil-based and the 6-hourly operational. They are referred to herein as SOIL and OPER, respectively. The 6-hourly operational, or initial OPER, represents the SAC parameter estimates that have been developed over the years at the RFC and are currently used operationally for model runs at 6-hour time step. Given each set of parameter estimates, the correction factors to MAPX and MAPE and the empirical UHs were estimated as described in Section 'Study area and data', and eleven SAC parameters (see Table 3) were optimized using SLS. For comparison, bounded global optimization of the eleven SAC parameters using the SCE algorithm was also carried out. The upper and lower bounds used for SCE are based on Koren et al. (2003) and are shown in Table 3.

Calibration results

Table 2 shows the long-term biases in MAPX and MAPE as estimated from the procedure described in Section 'Proposed approach'. The bias estimation procedure assumes that all sources and sinks are accounted for by the hydrologic model. As noted in this section, however, parts of the study area are characterized by karst geology. Also, a number of basins are known to have significantly large impervious areas (Elvidge et al., 2004), a piece of information that was not available at the time of this work. Hence, it is acknowledged that some of the bias estimates shown in

Table 3 Calibrated SAC-SMA model parameters and their feasible ranges (reproduced from Koren et al., 2003)

No	Parameter	Description	Ranges
1	UZTWM	The upper layer tension water capacity, mm	10–300
2	UZFWM	The upper layer free water capacity, mm	5–150
3	UZK	Interflow depletion rate from the upper layer free water storage, day^{-1}	0.10–0.75
4	ZPERC	Ratio of maximum and minimum percolation rates	5–350
5	REXP	Shape parameter of the percolation curve	1–5
6	LZTWM	The lower layer tension water capacity, mm	10–500
7	LZFSM	The lower layer supplemental free water capacity, mm	5–400
8	LZFPM	The lower layer primary free water capacity, mm	10–1000
9	LZSK	Depletion rate of the lower layer supplemental free water storage, day^{-1}	0.01– 0.35
10	LZPK	Depletion rate of the lower layer primary free water storage, day^{-1}	0.001–0.05
11	PFREE	Percolation fraction that goes directly to the lower layer free water storages	0.0–0.8

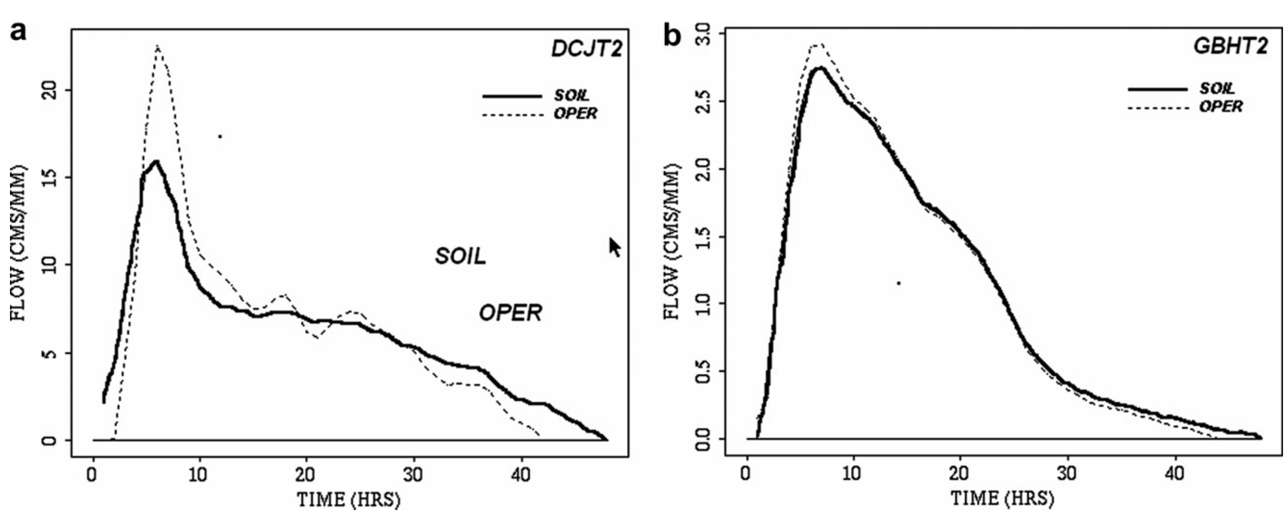

Figure 3 Example empirical unit hydrographs for six of the study basins.

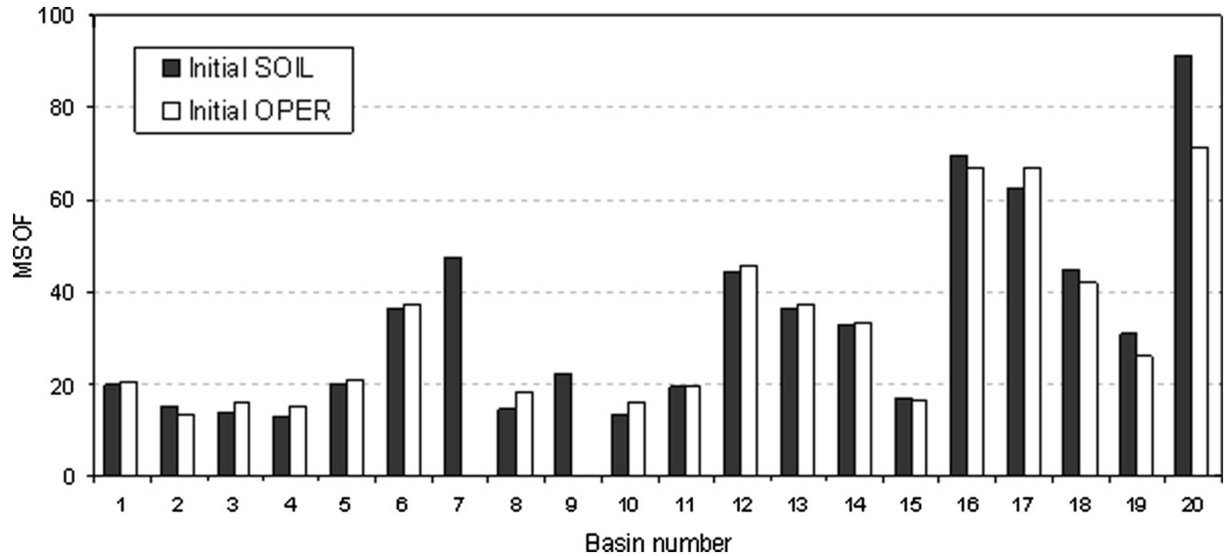

Figure 4 Multi-scale objective function values associated with a priori SOIL (solid bar) and initial OPER (empty bar) parameters. See Table 2 for basin numbers and names.

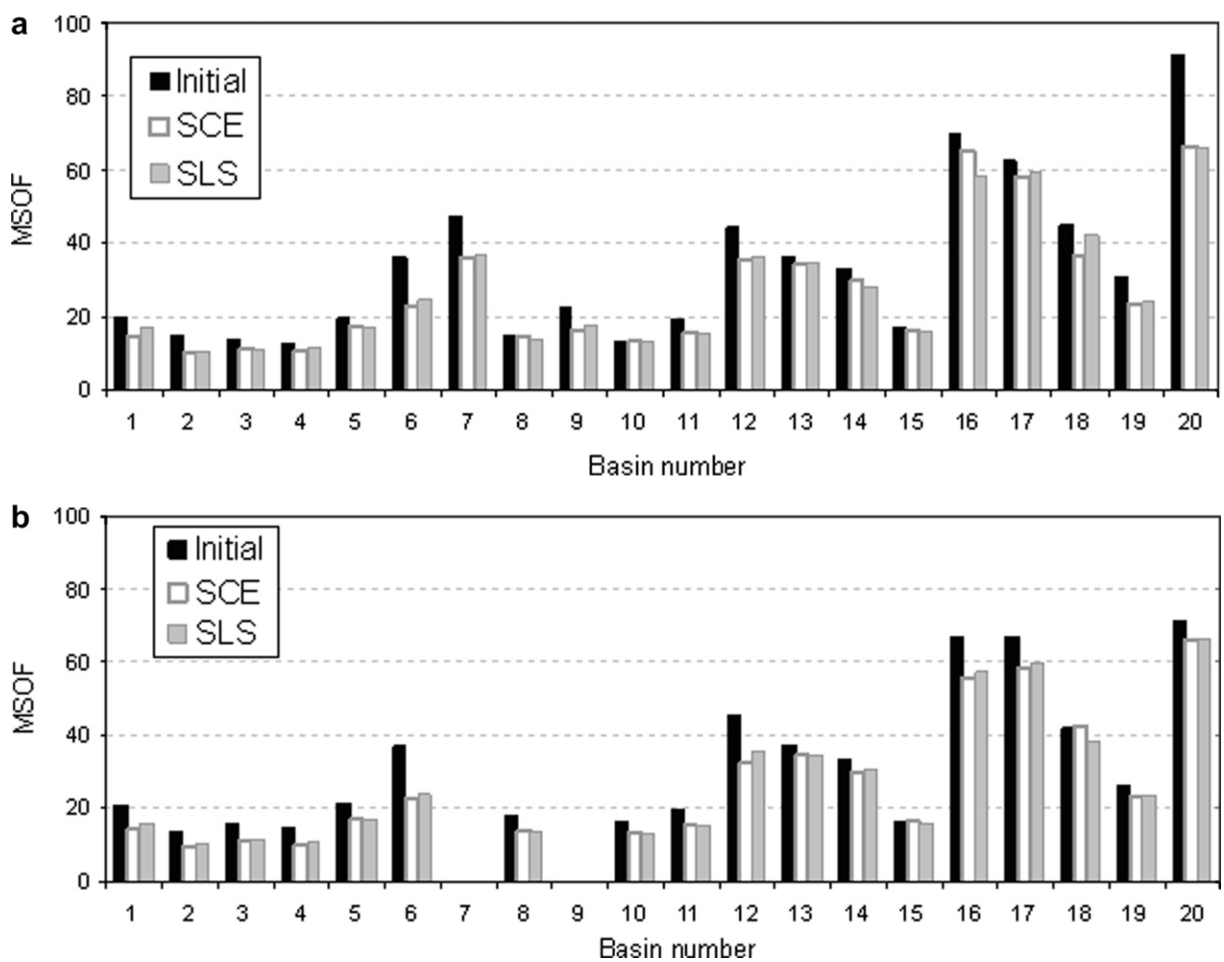

Figure 5 Multi-scale objective function values associated (a) with a priori and optimized SOIL parameters and (b) with initial and optimized OPER parameters. The optimization is by SLS and SCE. See Table 2 for basin numbers and names.

Table 2 may be subject to larger uncertainties than others. We may summarize Table 2 as follows: (1) the magnitude of the biases in MAPX or MAPE is modestly sensitive to the choice of the a priori SAC parameters in that $\beta_{\mathrm{P}}^{\mathrm{SOIL}}$ and $\beta_{\mathrm{P}}^{\mathrm{OPER}}$ as well as $\beta_{\mathrm{E}}^{\mathrm{SOIL}}$ and $\beta_{\mathrm{E}}^{\mathrm{OPER}}$ differ from each other only marginally, (2) for a given basin, the MAPX data may have a significant bias, and (3) the bias in the NDVI-adjusted MAPE climatology is relatively small.

Fig. 3 shows examples of the empirical UHs estimated from the procedure described in Section 'Proposed approach'. They are based on the bias-corrected MAPX data and MAPE climatology. The results for both SOIL and OPER are shown in the figure. It is seen that UH is modestly sensitive to the choice of the a priori SAC parameters. Perhaps the most notable feature in the UH's shown in Fig. 3 is that they do not, in general, fit very well widely used functional approximations for UH (e.g., two-parameter Gamma, Maidment et al., 1996), an indication that some basins studied in this work may not be appropriate for lumped modeling. The above results reinforce the importance of bias correction in the forcing data, accurate routing, and assessment of interdependence among all major sources of error that may impact calibration of the hydrologic model.

The multi-scale objective function values associated with the a priori SOIL and the initial OPER parameter estimates are shown in Fig. 4 for all basins. The figure indicates that the a priori soil-based estimates of the SAC parameters are of comparable quality, in terms of hydrograph simulation at the basin outlet, to the operationally-used estimates, and suggests that the soil-based estimates offer a very good initial guess for the SAC parameters. Fig. 5a shows the multi-scale objective function values associated with the a priori SOIL and optimized SOIL parameter estimates. Similarly, Fig. 5b shows the multi-scale objective function values associated with the initial and optimized OPER. In both figures, the optimization is carried out by both SLS and SCE (for SCE, the following parameter values were used maximum number of function evaluation $N_{FE,max} = 99,999$, number of complexes $p = 13$). The figures indicate that both the a priori SOIL and the initial OPER can be significantly improved by optimization and that, overall, the improvement by SLS is almost as good as that by SCE. That the objective function value for SCE is not always smaller that for SLS is a reflection of the quasi-global nature of the optimization by SCE (Fig. 1). Figs. 6a and b show examples of the simulated hydrographs based on various SAC parameter estimates optimized by SCE and SLS as compared with the observed. Overall, the simulated hydrographs based on the optimized parameter estimates are significantly better than those from the a priori SOIL or the initial OPER. The differences between SLS- and SCE-calibrated hydrographs, on the other hand, are rather minor.

Fig. 7a shows the multi-scale objective function value as a function of the number of function evaluations from the SLS and SCE iterations. Because SLS is a quasi-local optimizer with an informative prior whereas SCE is a quasi-glo-

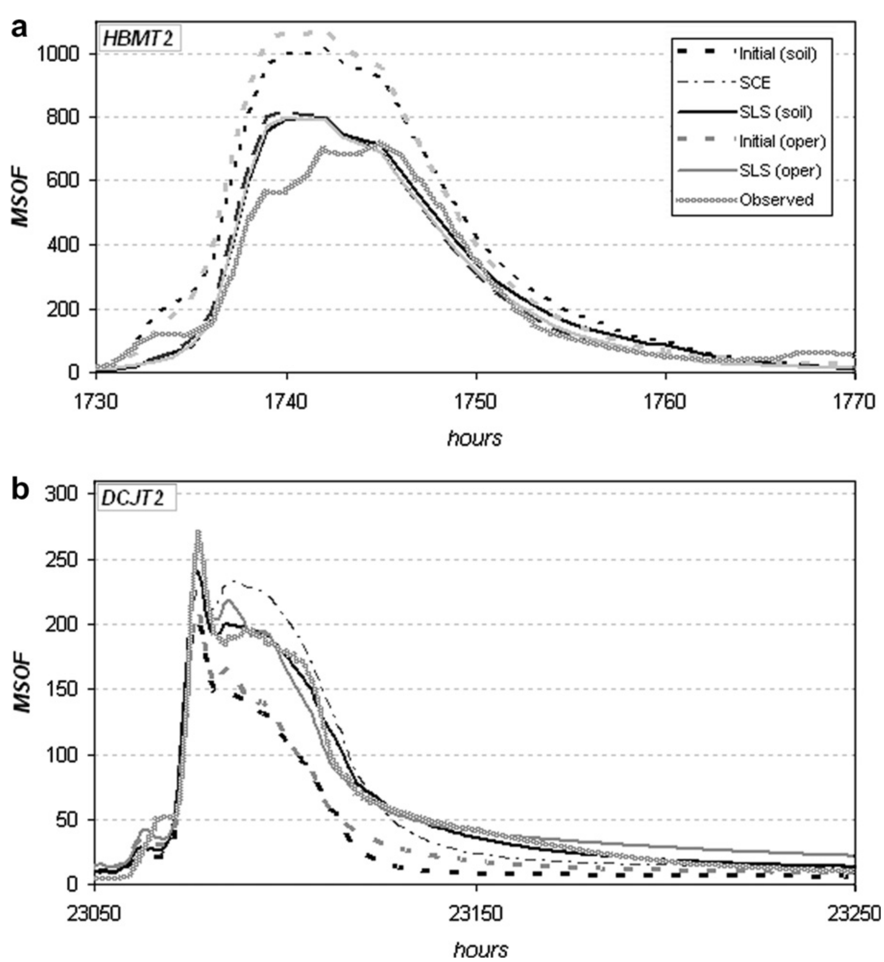

Figure 6 Comparison of observed hydrographs and hydrograph simulated based on the SCE-optimized parameters, the a priori and SLS-optimized SOIL parameters and the initial and SLS-optimized OPER parameters. The results are for HBMT2 (a) and DCJT2 (b).

bal optimizer with a diffuse prior (in that only the physically reasonable bounds are known), the two scatter plots in the figure may not be compared directly against each other. Here, they are shown together to illustrate the informativeness of the a priori SOIL estimates and the effectiveness of SLS. That the a priori soil-based estimates are of high quality may be seen from the objective function value associated with SLS at the origin of the plot, relative to the range of objective function values associated with SCE near the origin (see also Table 4). The computational economy of SLS is also clearly seen. The figure shows that SLS is extremely efficient; the rate of convergence is at least as fast as the fastest by SCE at any point in the iteration. As in any filtering, the premise behind the proposed approach is that, in the absence of strong observational evidence, the posterior parameter estimates should not deviate much from the a priori (Jazwinski, 1970). Fig. 7b shows the distance between the initial parameter estimates and the parameter values being optimized as a function of the number of function evaluations, where the distance is defined as the Euclidean norm of the vector connecting the a priori and the optimized points in the normalized parameter space. Figs. 7a and b indicate that the final parameter estimates from SLS remain much closer to the a priori estimates (0.36 for SLS vs. 1.78 for SCE), while the multi-scale objective function values are very similar (23.13 for SLS vs. 21.83 for SCE). Similar performance measures were used by Mertens et al. (2004) in their investigation of the trade-offs between goodness-of-fit and deviation from the a priori parameter values using SCE.

Fig. 8 shows the contributions of the individual terms in the multi-scale objective function in Eq. (3). The optimization is based on SLS and SCE, and the a priori estimates used for SLS is SOIL. For SCE, the choice of a priori estimates is irrelevant. The figure indicates that the rate of reduction in the objective function value generally differs among different time scales of aggregation. At the beginning of the search, the long-term components, corresponding to 10- and 30-day aggregation scales, contribute more to the overall improvement than the short-term. Later in the optimization, however, short-term components, 1-hour and 1-day scales, become dominant in the optimization. This behavior, observed for most of the basins studied in this work, partially mimics the manual calibration process that generally addresses low flow components first (Smith et al., 2003).

Fig. 9 shows the scale-specific root mean square error (RMSE) of simulated flow normalized by the standard deviation of observed flow at that scale. It shows how the residual errors in the simulated flow are decomposed into scale-specific contributions under the MSOF criterion. The results shown are for three basins, LYNT2, MCKT2 and ATIT2. There are three sets of curves of scale-specific normalized RMSE (NRMSE): (1) all four scales were included in the objective function (NRMSE4); (2) only the hourly and monthly scales were included (NRMSE2), and (3) only the hourly scale was

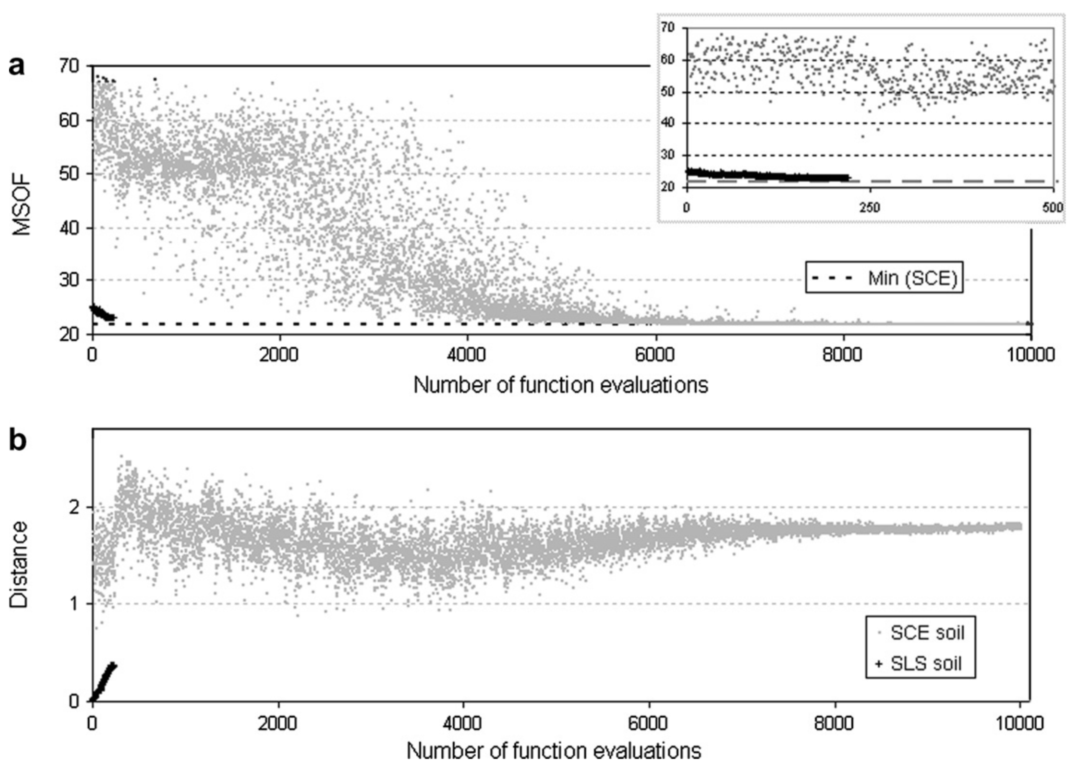

Figure 7 (a) Multi-scale objective function value vs. the number of function evaluations in the optimization sequence. The upper-right plot is a zoom-up near the origin. (b) Distance from the a priori parameters vs. the number of function evaluations in the optimization sequence. The light and dark markers denote the multi-scale objective function value and the distance associated with the SCE- and SLS-optimized parameters, respectively. The results are for ATIT2. The dotted line marked 'Min (SCE)' corresponds to the minimum from SCE.

Table 4 Comparison of search characteristics of SCE and SLS for the calibration period

Basin	Initial MSOF	SCE			SLS			Percent difference in final MSOF	Δ, %	$\frac{K_{SCE}}{K_{SLS}}$
		Final MSOF	Normalized distance D_{SCE}	Number of function evaluations K_{SCE}	Final MSOF	Normalized distance D_{SLS}	Number of function evaluations K_{SLS}			
1	2	3	4[a]	5	6	7[a]	8	9[b]	10[c]	11
ATIT2	23.21	**19.36**[d]	1.59	17,089	20.84	0.90	174	7.6	43.4	98.2
DCJT2	18.47	**16.13**	1.87	30,603	16.57	0.51	202	2.7	72.7	151.5
GBHT2	13.82	**11.35**	1.83	12,798	11.65	0.55	191	2.6	69.9	67.0
GETT2	17.39	**16.22**	1.39	18,194	16.52	0.57	176	1.8	59.0	103.4
GNVT2	16.89	**14.39**	2.30	12,669	14.60	1.23	292	1.5	46.5	43.4
HBMT2	35.69	**27.02**	2.09	16,450	28.52	0.93	234	5.6	55.5	70.3
HNTT2	39.50	**30.99**	1.87	8873	31.01	1.01	313	0.1	46.0	28.3
JTBT2	13.73	**12.19**	1.21	37,083	12.86	0.20	121	5.5	83.5	306.5
KNLT2	18.38	**11.55**	2.01	16,597	13.67	0.83	191	18.4	58.7	86.9
LYNT2	10.51	**10.22**	1.97	7604	10.37	0.32	124	1.5	83.8	61.3
MCKT2	16.84	**13.87**	2.18	12,288	14.18	1.03	315	2.2	52.8	39.0
MDST2	33.92	**25.79**	2.26	17,575	28.56	0.80	216	10.7	64.6	81.4
MTPT2	35.43	33.92	2.19	7589	**33.83**	0.96	283	**−0.3**	56.2	26.8
PICT2	38.99	38.00	1.50	10,826	**37.68**	0.28	129	**−0.8**	81.3	83.9
QLAT2	15.63	**14.38**	2.09	11,540	14.67	0.45	181	2.0	78.5	63.8
REFT2	55.13	**44.95**	2.28	18,978	46.66	1.02	326	3.8	55.3	58.2
SBMT2	56.66	**53.92**	1.35	14,317	54.57	0.78	180	1.2	42.2	79.5
SCDT2	49.03	**47.27**	2.42	14,548	47.55	0.74	185	0.6	69.4	78.6
SOLT2	29.02	**25.04**	1.66	13,045	25.48	0.52	222	1.8	68.7	58.8
UVAT2	107.27	**53.66**	2.37	7271	53.71	1.29	306	0.1	45.6	23.8

[a] The normalized distance is defined as the Euclidean norm of the vector connecting the a priori and the optimized points in the normalized parameter space.
[b] Percent difference in final is defined as [(MSOF (SLS) − MSOF (SCE))/MSOF (SCE) * 100%.
[c] Relative departure from the initial point $\Delta = (D_{SCE} - D_{SLS}/D_{SCE})$*100%.
[d] The smaller MSOF between SCE and SLS is in bold.

included (NRMSE1). Note that NRMSE4 is generally smaller than NRMSE2. Note also that, while NRMSE1 is smaller than NRMSE4 at the hourly scale, it is larger than NRMSE beyond the hourly scale. These observations confirm the expected scale dependence of calibration errors under the MSOF criterion, and support the use of MSOF in automatic calibration of hydrologic models.

Another important aspect of MSOF is that, owing to the contributions from a wide range of frequencies of flow, MSOF helps smooth (i.e. regularize) the objective function surface, thus enabling gradient-based search to be more effective. Fig. 10a shows an example of the objective function value as a function of the two SAC parameters, UZTWM and UZFWM (see Table 3 for explanation), in which the objective function is made only of hourly and monthly contributions. Figs. 10b and c show the partial derivatives of the objective function with respect to UZTWM and UZFWM, respectively. Figs. 10d–f are the same as Fig. 10a–c, respectively, except that they are based on the MSOF composed of contributions from all four scales (i.e. hourly, daily, weekly and monthly). Note that the MSOF in Fig. 10d is smoother than the objective function surface in Fig. 10a, and that the gradients in Figs. 10e and f are generally smaller and less spatially complex (i.e. in the parameter space) than those in Fig. 10b and c, respectively. The significance of a less complex gradient surface is that the zero-gradient isolines in Figs. 10b and c are significantly shorter in length than those in Fig. 10e and f, respectively. Hence, on average, there is a smaller chance of being wrongly attracted to an area of zero gradient even though the area may in fact be distant from a substantial minimum.

Validation results

All results presented above are based on the same data sets used to estimate the parameters themselves. Due to the shortness of the radar-based precipitation data, true validation over a long period could not be performed. As an alternative, we carried out a cross validation experiment in which each year in the five-year period (July 1998 through July 2003) was withheld from calibration and was used only for validation. The results are then plotted in the form of a wind rose where wind directions are replaced by the eleven SAC parameters (see Table 3) and the wind speed is replaced by the optimized parameter estimates normalized by the feasible region of the parameters. The parameter estimates at the outer- and innermost circles in the rose thus correspond to the maximum and the minimum shown in Table 3, respectively. The resulting plot is referred to herein as the 'Parameter Rose.' Figs. 11a and b show the

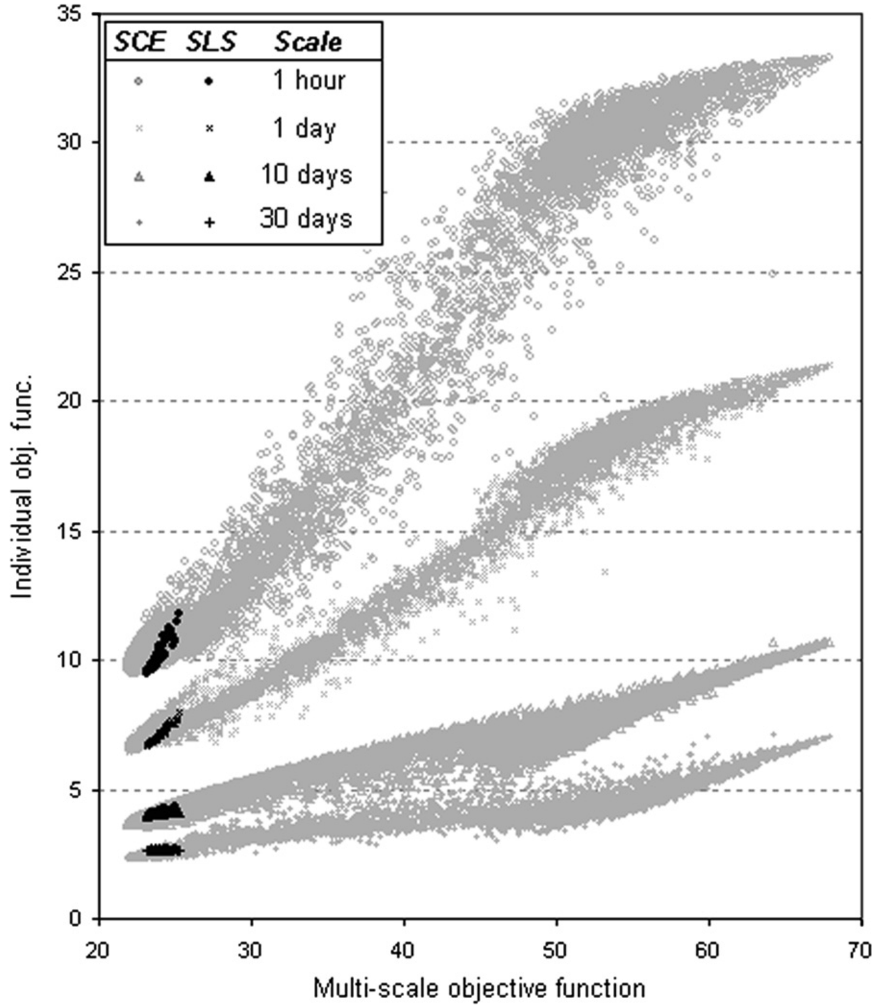

Figure 8 Decomposition of the multi-scale objective function into scale-specific contributions over the course of a parameter optimization run by SCE (light markers) and SLS (dark markers) for ATIT2.

Parameter Roses for two basins, ATIT2 and GNVT2, respectively. In each figure, the black solid line represents the soil-based a priori estimates of the SAC parameters. The different markers represent the SLS-optimized (open markers) and the SCE-optimized (shaded markers) parameter estimates obtained from m different calibration periods, each with $(m - 1)$ years in duration (m was equal to 4 or 5 years depending on the available period of observation for the basin). The two outer and inner roses connect the largest and the smallest parameter estimates among the four or five optimized parameter sets. Hence, the area between the two roses reflects the parametric uncertainty due to the uncertainty in the data (for stationary time series, the two roses would coincide if the period of record is sufficiently large). The multi-scale objective function values from SCE and SLS associated with the different calibration periods each with $(m - 1)$ years in duration are also plotted in the figure. Similarly, the four multi-scale objective function values associated with the m independent periods each with one year in duration are shown as well. Figs. 11a and b may be summarized as follows. The SCE-optimized parameter estimates usually span the entire parameter space whereas the SLS-optimized parameters tend to stay well within the feasible region of the parameter space and form relatively tight bounds. It is interesting to note that, for semi-arid ATIT2 and GETT2, the SLS-optimized parameter estimates form particularly tight bounds and lie very closely to the a priori estimates. In general, the performance of SCE is only marginally better than that of SLS in reducing the objective function value in the calibration (i.e. dependent) periods. In the validation (i.e. independent) periods, however, the performance of SLS is comparable to or better than (for 65% of the cases) that of SCE. It suggests that, even if one discounts the physical realism of the optimized parameter estimates, the quasi-local optimization of a priori parameter estimates offers significant advantages over 'global' optimization given the uncertainty inherent in the input data alone.

Sensitivity to the quality of a priori estimates

As described above, SLS is a quasi-local optimization procedure for locating minima in the vicinity of the a priori parameter estimates. Hence, unlike SCE, calibration results from SLS depend significantly on the quality of the a priori parameter estimates. To assess the effect of

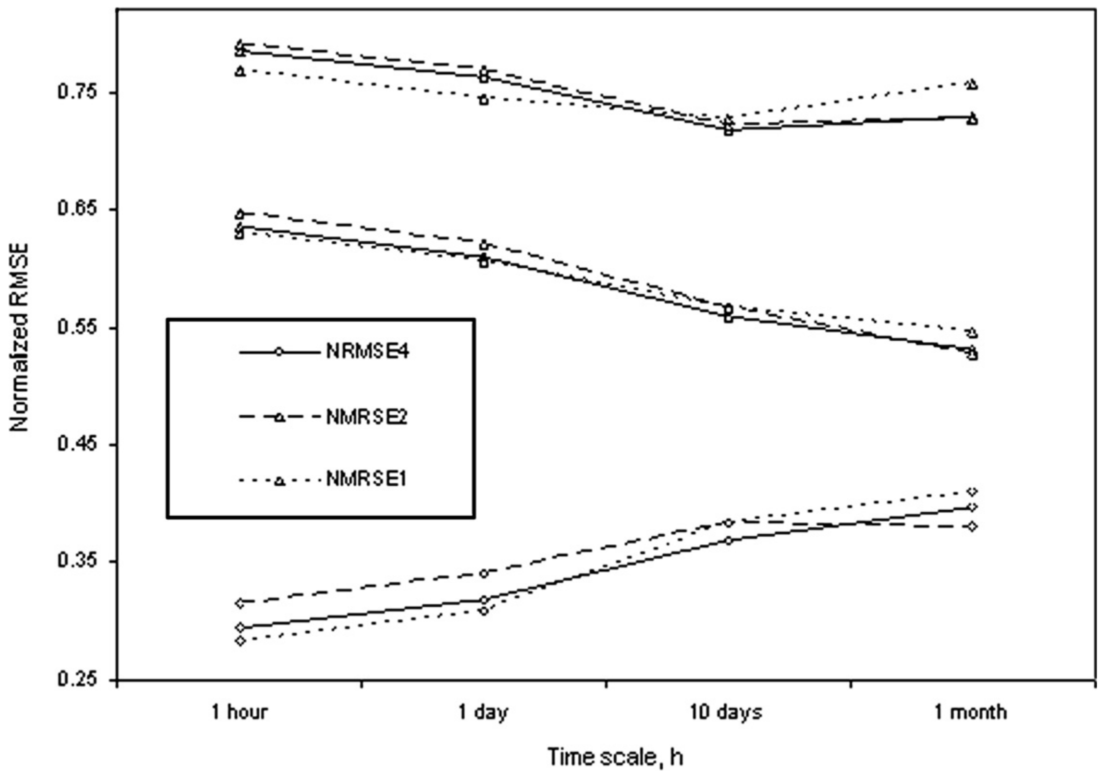

Figure 9 Scale-specific normalized RMSE under a varying number of time scales of aggregation in the objective function. NRMSE1, NRMSE2 and NRMSE4 show the normalized RMSEs when the objective function is made of one, two and four scale-specific error terms, respectively. The results are for LYNT2, MCKT2, ATIT2 (from the top to the bottom).

uncertainty in the a priori parameter estimates on the optimization results, we carried out a series of sensitivity analysis as described below. To degrade artificially the quality of the soil-based a priori estimates, we added zero-mean white noise to the a priori estimates with standard deviations of 5%, 10%, and 20% of those of the a priori estimates. The degraded a priori SAC parameter estimates were then optimized using SLS starting from each of the randomly perturbed parameter estimates. A total of 500 trials was carried out. Figs. 12a–c show the empirical probability density functions (PDF) of the three calibrated parameters, UZTWM, UZFWM, and ZPERC for ATIT2. Note that, over the range of the noise levels prescribed, the PDFs of the optimized parameter estimates cover a rather narrow region in the parameter space, and that the most probable parameter values differ very little across different noise levels and from the parameter values calibrated with no noise added. It was also observed that ZPERC, which is responsible for percolation, is the least identifiable among the parameters in that its PDF's are significantly more diffuse. This is in agreement with finding of Gupta and Sorooshian (1983) and Sorooshian and Gupta (1983) in which they attribute the behavior to the model structure. Fig. 12d shows the PDF of the multi-scale objective function value corresponding to Fig. 12a–c, expressed as the percent difference in the MSOF associated with the a priori parameter estimates. Note that the variability of the multi-scale objective function is rather small (in the range of ±3%) even when a noise of 20% was applied to the a priori parameters. It suggests that sensitivity of SLS to the goodness of the initial (i.e. a priori) parameter sets may be relatively modest.

Sensitivity to the step size

Unlike SCE, which has multiple parameters to contend with (though only one, the number of complexes, may be considered critical), SLS has only one adaptable parameter, the search step size (expressed as a fraction of the feasible parameter region). Generally speaking, the choice of the step size depends on the objective function surface in the parameter space and, to a lesser degree, computational considerations. If model parameters are highly collinear, it is very likely that the objective function surface is flat and has a wide minimum region. Moreover, there may be numerous 'pits' in the minimum region due to data and model uncertainties. As such, too small a step size can prematurely terminate the search at a 'pothole' nearest to the a priori parameter estimates. Too large a step size, on the other hand, may skip over the true minimum. A comprehensive and rigorous analysis of the dependence of the optimal step size on the objective function surface is left as a future endeavor. In this work, we performed a numerical experiment to determine the step size empirically. Fig. 13 shows the minimum values of MSOF for nine WGRFC basins as obtained using various step sizes, ranging from 1/5000 to 1/3 of the entire feasible region of the parameter space (Table 3). The figure suggests that the step size for SLS does not have to be known precisely, and that any step size within the interval [1/200; 1/20] leads to similar performance.

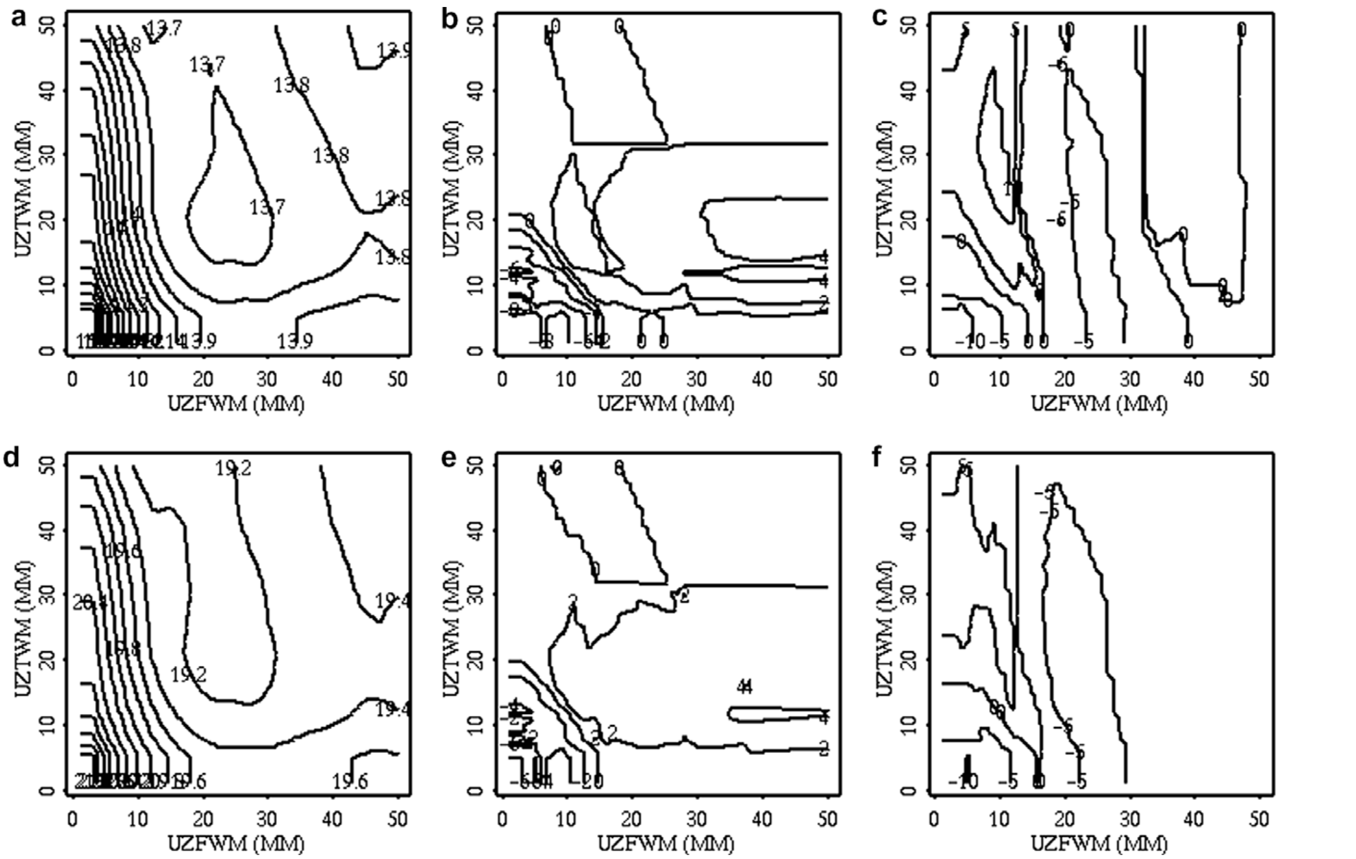

Figure 10 (a) The objective function surface with respect to UZFWM and UZTWM. The objective function is made of model simulation errors at the hourly and monthly scales of aggregation only. (b) Gradient of the objective function shown in (a) with respect to UZFWM. (c) Same as (b), but with respect to UZTWM. Figures d–f are the same as figures a–b, respectively, except that the objective function is made of simulation errors at the hourly, daily, weekly and monthly scales of aggregation.

All results presented in this paper were based on the step size 1/50.

Summary and conclusions

Automatic calibration of hydrologic models via local optimization of a priori parameter estimates is explored as a computationally inexpensive and physically appealing alternative to bounded global optimization. The hydrologic model used is the Sacramento soil moisture accounting model driven by hourly multisensor (radar and rain gauge) precipitation forcing. The a priori estimates of the SAC parameters used are from the soil-based parameterization of Koren et al. (2000, 2003). The technique developed in this work for local optimization is a simple but very effective and efficient pattern search algorithm referred to as the Stepwise Linear Search (SLS). To evaluate the technique, calibration and validation experiments were carried out using 20 basins in the National Weather Service West Gulf River Forecast Center's (NWS/WGRFC) service area in Texas, USA. For comparison, global optimization using the Shuffled Complex Evolution algorithm (SCE, Duan et al., 1992; Duan, 2003) was also carried out. To correct biases in the precipitation data and potential evaporation climatology and to estimate hourly empirical unit hydrographs, variational assimilation techniques were used. To examine the sensitivity of bias correction and unit hydrograph estimation to the quality of the initial hydrologic model parameters, two sets of hydrologic model parameters were used: the soil-based (Koren et al., 2000, 2003) and the operationally-used.

The results indicate that, in general, the quality of the soil-based a priori estimates of the SAC parameters is comparable to that of the operationally-used estimates at the RFC, and that both the soil-based and the operationally-used parameters can be improved significantly by optimization. In the dependent validation that covered a 5-year period, the performance of SLS was found to be only slightly inferior to that of SCE in terms of reduction in the objective function value (see Table 4). In the independent validation that covered five 1-year periods, however, the performance of SLS was found to be comparable to or somewhat better than that of SCE (see Table 5). Furthermore, validation results showed that, whereas the posterior (i.e. optimized) estimates from SLS remain consistently close to the a priori estimates, the globally optimized estimates from SCE are spread over the entire feasible region of the parameter space. In terms of computational economy, SLS needed only

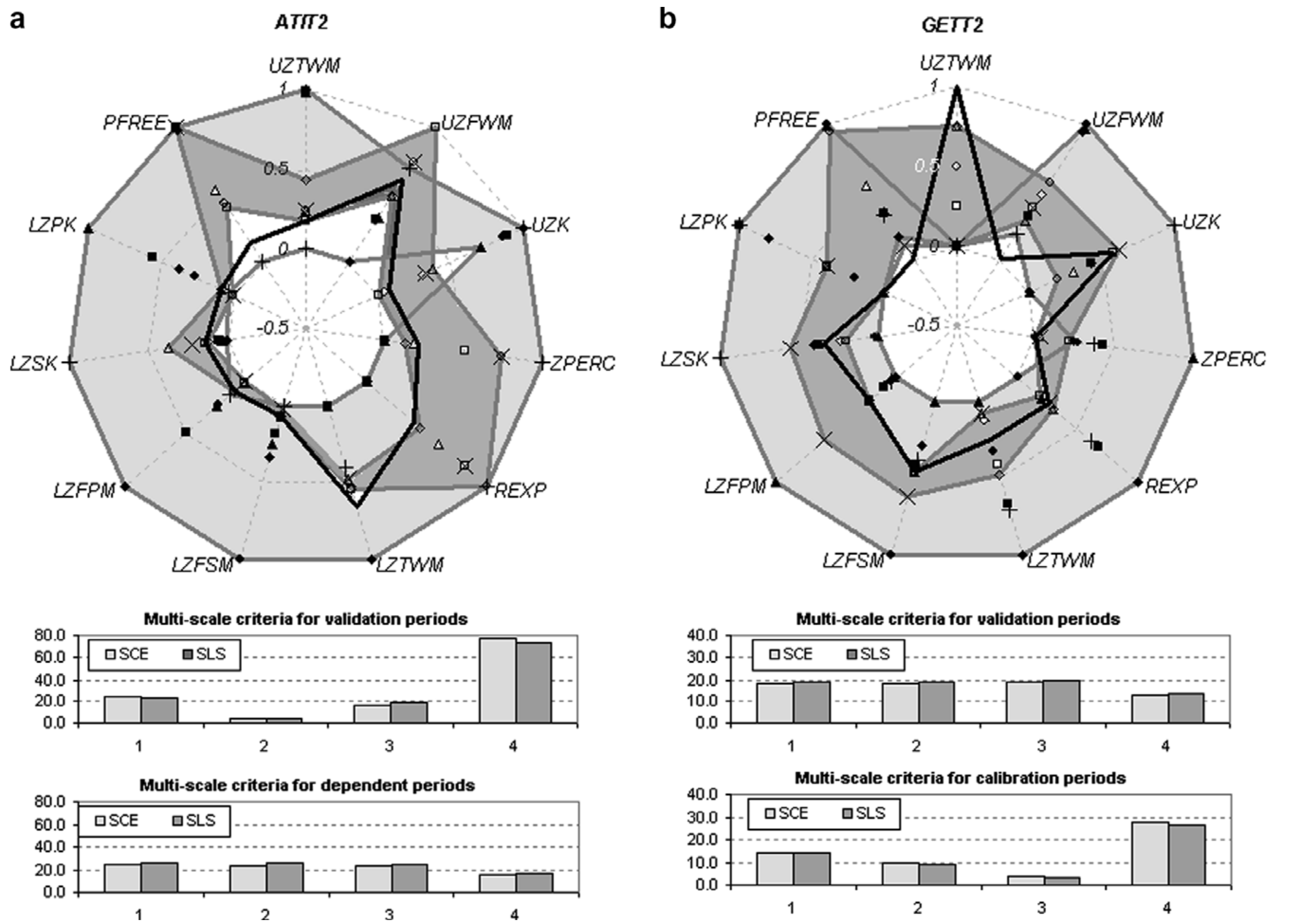

Figure 11 Examples of the normalized Parameter Rose for ATIT2 and GNVT2. In each rose, the 'doughnut' represents the entire parameter space (normalized to [0, 1]). The solid line represents the a priori parameter estimates. The light- and dark-shaded areas represent the sub-space spanned by SCE- and SLS-optimized parameter estimates, respectively. MSOF of calibration and validation periods is shown at the bottom.

about 1/20–1/300 of the function evaluations required by SCE. It suggests that, given the uncertainties in the input data alone (be they due to lack of data or poor quality), the filtering approach to hydrologic model calibration via local optimization of a priori parameter estimates offers a physically-consistent and computationally very efficient alternative to global optimization.

In an attempt to emulate the manual calibration process and to improve the efficiency and effectiveness of SLS, we also explored the use of a multi-scale objective function (MSOF), composed of contributions from simulation errors aggregated from hourly to monthly scales. The results show that the MSOF helps smooth, or regularize, the objective function surface and reduces the likelihood of the algorithm being trapped in a 'pothole.' It is also shown that MSOF yields smaller errors in simulated flow across a range of temporal scales of aggregation than an objective function made only of short- and long-term errors (e.g. a combination of root mean square and mean errors).

To assess the sensitivity of the proposed methodology to the goodness of the a priori parameter estimates, a numerical experiment was carried out. It is shown that, in the probabilistic sense, the resulting posterior (i.e. optimized) estimates do not differ significantly from one another whether an additional noise of 5%, 10% or 20% of the variability of the parameters is added to the a priori estimates, an indication that the procedure is rather robust. It is also important to note that the reduction in MSOF was significantly lower than the level of added noise. These results are necessarily limited to SAC and the soil-based parameters derived from Koren et al. (2003). While application of SLS to other lumped models and a priori parameters (presumably of widely varying quality) is seemingly straightforward, it is difficult to predict how SLS may perform, an area of research left as a future endeavor.

As stated in the Introduction Section, the ultimate goal of the proposed approach is automatic calibration of distributed-parameter hydrologic models. Even with greatly reduced computational burden, exhaustive optimization of distributed parameters is highly impractical, if not infeasible. As such, efforts should be directed to attaining a practical balance between reducing the number of control variables and increasing the spatial resolution of the model while accounting for input, model and parametric uncertainties.

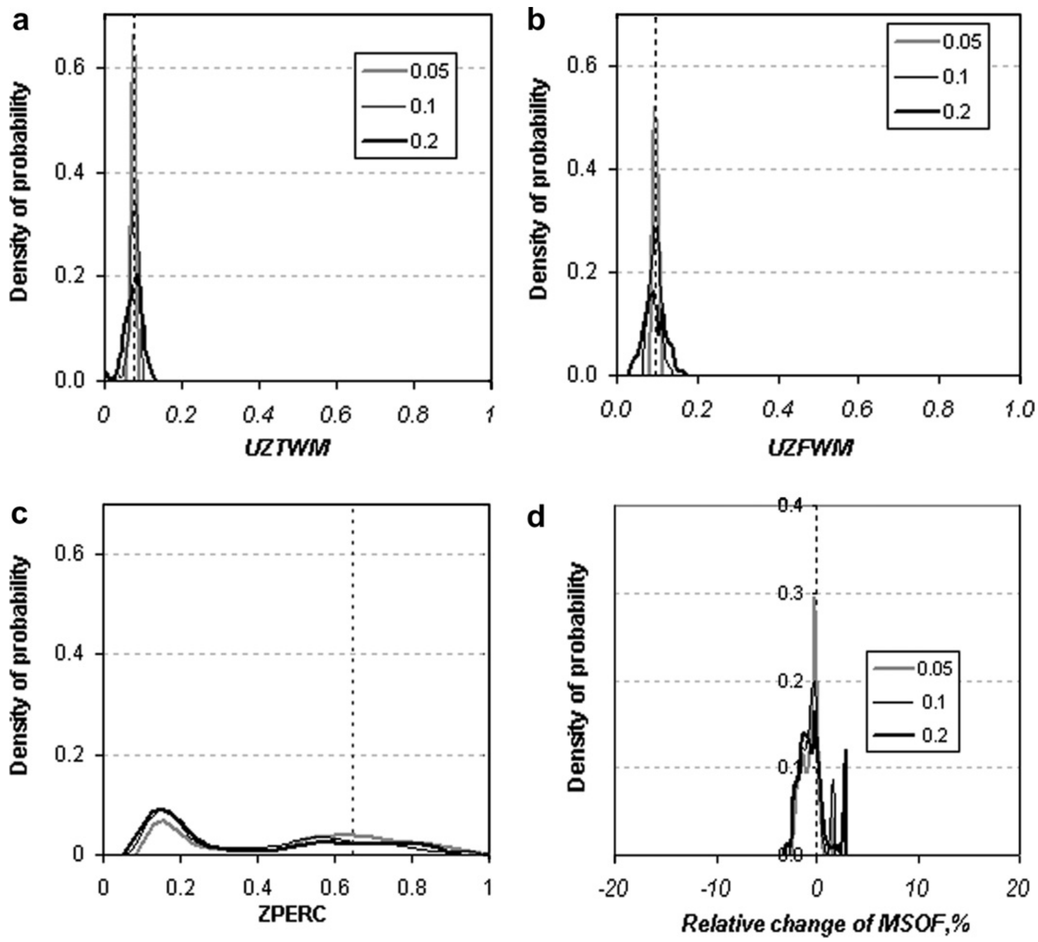

Figure 12 (a) Empirical probability density functions of the posterior (i.e. optimized) estimate of UZTWM. The optimization is by SLS. The a priori value of UZTWM was given a random noise ranging from 5%, 10% to 20% (denoted as 0.05, 0.1 and 0.2, respectively) of the spatial variability of the a priori UZTWM in the basin. (b) Same as (a), but for UZFWM. (c) Same as (a), but for ZPERC. (d) Empirical probability density functions of the relative difference between the posterior multi-scale objective function value based on the noise-added a priori parameters and that based on the a priori parameters without noise.

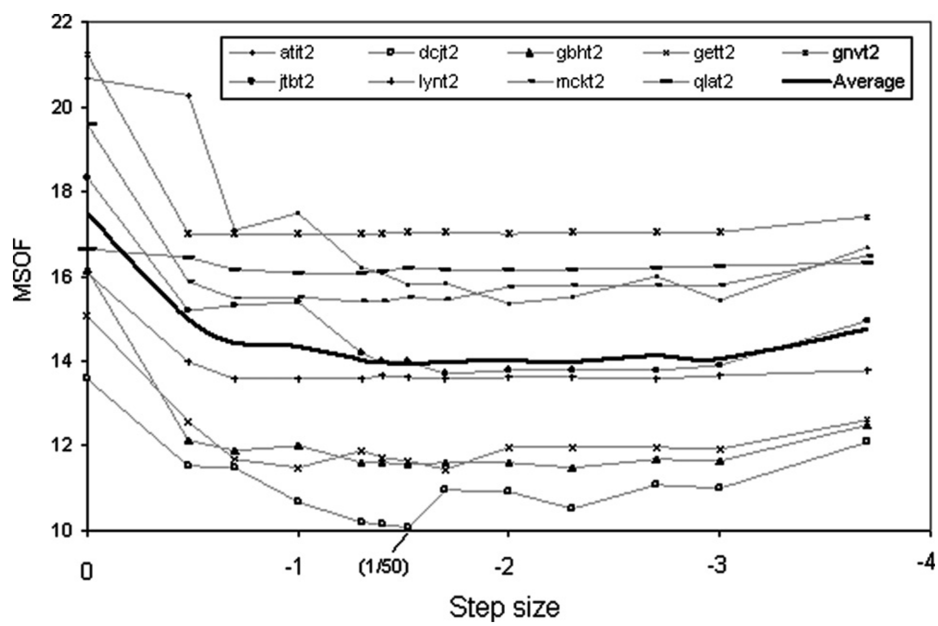

Figure 13 Sensitivity of the final multi-scale objective function value (MSOF) to the step size used in SLS. The values on the *y*-axis represent the MSOFs associated with the posterior parameter estimates.

Table 5 Comparison of multi-scale objective function values from SCE and SLS for five calibration and validation periods of 4- and 1-years durations, respectively

Basin	SCE					SLS				
	1	2	3	4	5	1	2	3	4	5
Calibration										
GBHT2	**14.1**[a]	**14.0**	**13.8**	**14.3**	**13.8**	14.3	14.4	14.0	14.6	14.1
GETT2	**18.3**	**18.6**	**18.9**	**12.6**	**18.3**	18.8	18.9	19.5	12.9	18.8
HBMT2	**31.9**	**33.5**	**32.9**	**33.3**	**32.1**	33.4	35.4	34.8	35.0	33.0
HNTT2	**36.8**	38.8	**28.3**	**34.2**	**36.5**	36.9	38.8	28.6	34.5	36.7
JTBT2	**11.7**	12.6	7.21	15.7	**13.2**	12.6	**12.2**	**7.15**	**15.6**	13.8
KNLT2	**15.0**	**18.7**	**18.0**	**18.7**	**10.9**	17.3	20.1	19.8	19.7	11.2
LYNT2	12.7	**12.8**	**12.3**	**12.2**	**8.54**	12.7	13.0	12.4	12.3	8.69
MTPT2	38.0	41.7	41.3	**40.0**	38.0	**37.9**	**41.5**	41.3	40.1	**37.9**
Validation										
GBHT2	**13.0**	14.6	**15.4**	**10.3**	**15.0**	14.8	**14.3**	15.7	11.4	15.4
GETT2	14.2	9.73	3.57	27.7	13.9	**14.1**	**8.71**	**3.50**	**26.2**	**13.0**
HBMT2	29.9	**27.9**	**25.1**	**21.5**	**35.9**	**27.0**	34.6	27.6	25.2	47.1
HNTT2	33.3	4.81	**66.1**	47.4	32.0	**32.0**	**4.51**	66.3	**44.1**	32.0
JTBT2	12.4	4.32	25.9	9.59	26.2	**4.96**	**3.79**	**24.7**	**6.47**	**17.6**
KNLT2	31.9	**4.38**	13.7	18.0	47.9	**28.5**	11.1	**10.8**	**15.3**	**43.1**
LYNT2	**11.4**	5.89	11.0	11.4	**36.9**	11.8	**4.92**	**10.3**	**11.1**	37.3
MTPT2	**45.1**	16.2	20.9	34.2	52.4	45.4	**14.5**	**19.6**	**33.7**	**52.0**

[a] The smaller MSOF between SCE and SLS is in bold.

Acknowledgements

This work is supported by the Advanced Hydrologic Prediction Service (AHPS) program of the National Weather Service (NWS), by the Climate Predictions Program for the Americas (CPPA) of the Climate Program Office (CPO), both of the National Oceanic and Atmospheric Administration (NOAA), and by the Integrated Systems Solutions component of the NASA Decision Support through Earth Science Results Cooperative Agreement Notice (CAN) NN-H-04-Z-YO-010-C of the National Aeronautic and Space Administration (NASA). These supports are gratefully acknowledged. The views expressed are those of the authors and do not necessarily represent those of the National Weather Service. The authors would like to thank Robert Corby and Tracy Howieson of the NWS/West Gulf River Forecast Center (WGRFC) for providing the 1-hour precipitation and streamflow data used in this work and for many helpful discussions throughout the course of this work. The authors would also like to thank Michael Smith, Seann Reed and John Schaake of the NWS Hydrology Laboratory (HL) and Qingyun Duan of the Lawrence Livermore Laboratory for many helpful discussions. The authors are grateful to the two anonymous reviewers for many very helpful comments which significantly improved the manuscript.

References

Archetti, F., Schoen, F., 1984. A survey on the global optimization problem: general theory and computational approaches. Ann. Oper. Res. 1, 87–110.

Beven, K.J., Kirkby, M.J., 1979. A physically based variable contributing area model of basin hydrology. Hydrol. Sci. Bull. 24 (1), 43–69.

Brazil, L.E., 1988. Multilevel calibration strategy for complex hydrologic simulation models, PhD dissertation, Dept. of Civil Eng., Colorado State University, Fort Collins, Colorado, p. 235.

Burnash, R.J.C., Ferral, R.L., McGuire, R.A., 1973. A Generalized Streamflow Simulation System: Conceptual Models for Digital Computers. Joint Fed.-State River Forecast Center, Sacramento, CA.

Chaubey, I., Haan, C.T., Grunwald, S., Salisbury, J.M., 1999. Uncertainty in the model parameters due to spatial variability of rainfall. J. Hydrol. 220 (1–2), 48–61.

Chow, V.T., Maidment, D.R., Mays, L.W., 1988. Applied Hydrology. McGraw Hill, p. 572.

Cosgrove, B., Lohmann, D., Mitchell, K., Houser, P., Wood, E., Schaake, J., Robock, A., Marshall, C., Sheffield, J., Duan, Q., Luo, L., Higgins, W., Pinker, R., Tarpley, D., Meng, J., 2003. Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project. J. Geophys. Res. 108 (D22), 12.

Duan, Q., 2003. Global optimization for watershed model calibration. In: Duan et al. (Eds.), Calibration of Watershed Models, Water Science and Application, vol. 6. American Geophysical Union, 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.

Eckhardt, K., Arnold, J.G., 2001. Automatic calibration of a distributed catchment model. J. Hydrol. 251, 103–109.

Elvidge, C.D., Milesi, C., Dietz, J.B., Tuttle, B.T., Sutton, P.C., Nemani, R., Vogelmann, J.E., 2004. U.S. constructed area approaches the size of Ohio. EOS 85 (24), 233.

Evtushenko, Y.G., 1973. Numerical algorithms for global extremum search (case of a non-uniform mesh). USSR Comp. Math. Math. Phys. 11, 1390–1403.

Francés, F., Veléz, J.I., Veléz, J.J., 2007. Split-parameter structure for the automatic calibration of distributed hydrological models. J. Hydrol. 332, 226–240.

Gupta, H.V., Bastidas, L.A., Vrugt, J.A., Sorooshian, S., 2003. Multiple criteria global optimization for watershed model

calibration. In: Duan et al. (Eds.), Calibration of Watershed Models, Water Science and Application, vol. 6. American Geophysical Union, Washington, DC, pp. 125–132.
Gupta, V.K., Sorooshian, S., 1983. Uniqueness and observability of conceptual rainfall–runoff model parameters: the percolation process examined. Water Res. Res. 19, 269–276.
Heuvelmans, G., Muys, B., Feyen, J., 2006. Regionalisation of the parameters of a hydrological model: comparison of linear regression models with artificial neural nets. J. Hydrol. 319, 245–265.
Hudlow, M.D., 1988. Technological developments in real-time operational hydrologic forecasting in the United States. J. Hydrol. 102, 69–92.
Jazwinski, A.H., 1970. Stochastic Processes and Filtering Theory. Academic Press, p. 376.
Koren, V.I., Smith, M., Wang, D., Zhang, Z., 2000. Use of soil property data in the derivation of conceptual rainfall–runoff model parameters, preprints. In: 15th Conf. on Hydrol., Long Beach, CA, Amer. Meteor. Soc., 10–14 January, 2000, pp. 103–106.
Koren, V., Smith, M., Duan, Q., 2003. Use of a priori parameter estimates in the derivation of spatially consistent parameter estimates of rainfall–runoff models. In: Duan et al. (Eds.), Calibration of Watershed Models, Water Science and Application, vol. 6. American Geophysical Union, Washington, DC, pp. 239–255.
Koren, V., Reed, S., Smith, M., Zhang, Z., Seo, D.-J., 2004. Hydrology laboratory research modeling system (HL-RMS) of the US National Weather Service. J. Hydrol. 291, 297–318.
Koren, V., Schaake, J., Duan, Q., Smith, M., Cong, S., September 1998. PET upgrades to NWSRFC – Project plan, HRL Internal Report (copy available on request from: Hydrology Laboratory, Office of Hydrologic Development, NOAA, National Weather Service, 1325 East-West Highway, Silver Spring, MD 20910, USA).
Kuczera, G., Parent, E., 1998. Monte Carlo assessment of parameter uncertainty in conceptual catchment models: the Metropolis algorithm. J. Hydrol. 211 (1–4), 69–85.
Leavesley, G.H., Hay, L.E., Viger, R.J., Markstrom, S.L., 2003. Use of a priori parameter-estimation methods to constrain calibration of distributed-parameter models. Calibration of Watershed Models. In: Duan et al. (Eds.), Calibration of Watershed Models, Water Science and Application, vol. 6. American Geophysical Union, Washington, DC, pp. 255–266.
Li, Z., Navon, I.M., 2001. Optimality of variational data assimilation and its relationship with Kalman filter and smoother. Quart. J. Roy. Meteor. Soc. 127, 661–683.
Lohmann, D., Mitchell, K., Houser, P., Wood, E., Schaake, J., Robock, A., Cosgrove, B., Sheffield, J., Duan, Q., Luo, L., Higgins, W., Pinker, R., Tarpley, D., 2004. Streamflow and water balance intercomparisons of four land surface models in the North American Land Data Assimilation System project. J. Geophys. Res. 109 (D07S91), 22.
Madsen, H., 2000. Automatic calibration of a conceptual rainfall–runoff model using multiple objectives. J. Hydrol. 235 (3–4), 276–288.
Madsen, H., 2003. Parameter estimation in distributed hydrological catchment modelling using automatic calibration with multiple objectives. Adv. Water Resour. 26, 205–216.
Maidment, D.R., Olivera, J.F., Calver, A., Eatherall, A., Fraczek, W., 1996. A unit hydrograph derived from a spatially distributed velocity field. J. Hydrol. Process. 10 (6), 45–52.
Mazi, K., Koussis, A.D., Restrepo, P.J., Koutsoyiannis, D., 2004. A groundwater-based, objective-heuristic parameter optimisation method for a precipitation-runoff model and its application to a semi-arid basin. J. Hydrol. 290 (3–4), 243–258.
McCabe, M.F., Franks, S.W., Kalma, J.D., 2005. Calibration of a land surface model using multiple data sets. J. Hydrol. 302 (1–4), 1209–1222.
Mertens, J., Madsen, H., Feyen, L., Jacques, D., Feyen, J., 2004. Including prior information of effective soil parameters in unsaturated zone modeling. J. Hydrol. 294, 251–269.
Miller, D.A., White, R.A., 1999. A conterminous United States multilayer soil characteristics data set for regional climate and hydrology modeling. Earth Interact. 2. <http://EarthInteractions.org>.
Misirli, F., Gupta, H.V., Sorooshian, S., Thiemann, M., 2003. Bayesian recursive estimation of parameter and output uncertainty for watershed models. In: Duan et al. (Eds.), Calibration of Watershed Models, Water Science and Application, vol. 6. American Geophysical Union, Washington, DC, pp. 113–124.
Mitchell, K., Lohmann, D., Houser, P., Wood, E., Schaake, J., Robock, A., Cosgrove, B., Sheffield, J., Duan, Q., Luo, L., Higgins, W., Pinker, R., Tarpley, D., Lettenmaier, D., Marshall, C., Entin, J., Pan, M., Shi, M., Koren, V., Meng, J., Ramsay, B., Bailey, A., 2004. The multi-institution North American Assimilation System (NLDAS): utilizing multiple GCIP products and partners in a continental distributed modeling system. J. Geophys. Res. 109 (D07S90), 32.
Parada, L.M., Fram, J.P., Liang, X., 2003. Multi-resolution calibration methodology for hydrologic models: Application to a subhumid catchment. In: Duan et al. (Eds.), Calibration of Watershed Models, Water Science and Application, vol. 6. American Geophysical Union, Washington, DC, pp. 132–139.
Press, W.H., Flannery, B.P., Teukolsky, S.A., Vetterling, W.T., 1986. Numerical Recipes. Cambridge University Press, p. 818.
Reed, S., Koren, V., Smith, M., Zhang, Z., Moreda, F., Seo, D.-J., 2004. Overall distributed model intercomparison project results. J. Hydrol. 298, 27–60.
Refsgaard, J.C., 1997. Parameterization, calibration, and validation of distributed hydrological models. J. Hydrol. 198, 69–97.
Schaake, J., Duan, Q., Koren, V., Hall, A., 2001. Toward improved parameter estimation of land. In: Dolman, Hall, Kavvas, Oki, Pomeroy (Eds.) Soil–Vegetation–Atmosphere Transfer Schemes and Land-Scale Hydrological Models, IAHS Pub. No. 270, pp. 91–97.
Schweppe, F.C., 1973. Uncertain Dynamic Systems. Prentice-Hall, Inc., Englewood Cliffs, NJ, p. 563.
Seibert, J., 2000. Multi-criteria calibration of a conceptual runoff model using a genetic algorithm. Hydrol. Earth Syst. Sci. 4, 215–224.
Seo, D.-J., Breidenbach, J.P., Johnson, E.R., 1999. Real-time estimation of mean field bias in radar rainfall data. J. Hydrol. 223, 131–147.
Seo, D.-J., Breidenbach, J.P., 2002. Real-time correction of spatially nonuniform bias in radar rainfall data using rain gauge measurements. J. Hydrometeorol. 3, 93–111.
Seo, D.-J., Koren, V., Cajina, N., 2003a. Real-time variational assimilation of hydrologic and hydrometeorological data into operational hydrologic forecasting. J. Hydrometeorol. 4, 627–641.
Seo, D.-J., Koren, V., Cajina, L., Corby, R., Finn, B., Bell, F., Howieson, T., 2003b. Real-time variational assimilation of streamflow and radar-based precipitation data into operational hydrologic forecasting. Paper: EAE03-A-14671; NH4-1MO4P-1416, AGU-EGS Meeting, Nice, France.
Smith, M., Laurine, D.P., Koren, V.I., Reed, S., Zhang, Z., 2003. Hydrologic model calibration in the national weather service. In: Duan, Gupta, Sorooshian, Rousseau, Turcotte (Eds.), Calibration of Watershed Models, Water Science and Application, vol. 6. American Geophysical Union, pp. 133–152.
Smith, J.A., Seo, D.-J., Baeck, M.L., Hudlow, M.D., 1996. An intercomparison study of NEXRAD precipitation estimates. Water Resour. Res. 32, 2035–2045.
Smith, M., Seo, D.-J., Koren, V., Reed, S., Zhang, Z., Duan, Q., Moreda, F., Cong, S., 2004. The distributed model intercomparison project (DMIP): motivation and experiment design. J. Hydrol. 298, 4–26.

Sorooshian, S., Gupta, V.K., 1983. Automatic calibration of conceptual rainfall—runoff models: the question of parameter observability and uniqueness. Water Resources Res. 19 (1), 251—259.

Teleb, R., Azadivar, F., 1994. A methodology for solving multi-objective simulation-optimization problems. Eur. J. Oper. Res. 72, 135—145.

Torczon, V., 1997. On the convergence of pattern search algorithms. SIAM J. Optim. 7, 1—25.

Vieux, B.E., Cui, Z., Gaur, A., 2004. Evaluation of a physic-based distributed hydrologic model for flood forecasting. J. Hydrol. 298, 155—177.

Vieux, B.E., Moreda, F., 2003. Ordered physics-based parameter adjustment of a distributed model. In: Duan, Q., Sorooshian, S., Gupta. H., Rosseau, A., Turcotte, R. (Eds.), Advances in the Calibration of Watershed Models, AGU Water Science and Applications Series, pp. 267—281.

Vrugt, J.A., Bouten, W., Gupta, H.V., Hopmans, J.W., 2003a. Toward improved identifiability of soil hydraulic parameters: on the selection of a suitable parametric model. Vadose Zone J. 2, 98—113.

Vrugt, J.A., Gupta, H.V., Bouten, W., Soroosian, S., 2003b. A shuffled complex evolution metropolis algorithm for estimating posterior distribution of watershed model parameters. In: Duan et al. (Eds.), Calibration of Watershed Models, Water Science and Application, vol. 6. American Geophysical Union, Washington, DC, pp. 105—112.

Whitehead, P.G., Neal, C., Seden-Perrington, S., Christophersen, N., Langan, S., 1986. A time-series approach to modelling stream acidity. J. Hydrol 85, 281—303.

Wilson, J., Brandes, E., 1979. Radar measurement of rainfall — a summary. Bull. Am. Meteorol. Soc. 60, 1048—1058.

Yapo, P.O., Gupta, H.V., Soroosian, S., 1997. Multi-objective global optimization for hydrologic models. J. Hydrol. 204, 83—97.

Young, C.B., Bradley, A.A., Krajewski, W.F., Kruger, A., 2000. Evaluation of NEXRAD multisensor precipitation estimates for operational hydrologic forecasting. J. Hydrometeorol. 1, 241—254.