Available online at www.sciencedirect.com
ScienceDirect
Physics and Chemistry of the Earth xxx (2008) xxx–xxx
PHYSICS and CHEMISTRY of the EARTH
www.elsevier.com/locate/pce

Use of soil moisture observations to improve parameter consistency in watershed calibration

Victor Koren *, Fekadu Moreda, Michael Smith

Office of Hydrologic Development, NOAA National Weather Service, 1325 East–West Highway, Silver Spring, Maryland 20910, USA

Received 17 October 2007; received in revised form 15 January 2008; accepted 28 January 2008

Abstract

Calibration is a critical component in the implementation of operational models for river forecasting. It has traditionally relied on minimizing the errors between simulated and observed basin outlet hydrographs. However, considering numerous sources of uncertainty and the complexity of recently-developed models, this approach often fails to reduce parameter uncertainties. One of the possibilities to reduce parameter uncertainty would be use of additional independent data in the model evaluation. Unfortunately, such data are limited and their quality is usually not well defined. This study investigates the potential use of soil moisture measurements in the model calibration process. While these data are not commonly available, there is potential for considerable expansion of soil moisture measurements in the near future. Comprehensive soil moisture measurements from the Oklahoma Mesonet are used in the analysis. The Sacramento Soil Moisture Accounting model with a new heat transfer component (SAC-HT) is applied to more than 20 watersheds of sizes ranging from 200 to 4000 km^2 to answer the question: can the use of soil moisture data improve calibration reliability without an unacceptable reduction in the accuracy of the simulated outlet hydrograph. Three cases of simulated soil moisture and hydrographs are analysed: (1) the control run with the use of *a priori* parameters; (2) automatic calibration based on outlet hydrograph goodness-of-fit only; and (3) automatic calibration based on outlet hydrographs and basin average soil moisture computed at two depths. Results show deficiencies in model calibration using only outlet hydrograph goodness-of-fit as a measure. The automatic calibration in this case improves runoff simulation results on average by 45% compared to the use of *a priori* parameters. Soil moisture dynamics and trends are also reproduced reasonably well; however, large soil moisture biases can be seen. These biases in the top soil layer are 36% higher than in the control run. Addition of soil moisture measurements into the calibration process reduces soil moisture biases at the both soil layers by 40% without considerable reduction in runoff accuracy (5%) and improves internal consistency of calibration. The use of soil moisture measurements provides more benefit for 'dry' watersheds when there is no strong direct interconnection between runoff and soil moisture.

Keywords: Automatic calibration; Soil moisture; Optimization criteria; Parameter consistency

1. Introduction

Calibration is a critical component in the implementation of operational models for river forecasting. Traditionally, calibration of watershed models has relied on minimizing the errors between simulated and observed basin outlet hydrographs. However, considering numerous sources of uncertainty and the complexity of recently-developed models, this approach often fails to generate consistent parameter sets (Bastidas et al., 2003; Seibert and McDonnell, 2003). Reasons for this failure are well documented. For example, Jakeman and Hornberger (1993) argued that the information content in a rainfall-runoff record (i.e., 'hard data') is sufficient to support models of only very limited complexity with a few model parameters to calibrate. Kuczera and Mroczkowski

* Corresponding author.
E-mail address: Victor.Koren@noaa.gov (V. Koren).

1474-7065/$ - see front matter © 2008 Published by Elsevier Ltd.
doi:10.1016/j.pce.2008.01.003

(1998) reported similar finding that models with more than four parameters calibrated to streamflow data often have poorly identified parameters.

One possibility to reduce calibration uncertainty is to utilize additional observations in the process of model calibration and evaluation (Ambroise et al., 1995; Refsgaard, 1997; Kuczera and Mroczkowski, 1998; Bastidas et al., 2003; Seibert and McDonnell, 2003). Bastidas et al. (2003) used point measurements of near-surface soil moisture and temperature as well as heat fluxes to constrain the land surface model parameters via multi-objective calibration. The compromise solution was found from a set of Pareto optimization results. They found that additional data increased the model consistency but the accuracy of soil moisture and temperature simulations deteriorated with depth. Kuczera and Mroczkowski (1998) found that augmenting streamflow data with other measurements may not reduce parameter uncertainty. For example, use of groundwater level data in tests of a hydrosalinity model did little to reduce the uncertainty in poorly defined parameters, whereas use of stream salinity data substantially reduced parameter uncertainty. They recommend performing an assessment of the worth of additional data prior to wide-scale application.

Other attempts to supplement streamflow data for calibration include what is referred as the use of 'soft' or qualitative data. For example, Seibert and McDonnell (2003) demonstrated the use of additional soft data such as isotope-based 'new' water information transformed into quantitative data through fuzzy measures of model-simulation and parameter-value acceptability. They observed that parameter uncertainty measured as a normalized parameter variation was reduced on average by 60% when additional soft data were used. Seibert and McDonnell (2003) advocate that the use of additional data might allow for assessing internal model consistency and, as a result, lead to a more realistic model structure. Also, Casper et al. (2007) used soil moisture measurements at a local site in a fuzzy rule-based system to improve model calibration and discharge prediction at the watershed outlet.

This brief introduction highlights the importance of exploiting additional 'hard' or 'soft' information in the calibration process. However, there are still many questions regarding this approach such as how to combine additional information with basic data to achieve maximum benefit and how much value can truly be extracted from limited point measurements. This paper investigates the potential use of soil moisture measurements as additional 'hard' information in the watershed model calibration process. We define a single-criterion objective function that measures the combined goodness-of-fit of simulated soil moisture and outlet hydrographs. Comprehensive multi-layer soil moisture measurements from the Oklahoma Mesonet, USA are used in the analysis. A simple local calibration technique was applied to the modified Sacramento Soil Moisture Accounting model (SAC-HT) (Koren et al., 2006). Calibration tests with and without soil moisture data are performed for 20 river basins to analyse parameter consistency.

2. Study area and data

Twenty watersheds with areas ranging from 200 to 4000 km^2 were selected within the Arkansas-Red River basin in Oklahoma as shown in Fig. 1. The watershed properties are shown in Table 1. The area encompasses a wide variety of climatic conditions, ranging from an arid region in the western part to a humid region in the eastern part. The ratio of annual precipitation to potential evaporation (P/PE) varies from 0.57 in the western portion to 1.18 in the eastern portion of the study domain, revealing a strong gradient. This area has the longest archive of 4 km NEXRAD-based multi-sensor precipitation grids, and these rainfall estimates have been thoroughly evaluated (Johnson et al., 1999; Young et al., 2000) and used for major model evaluation studies (e.g., Reed et al., 2004). Observed hourly streamflow data are available at each of the 20 watershed outlets. The SAC-HT model also requires potential evaporation demand input to calculate actual evapotranspiration. In this analysis, we used climatological monthly free surface water evaporation (Farnsworth et al., 1982) seasonally adjusted for vegetation effects.

The test area has a unique soil moisture data collection network, the Oklahoma Mesonet. The Oklahoma Mesonet provides real-time data including soil moisture measurements at four depths (5, 25, 60, 75 cm) from more than 100 sites since 1997. All sites are equipped with heat dissipation soil moisture sensors which measure the temperature change of a heat pulse (Brock et al., 1995). To determine the representativeness of these measurements, Illston et al. (2004) compared soil moisture measurements at Mesonet sites to soil core samples at 5 and 25 cm during the enhanced drying phase. They concluded that overall, the Oklahoma Mesonet sensors performed quite well in representing average soil moisture estimates. The average soil moisture at the 5 cm was 0.22 and 0.25 $m^3\,m^{-3}$ from Mesonet and soil cores measurements, respectively. For 25 cm, they found that the average soil moisture was 0.27 $m^3\,m^{-3}$ for both measurements. However, they also uncovered a significant decrease in the soil moisture variability for the Mesonet observations. The standard deviation of soil moisture at the 5 cm was 0.06 and 0.11 $m^3\,m^{-3}$ from Mesonet and soil core measurements, respectively. Similarly, for 25 cm, the standard deviation was 0.05 and 0.09 $m^3\,m^{-3}$ for Mesonet and soil cores, respectively.

There are two issues to consider while using the volumetric soil moisture data from the Mesonet sites (Koren et al., 2006). First, the instantaneous volumetric soil moisture measurement at a station is related to the soil type and the physiographic properties of the location in addition to the availability of moisture supply, i.e., precipitation in the area. This hampers comparisons of stations located in

Please cite this article in press as: Koren, V. et al., Use of soil moisture observations to improve parameter consistency ..., J. Phys. Chem. Earth (2008), doi:10.1016/j.pce.2008.01.003

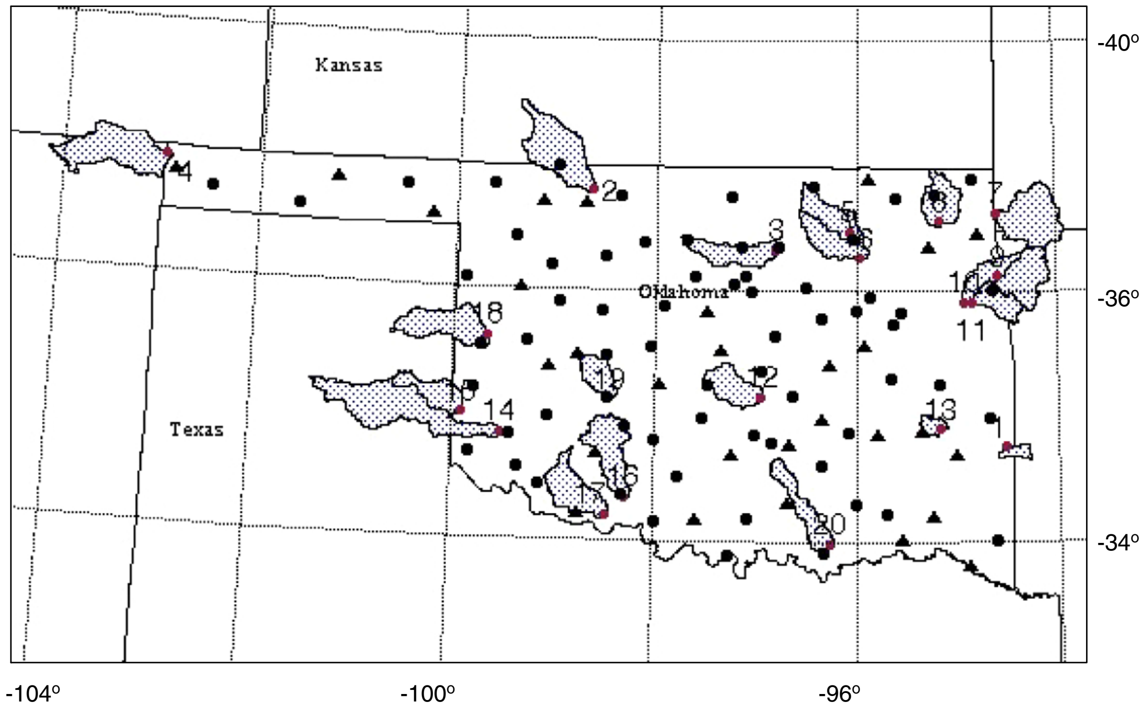

Fig. 1. Map of the Oklahoma Mesonet region with location of tested watersheds (shaded areas) and soil moisture site measurements at four layers (filled circles) and only two top layers (filled triangles).

Table 1
Tested watersheds and some properties

	USGS ID	Name of the basin	Area (km^2)	ELV (m)	P (mm)	PE (mm)	P/PE	Runoff (mm)
1	7247250	Black Fork below Big Creek near Page OK	193	684	1356	1146	1.18	467
2	7148400	Salt Fork Arkansas River near Alva OK	2613	1292	758	968	0.78	81
3	7153000	Black Bear Creek at Pawnee OK	1492	803	858	1021	0.84	224
4	7154500	Cimarron River near Kenton OK	2864	4262	412	728	0.57	18
5	7176500	Bird Creek at Avant OK	943	651	967	1073	0.90	365
6	7177500	Bird Creek near Sperry OK	2344	579	959	1070	0.90	352
7	7189000	Elk River near Tiff City Mo	2258	751	1119	1116	1.00	320
8	7191000	Big Cabin Creek near Big Cabin OK	1165	622	1078	1085	0.99	387
9	7195500	Illinois River near Watts OK	1644	894	1146	1106	1.04	345
10	7196500	Illinois River near Tahlequah OK	2484	664	1154	1123	1.03	385
11	7197000	Baron Fork at Eldon OK	795	701	1168	1125	1.04	337
12	7230500	Little River near Tecumseh OK	1181	899	914	1068	0.86	148
13	7247500	Fourche Maline near Red Oak OK	316	541	1182	1170	1.01	677
14	7300500	Salt Fork Red River near Mungun,OK	4009	1490	565	934	0.60	42
15	7303400	Elm Fork of NF Red River near Carl OK	1077	1715	580	949	0.61	73
16	7311000	East Cache Creek near Walters OK	1748	938	777	1015	0.77	126
17	7311500	Deep Red Creek near Randlett OK	1598	924	733	947	0.77	142
18	7316500	Washita River near Cheyenne OK	2056	1901	575	984	0.58	54
19	7326000	Cobb Creek near Fort Cobb OK	795	1254	745	1042	0.72	78
20	7332500	Blue River near Blue OK	1233	504	1034	1092	0.95	302

different areas even during similar weather conditions. Secondly, hydrologic model states and volumetric soil moisture measurements may not have a one-to-one correspondence; therefore one may not be able to compare these two quantities objectively. To reduce the impacts of these issues, we will use a saturation ratio (S_R):

$$S_R = \frac{\theta - \theta_r}{\theta_s - \theta_r} \tag{1}$$

where θ is a volumetric water content (m^3 m^{-3}), θ_s is the saturation volumetric water content (m^3 m^{-3}), and θ_r is a residual volumetric water content (m^3 m^{-3}). $S_R = 0$ corresponds to dry soil conditions while $S_R = 1$ corresponds to saturation or wet soil conditions. The saturation ratio attempts to reduce the effects of the individual soil property variation for generating soil moisture maps and estimating basin averages. It should be noted that Oklahoma Mesonet soil moisture measurements were designed for drought

Please cite this article in press as: Koren, V. et al., Use of soil moisture observations to improve parameter consistency ..., J. Phys. Chem. Earth (2008), doi:10.1016/j.pce.2008.01.003

monitoring over a large area (average coverage is one site per 3000 km^2). As a result, they do not represent soil moisture variability at a hillslope-type scale and can be used as indexes of soil moisture states over mid- and large-size watersheds.

Our analyses are performed for averages of soil moisture over two soil layers: the top 0–25 cm layer, and the deeper layer (25–75 cm). The soil moisture measurements are automatically taken every 30 min, but we aggregated them into daily average values. For each layer, point saturation ratio values are interpolated to 4 km grid cells for the entire Oklahoma state using an inverse distance weighting method. Weights are computed on a daily basis depending on station locations with available data at a given day. Later, the gridded daily maps of S_R are used to generate daily time series of basin average soil moisture for the period from January 1997 till December 2002.

3. Methodology

3.1. The model

We use the SAC-HT model in our analysis. The SAC-HT is an extension of the Sacramento Soil Moisture Accounting (SAC-SMA) model (Burnash, 1995) that allows linking conceptual storage-type states to soil moisture states at a soil profile. A description of the SAC-HT model can be found in Koren (2006) and Koren et al. (2006). The SAC-HT has its origins in work performed by Koren et al. (1999) in which the land surface component of a numerical weather prediction model was heavily modified for cold season effects. Since the structure of SAC-HT is based on the SAC-SMA with the addition of two physically based parameters, the remaining parameters of the SAC-HT model are exactly same as SAC-SMA (see Table 2 for the SAC-HT parameters list).

Koren et al. (2003) developed a set of physical relationships that link the SAC-SMA parameters to soil properties such as porosity, field capacity, wilting point, and hydraulic conductivity (note that these relationships are valid for the SAC-HT too). They assume that tension water storages are related to available soil water, and that free water storages are related to gravitational soil water. These relationships allow recalculating the upper and lower soil moisture capacities into soil moisture contents at a number of soil layers which are the soil moisture states of SAC-HT. Five soil layers are predefined to cover a 2 m soil profile with thinner layers closer to the soil surface. However, the actual number of soil layers and their thicknesses are automatically adjusted using actual SAC-HT parameter-values. Because of this, the number of soil layers may be less than five and can be different for different watersheds. For more detail on this procedure see Koren et al. (2002) and Koren (2006). At each time step, the liquid water storage changes due to rainfall are computed, and then transformed into soil moisture states. The heat transfer component calculates the temperature of each soil layer. Consequently, based on the simulated soil temperature profile, the total water content is split into frozen and liquid water portions. Estimated new soil moisture states are then converted back into model storages.

Although there are strong physical arguments to support the SAC-HT model, its 18 parameters (Table 2) derived from the procedures described above or from traditional hydrograph analyses require further calibration for optimal results. For this, well defined manual and automatic calibration procedures for lumped model applications are available (e.g., Smith et al., 2003).

The total runoff output from SAC-HT is routed downstream using a simple unit hydrograph (UH) technique derived from Clark's time–area method (Kull and Feldman, 1998). With readily available DEM and GIS pack-

Table 2
SAC-HT model parameters and their feasible ranges or default values for the last five parameters

No.	Parameter	Description	Ranges or default value
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
12	PCTIM	Permanent impervious area fraction	0.001
13	ADIMP	Maximum fraction of an additional impervious area due to saturation	0.0
14	RIVA	Riparian vegetation area fraction	0.001
15	SIDE	Ratio of deep percolation from lower layer free water storages	0.3
16	RSERV	Fraction of lower layer free water not transferable to lower layer tension water	0.0
17	STXT	Soil texture of the upper layer	
18	TBOT	Climatological annual air temperature	

Parameters calibrated in options 1 and 2 are highlighted.

ages, the time–area histogram can be derived for each test basin.

3.2. The tests performed

The SAC-HT was applied in a lumped approach to the 20 test basins. Input data and model parameters are aggregated over each selected watershed. All model-simulations are performed at the 1 h time step. We define three test cases using soil moisture and hydrograph data: (1) control run with no parameter calibration ('control run'), (2) model calibration run to fit only the outlet hydrograph ('option 1'), and (3) model calibration run to fit both the outlet hydrograph and basin average soil moisture at two depths ('option 2').

3.2.1. Control run parameters

A priori SAC-HT parameter estimates are used in the control run. Koren et al. (2003) generated *a priori* grids of 11 major SAC-HT parameters (highlighted in Table 2) covering the conterminous US at 1 and 4 km resolution. The control run parameters for the tested watersheds are derived from these 4 km resolution grids by a simple arithmetic averaging. Control values of the other five minor parameters are defined as recommended values from manual calibration experience (see Table 2).

Two UH parameters, the overland flow lag time, t_h, and channel concentration time, t_c, are estimated from empirical relationships following Moreda et al. (2006):

$$t_h = 0.95\left(\frac{A}{L_{max}}\right)^{2/3} \quad (2)$$

$$t_c = 5.0\left(\frac{L_{max}}{\sqrt{s}}\right)^{0.5} \quad (3)$$

where A is the watershed area in mi^2, s is the main channel slope in feet per mi, and L_{max} is the distance from the outlet to the farthest contributing point of the watershed in mi. Variables t_h and t_c are in hours.

3.2.2. Parameter calibration tests

Eleven SAC-HT and two UH parameters are calibrated in options 1 and 2. The remaining five minor SAC-HT parameters are kept constant. Also, a scale factor which corrects overall biases in potential evapotranspiration is calibrated.

The soil-based estimates of the SAC-HT parameters have been used extensively with generally favorable results in various applications (Seo et al., 2003; Koren et al., 2004,2006; Reed et al., 2004; Lohmann et al., 2004). The approach taken here is to start from the *a priori* parameter estimates, and locate the nearest minimum via a pattern search technique. We utilize the Stepwise Line Search (SLS) procedure (Kuzmin et al., in preparation) that performs a successive minimization along each parameter with a fixed step size. As shown by Kuzmin et al. (in preparation), this procedure, if started from the soil-based parameter estimates, is very efficient and provides more consistent results on independent data sets comparing to the global SCE-UA algorithm.

We define the optimization objective function as the summation of a number of specific goodness-of-fit measures:

$$F = \sqrt{\sum_{i=1}^{n}\left(\frac{\sigma_1}{\sigma_i}f_i\right)^2} \quad (4)$$

where σ_i is the standard deviation of i-th tested variable, σ_1 is the standard deviation of a variable normalized to, the first tested variable in this case, n is the number of tested variables, and f_i is i-th specific measure; the root mean square error (RMSE) is used in this study for all variables. Note that the weight associated with each goodness-of-fit measure is given by the inverse of the standard deviation of the respective variables. This weighting scheme assumes that the uncertainty in each measure is proportional to the variability of the related property.

In option 1, RMSE values are estimated at four different time scales: 1 h, one day, 10 days, and one month, resulting in the objective function F_Q:

$$F_Q = \sqrt{\left[\sum_{i=1}^{4}\left(\frac{\sigma_{Q,1}}{\sigma_{Q,i}}\right)^2 \frac{1}{M_i}\sum_{j=1}^{M_{Q,j}}\left(Q_{s,ij} - Q_{o,ij}\right)^2\right]} \quad (5)$$

where $Q_{s,ij}$ and $Q_{o,ij}$ are simulated and observed outlet hydrograph ordinates averaged over the i-th time interval, and $M_{Q,j}$ is the number of hydrograph ordinates at i-th averaging interval.

In option 2, we try to use soil moisture measurements at two layers in addition to outlet hydrographs. The objective function in Eq. (5) is thus transformed to

$$F_{combined} = \sqrt{\left[F_Q^2 + \sum_{i=1}^{2}\left(\frac{\sigma_{Q,1}}{\sigma_{S_R,i}}\right)^2 \frac{1}{M}\sum_{j=1}^{M_{s,j}}\left(S_{Rs,ij} - S_{Ro,ij}\right)^2\right]} \quad (6)$$

where $M_{S,j}$ is the number of daily soil moisture measurements at each layer, $S_{Rs,ij}$ and $S_{Ro,ij}$ are simulated and observed soil moisture at time j, and index $i = 1$ refers to the upper soil layer and $i = 2$ refers to the lower layer. Note that $\sigma_{Q,1}$ is the same value in Eqs. (5) and (6). Selection of the soil moisture uncertainty measure is critical. Considering that Oklahoma Mesonet measurement coverage is about one site per basin, one can use the standard deviation as the uncertainty measure of basin average soil moisture. As mentioned in Section 2 Oklahoma Mesonet measurements underestimate soil moisture variability by a factor of 1.8. To account for the underestimation of variability, standard deviation, $\sigma_{Sr,i}$, of soil moisture estimated from measured time series is increased by 1.8.

Please cite this article in press as: Koren, V. et al., Use of soil moisture observations to improve parameter consistency ..., J. Phys. Chem. Earth (2008), doi:10.1016/j.pce.2008.01.003

3.3. Parameter consistency test

Parameter consistency (uncertainty) from calibration is more critical than achieving a minimum value of the optimization criteria (e.g., Bastidas et al., 2003; Seibert and McDonnell, 2003). To analyse parameter consistency, a traditional split-sample calibration test is conducted for two watersheds. Due to the shortness of the overlapping radar-based precipitation and soil moisture data sets, validation over a long period could not be performed. As an alternative, we carry out a cross validation test in which each year in the six-year period is withheld from calibration (note that the year 2003 is excluded from validation because of missing soil moisture measurements). As a result, six calibrated parameter sets for each calibration option are generated. A normalized parameter variation from these cross validation tests for the two calibration options is used as the measure of parameter consistency:

$$V_i = \frac{\frac{1}{p}\sum_{j=1}^{p}\left|x_{i,j} - x_{\mathrm{avg},i}\right|}{x_{\max,i} - x_{\min,i}}, \quad i = 1, 2, \ldots N \qquad (7)$$

where $x_{i,j}$ is i-th calibrated parameter from the j-th split-sample test, $x_{\mathrm{avg},i}$ is the average value of the i-th calibrated parameter, $x_{\max,i}$, and $x_{\min,i}$, are the feasible maximum and minimum values of i-th parameter from Table 2, respectively, N is the number of calibrated parameters, and p is the number of split-sample tests. The overall measure of parameter consistency can be defined as an average value of the individual parameter measures.

4. Results and discussion

4.1. Parameter calibration tests

First, we compare overall water balance simulations when calibration was performed for the entire 1997–2003 period. As can be seen in Fig. 2a, there is good agreement between simulated and observed annual runoff averaged over the seven year period from the control, option 1, and 2 runs. Both simulated and observed runoff values display similar dependencies on the P/PE climate index (not shown) with much higher values for the wettest watersheds. Regarding soil moisture simulations, similar accuracy of the soil moisture saturation is achieved only from option 2 as shown in Fig. 2b and c. Soil moisture simulations from the control run and option 1 calibration are less accurate with considerable positive biases in the upper soil layer from option 1, and negative bias in the lower soil layer from both runs. Simulated soil moisture also reproduces reasonably well the dependency on the P/PE climate index.

Hourly runoff prediction is less accurate as shown in Fig. 3a. For predicted runoff, there is a trend for correlation (R) to decrease as watersheds become drier (decreasing P/PE), most notably for the control run. Both options of parameter calibration considerably improve hourly runoff prediction compared to the control run as seen in Table 2. Similar improvement is achieved for both RMSE and R^2: 45.5% and 38.0%, respectively from option 1, and 42.5% and 34.0%, respectively from option 2. It can be seen that the use of soil moisture data in the calibration process reduces the overall accuracy of hourly runoff prediction by 5.2% in RMSE and 1.0% in R^2. However, considering the uncertainty in rainfall and discharge measurements, such a reduction may be acceptable if it leads to more reliable parameters. In other words, we argue that a slight reduction in hydrograph simulation accuracy could be acceptable in exchange for having more confidence that the model robustly represents a broader set of watershed processes (i.e., runoff and soil moisture).

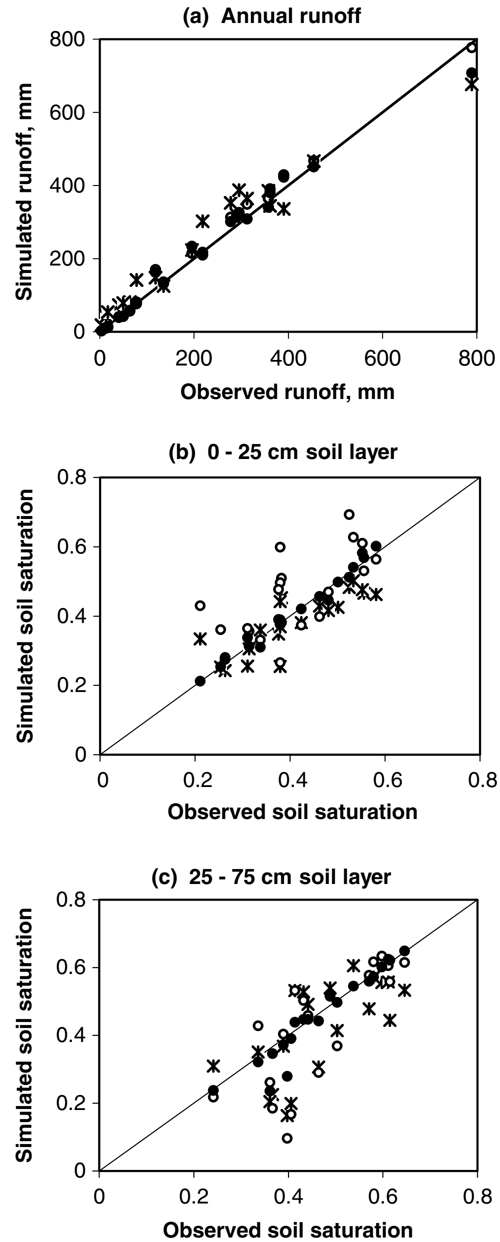

Fig. 2. Observed average annual runoff (a) and soil moisture (b, c) compared to simulated from the control run (star), calibration option 1 (open circle), and calibration option 2 (filled circle) for test watersheds.

Please cite this article in press as: Koren, V. et al., Use of soil moisture observations to improve parameter consistency ..., J. Phys. Chem. Earth (2008), doi:10.1016/j.pce.2008.01.003

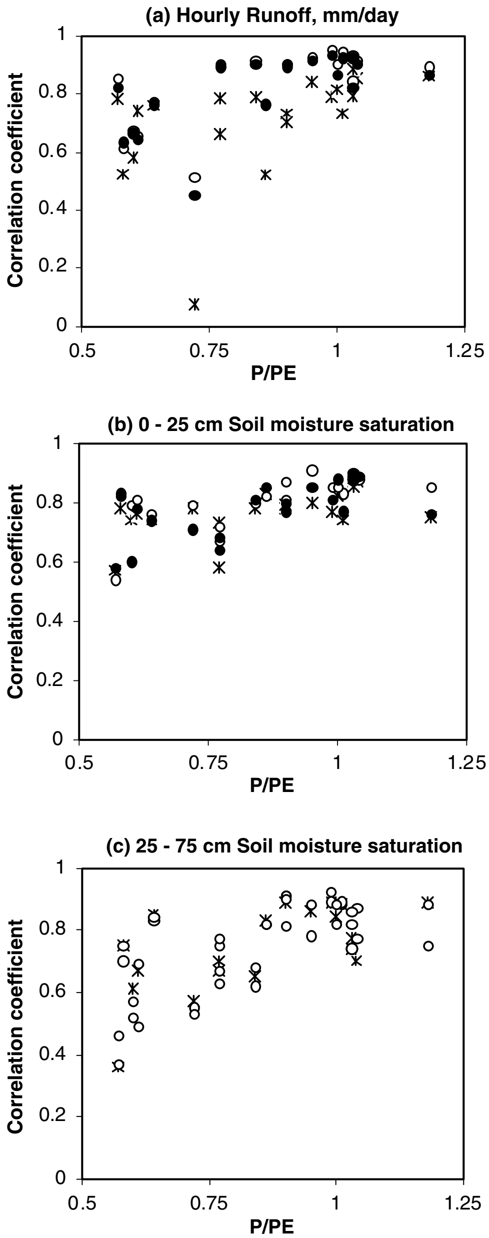

Fig. 3. Correlation coefficient of hourly runoff (a) and daily soil moisture (b, c) saturation from the control run (star), and calibration options 1 (open circle) and 2 (filled circle) vs. P/PE ratio.

As can be seen from Fig. 3b and c, and Table 2, parameter calibration has different effects on soil moisture simulation results. While option 1 leads to some improvement in the correlation of the upper layer soil moisture compared to measured values, the RMSE values degrade by 36.4%. The overall statistics of the lower layer soil moisture from option 1 are close to the control run results although there are a few outliers in terms of RMSE. These results suggest that the use of outlet discharge alone as a goodness-of-fit measure can lead to considerable biases in soil moisture while preserving soil moisture dynamics accuracy at about the same level. Calibration option 2 removes mostly biases in the upper and lower layers soil moisture without noticeable changes in soil moisture dynamics. This observation can be seen also in Fig. 4, which compares observed and simulated time series of monthly runoff and soil moisture saturation for the Cobb Creek watershed located in the dry western part of the study domain ($P/PE = 0.72$). All simulation options including the control run reproduce reasonably well monthly and seasonal soil moisture dynamics. However, the control run considerably overestimates the upper layer soil moisture during wet seasons. Parameter calibration using only discharge goodness-of-fit improves runoff and lower layer soil moisture time series while generating even higher biases in the upper layer soil moisture. The addition of soil moisture measurements in the calibration process (option 2) consistently improves soil moisture simulations for both soil layers with only a minimal reduction in runoff accuracy.

Different behavior can be seen in the Baron Fork basin located in the wetter eastern part of the region ($P/PE = 1.04$). Fig. 5 displays similar levels of improvement in simulations of soil moisture and runoff for the both calibration options 1 and 2. For this basin, the use of soil moisture measurements does not lead to noticeable differences in soil moisture and runoff simulations from the two calibration options. Similar results were obtained for the most basins in the wetter eastern part of the region where soil-based parameters provide reasonable soil moisture simulation results prior to calibration. The main reason of this behavior is much higher correlation between runoff and soil moisture for 'wet' basins compared to 'dry' one. More discussion on this is in the next section (see Table 3).

In general, the SAC-HT control run parameters as well as parameters from calibration options 1 and 2 averaged over all tested watersheds do not differ very much as shown in Table 4, with the exception of ZPERC, which controls percolation into the lower zone. However, considerable differences may be observed for some watersheds, especially in the dry western part of the region. Rather strong correlation exists between option 2 parameters and the control parameters for the majority parameters, Table 4. On the other hand, parameters from the discharge-based only calibration (option 1) are less correlated with the control parameters. Surprisingly, correlation between option 1 and 2 parameters is lower than correlation between control parameters and option 2 parameters. One of the reasons of this behavior could be higher uncertainties in option 1 calibration because it relies only on a basin aggregated outlet hydrograph.

4.2. Parameter consistency tests

Here we perform tests on two watersheds: Baron Fork located in the wet region and Cobb Creek located in the dry region of the study domain. The Baron Fork watershed represents cases in which both calibration options 1 and 2 produce accurate simulations of outlet hydrographs and soil moisture as shown in Fig. 5. On the other hand, the

Please cite this article in press as: Koren, V. et al., Use of soil moisture observations to improve parameter consistency ..., J. Phys. Chem. Earth (2008), doi:10.1016/j.pce.2008.01.003

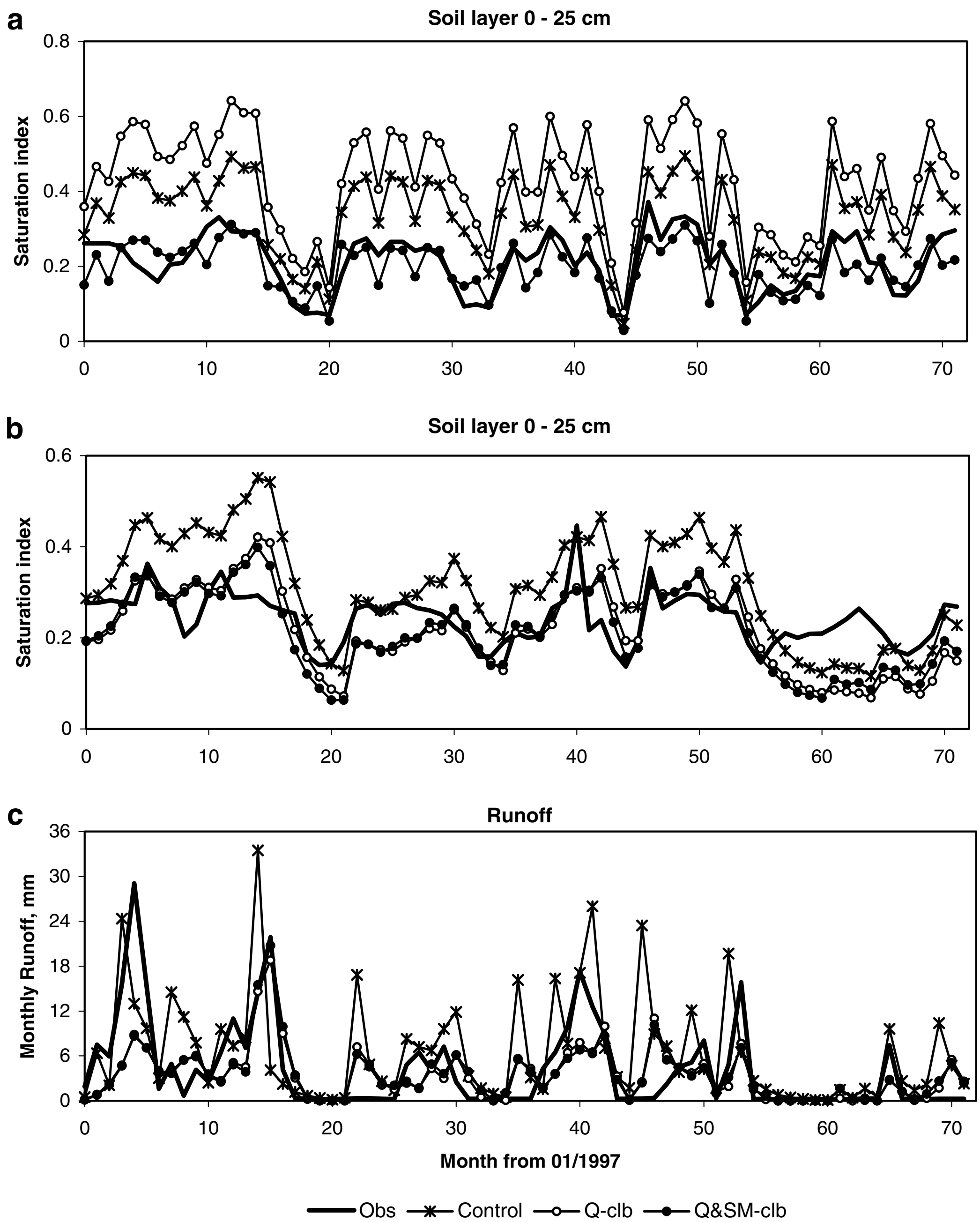

Fig. 4. Monthly runoff and soil moisture saturation time series for the Cobb Creek watershed (dry area, $P/PE = 0.72$) generated from the control run, and calibration options 1 (Q-clb) and 2 (Q&SM-clb).

Cobb Creek watershed represents cases in which calibration option 1 leads to large biases in soil moisture results compared to those from option 2 (Fig. 4). Cobb Creek and Baron Fork consistency test results are plotted in Figs. 6a and b and 7a and b, respectively. Fig. 6a shows that calibration option 1 for the Cobb Creek watershed generates wide spread in SAC-HT parameters from six slightly different input data sets (recall that each data set is created by removing one year from the seven year total period span). In addition, the values of the evaluation criteria also vary considerably, specifically the daily root mean square errors of soil moisture for both upper and lower layers. On the other hand, the addition of soil moisture measurements for calibration (option 2) reduces parameter and criteria value spread as shown in Fig. 6b. Moreover, overall parameter consistency is substantially improved, with consistency measure (Eq. (7)) from 0.105 to 0.045. Soil moisture simulation accuracy also improved compared to option 1. Here the six-run average root mean square error for the upper and lower layer soil moisture reduces to 0.067 and 0.070, respectively, compared to corresponding values of 0.385 and 0.202 from option 1. At the same time, we notice that hourly runoff accuracy degrades very slightly from 2.33 to 2.35 cm in terms of average RMSE. These

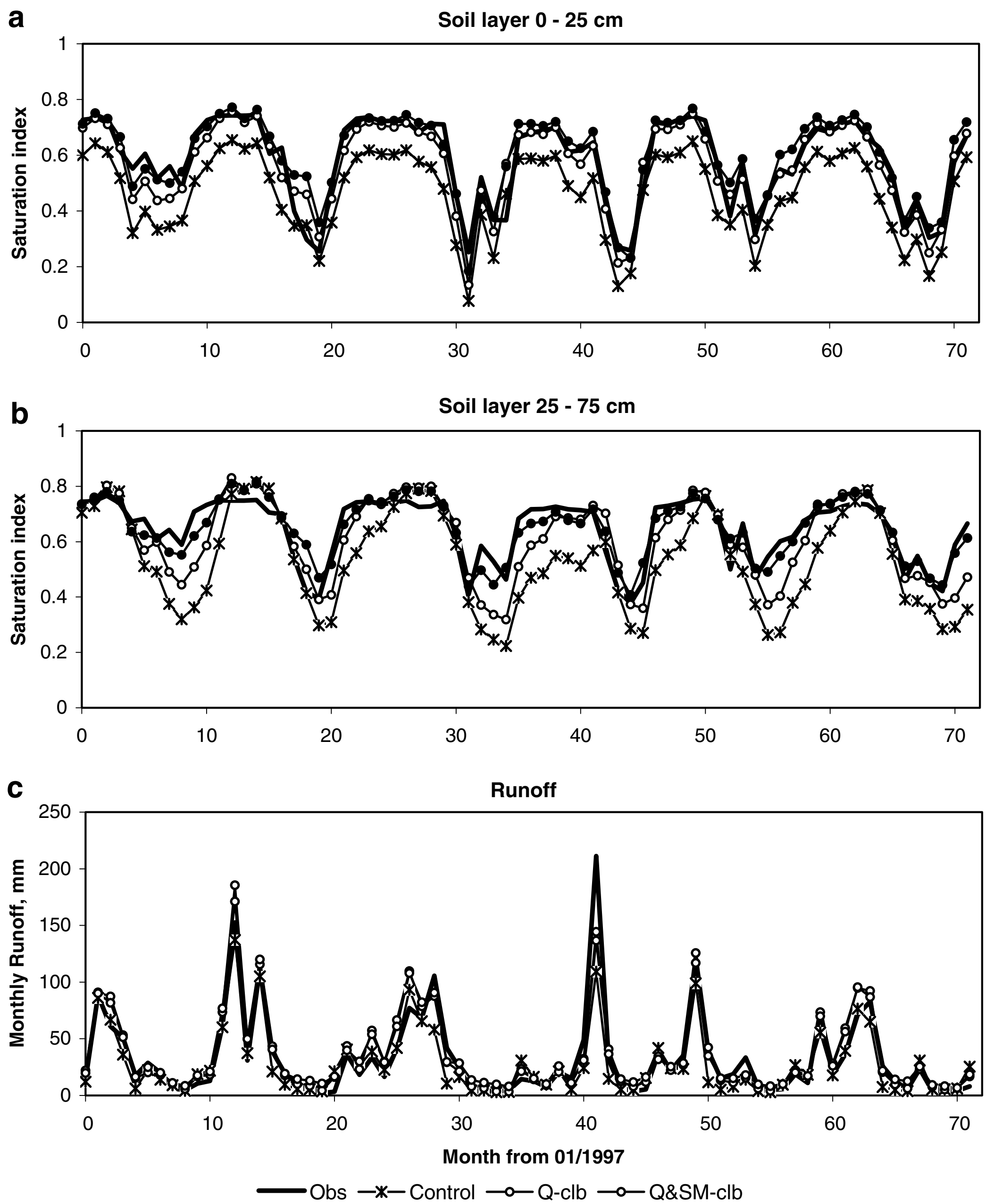

Fig. 5. Monthly runoff and soil moisture saturation time series for the Baron Fork watershed (wet area, $P/PE = 1.04$) generated from the control run, and calibration options 1 (Q-clb) and 2 (Q&SM-clb).

Table 3
Overall RMSE and coefficient of determination (R^2) averaged over all tested watersheds from control and two calibration option runs

Statistics	Control run	Calibration option 1	Calibration option 2	% Improvement		
				Option 1 vs. control	Option 2 vs. control	Option 2 vs. option 1
Discharge statistics						
RMSE	18.24	9.94	10.49	45.5	42.5	−5.2
R^2	0.50	0.69	0.67	38.0	34.0	−1.0
0–25 cm layer soil moisture statistics						
RMSE	0.11	0.15	0.09	−36.4	18.2	40.0
R^2	0.59	0.65	0.60	10.2	1.7	−7.7
25–75 cm layer soil moisture statistics						
RMSE	0.15	0.15	0.09	0.0	40.0	40.0
R^2	0.56	0.54	0.58	−3.6	3.6	7.4

Table 4
SAC-HT parameters averaged over all tested basins, and correlation coefficients between calibrated and control parameters as well as between two calibration options

SAC-HT Parameter	Twenty watershed average parameter-value			Twenty watershed correlation coefficient		
	Control	Calibration option 1	Calibration option 2	Option 1 vs. control	Option 2 vs. control	Option 1 vs. option2
UZTWM	53.4	86.7	65.0	−0.06	0.70	0.13
UZFWM	40.7	38.1	44.1	0.01	0.73	0.11
UZK	0.43	0.41	0.44	0.33	−0.27	0.03
ZPERC	54.1	123.0	155.3	−0.21	0.03	0.50
REXP	2.47	2.21	2.48	0.52	0.73	0.72
LZTWM	181	221	240	−0.05	0.80	−0.03
LZFSM	28.4	38.0	36.6	0.10	0.47	0.32
LZFPM	81.2	127.3	107.7	0.87	0.25	0.57
LZSK	0.17	0.18	0.19	−0.10	0.21	0.65
LZPK	0.01	0.01	0.01	0.68	0.51	0.49
PFREE	0.34	0.28	0.23	0.91	0.86	0.86
Average				0.27	0.45	0.40

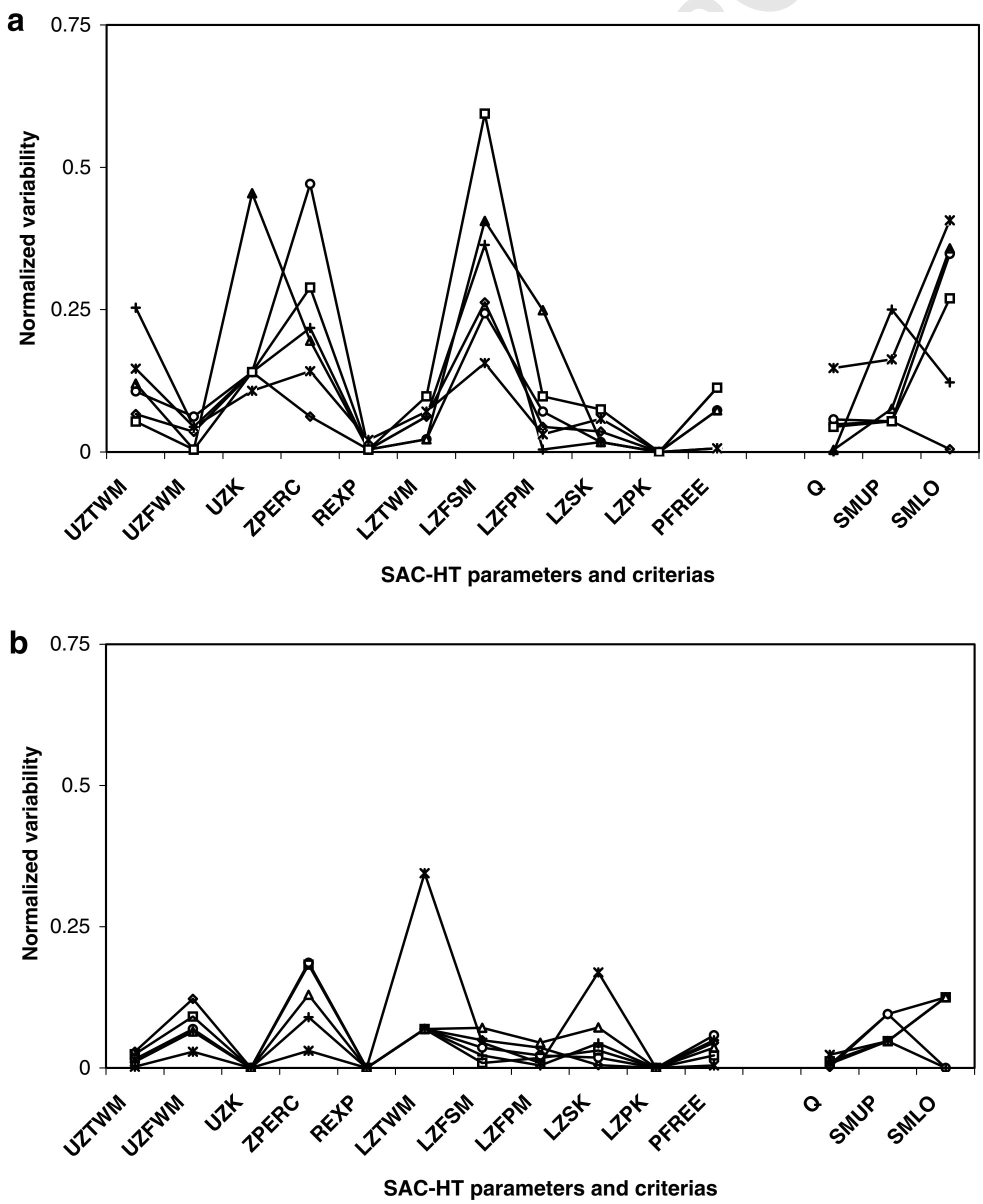

Fig. 6. Normalized variability of calibrated parameters and root mean square errors of hourly discharge (Q) and daily soil moisture of the upper (SMUP) and lower (SMLO) layers for the Cobb Creek watershed: (a) option 1 calibration, (b) option 2 calibration. Each line represents calibration results from one selected data set in split-sample tests.

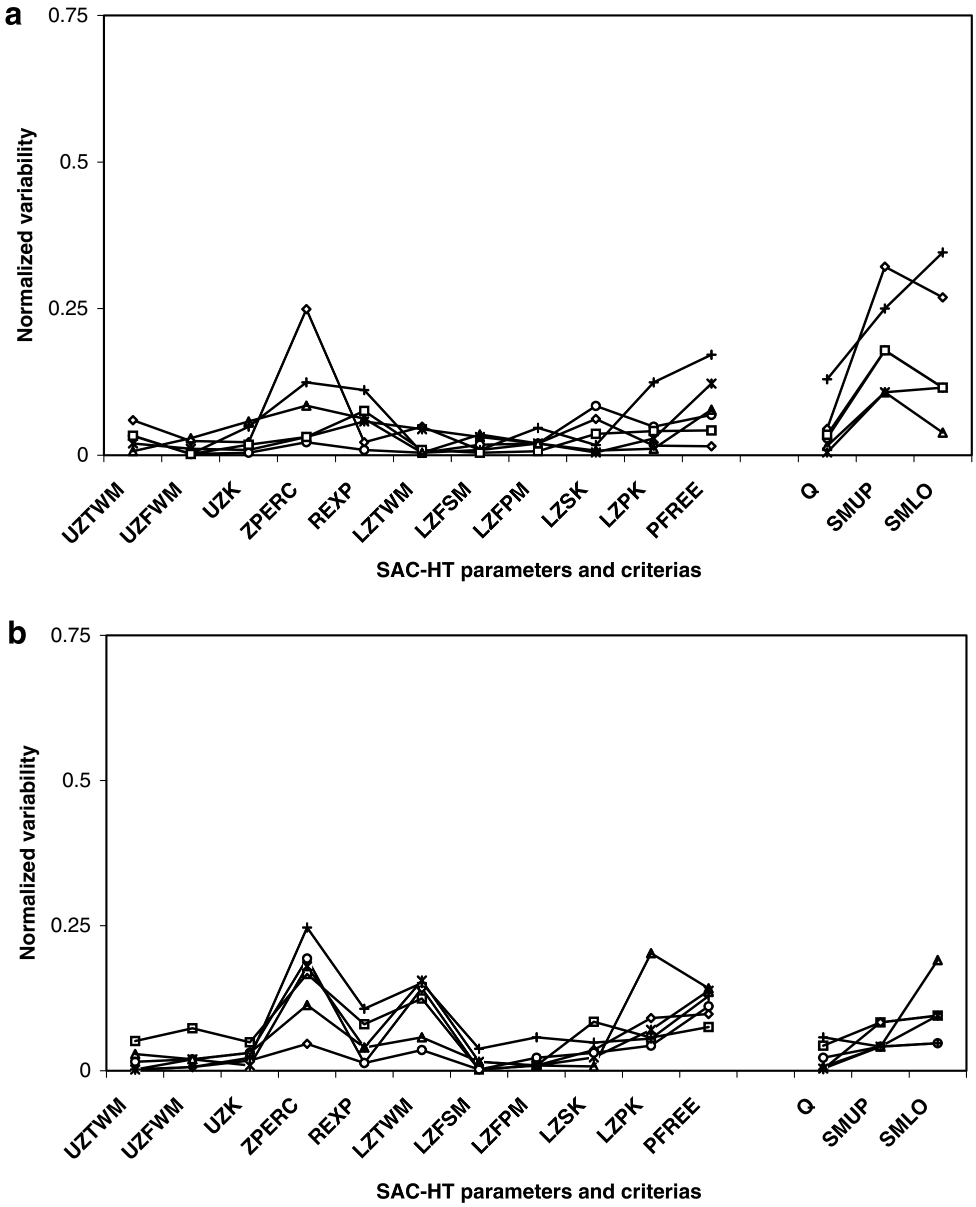

Fig. 7. Normalized variability of calibrated parameters and root mean square errors of hourly discharge (Q) and daily soil moisture of the upper (SMUP) and lower (SMLO) layers for the Baron Fork watershed: (a) option 1 calibration, (b) option 2 calibration. Each line represents calibration results from one selected data set in split-sample tests.

results suggest that the use of outlet response data alone in the model evaluation may lead to unreliable parameter sets while providing reasonable accuracy of the selected variable for the calibration period.

The Baron Fork watershed results (Fig. 7) are somewhat different. There is not much difference in parameter spread from the option 1 and 2 calibrations. A little bit more variability is observed in runoff and soil moisture accuracy from option 1. The model parameters vary much less in split-sample tests for this ‘wet’ watershed compared to the previous ‘dry’ watershed. This behavior can be expected considering close runoff and soil moisture simulation results from the basic calibration of option 1 and 2 (Fig. 5). The possible reason for this may be much higher correlation between runoff and soil moisture states for ‘wet’ watersheds compared to ‘dry’ watersheds. As a result, outlet runoff may be informative enough to derive physically consistent parameters.

5. Summary

The SAC-HT model driven by *a priori* parameters performs reasonably well on the water balance and allows explicit estimation of soil moisture at desired layers. Annual runoff and soil moisture agrees well with observed data for a range of spatial scales. However, deeper layer (25–75 cm) soil moisture has a negative bias for watersheds located in the dry western part of the region with a climatological index (P/PE) of less than 0.75.

Higher time resolution predictions with *a priori* parameters are less accurate. While hourly runoff and daily soil moisture dynamics are consistent with measurements (correlation coefficients are above 0.5 for most watersheds with the average values 0.71 and 0.77 for runoff and soil moisture, respectively), considerable biases in amplitude and timing are common for some watersheds. Automatic calibration based solely on outlet hydrograph goodness-of-fit improves runoff simulation results on average by 45%. However, it reduces the accuracy of soil moisture simulation in the top soil layer by 36% compared to *a priori* parameter simulations.

The use of soil moisture measurements in the calibration process reduces soil moisture simulation RMSE in both soil layers by 40% with only a marginal reduction of 5% in runoff accuracy. However, we note that the selected uncertainty level of soil moisture measurements in Eq. (6) can considerably affect calibration results. For example, selection of an uncertainty level below the soil moisture variability from measured data can lead to the reduction in runoff accuracy without measurable improvement in soil moisture results.

There is a tendency to improve internal consistency of calibration when soil moisture data are used. It is more noticeable for 'dry' watersheds where there is no strong direct interconnection between runoff and soil moisture.

Our study highlights deficiencies in model calibration that is based solely on outlet hydrograph goodness-of-fit and points to the need for ingesting additional information such as soil moisture data in the calibration process. This study uses comprehensive soil moisture data which are not commonly available. While there is a hope that improved satellite and surface observation techniques will provide comprehensive and reliable soil moisture information on a broader geographic scale, more analysis should be performed on the specifics of this information such as space-depth-time representation to realize the maximum benefit in practical applications.

Acknowledgments

We acknowledge the Oklahoma Mesonet team for providing us with comprehensive soil moisture data over the study region.

References

Ambroise, B., Perrin, J.L., Reutenauer, D., 1995. Multicriterion validations of a semidistributed conceptual model of the water cycle in the Fecht catchment (Vosges, Massif. France). Water Resources Research 31, 1467–1481.

Bastidas, L.A., Gupta, H.V., Hsu, K-l., Sorooshian, S., 2003. Parameter, structure, and model performance evaluation for land-surface schemes. In: Duan, Q., Sorooshian, S., Gupta, H., Rosseau, H., Turcotte, R. (Eds.), Calibration of Watershed Models, Water Science and Applications, vol. 6. AGU, pp. 239–254.

Brock, F.V., Crawford, K.C., Elliot, R.L., Cuperus, G.W., Stadler, S.J., Johnson, H., Eillts, M.D., 1995. The Oklahoma Mesonet: a technical overview. Journal of Atmospheric and Oceanic Technology 12, 5–19.

Burnash, R.J.C., 1995. The NWS river forecast system – catchment modeling. In: Singh, V.P. (Ed.), Computer Models of Watershed Hydrology. Water Resources Publications, Littleton, CO, pp. 311–366.

Casper, M., Gemmar, P., Gronz, O., Johst, M., Stuber, M., 2007. Fuzzy logic-based rainfall-runoff modeling using soil moisture measurements to represent system states. Hydrological Sciences Journal 52 (3), 478–490.

Farnsworth, R.K., Thompson, E.S., Peck, E.L., 1982. Evaporation atlas for the contiguous 48 United States. NOAA Technical Report NWS 33. US Department of Commerce, Washington, DC., p. 27.

Illston, B.G., Caldwell, J.C., Bodnar, S.G., 2004. Representativeness of soil moisture conditions in central Oklahoma during the enhanced drying phase. In: Proceedings of 15th Symposium on Global Change and Climate Variations, AMS, Seattle, Washington.

Jakeman, A.J., Hornberger, G.M., 1993. How much complexity is warranted in rainfall-runoff model? Water Resources Research 29 (8), 2637–2649.

Johnson, D., Smith, M., Koren, V., Finnerty, B., 1999. Comparing mean areal precipitation estimates from NEXRAD and rain gauge networks. Journal of Hydrologic Engineering 4 (2), 117–124.

Koren, V., 2006. Parameterization of frozen ground effects: sensitivity to soil properties. In: Prediction in Ungauged Basins: Promise and Progress, (Proceedings of Symposium S7at Foz do Iguacu, Brazil, April 2005), vol. 303. IAHS Publishers, pp. 125–133.

Koren, V., Schaake, J., Mitchell, K., Duan, Q.-Y., Chen, F., Baker, J.M., 1999. A parameterization of snowpack and frozen ground intended for NCEP weather and climate models. Journal of Geophysical Research 104 (D16), 19569–19585.

Koren, V., Smith, M., Duan, Q., 2003. Use of *a priori* parameter estimates in the derivation of spatially consistent parameter sets of rainfall-runoff models. In: Duan, Q., Gupta, H.V., Sorooshian, S., Rousseau, A.N., Turcotte, R. (Eds.), Calibration Watershed Models, Water Science and Applications, vol. 6. AGU, pp. 239–254.

Koren, V., Reed, S., Smith, M., Zhang, Z., Seo, D.J., 2004. Hydrology laboratory research modeling system (HL-RMS) of the national weather service. Journal of Hydrology 291, 297–318.

Koren, V., Fekadu, M., Reed, S., Smith, M., Zhang, Z., 2006. Evaluation of grid-based distributed hydrological model over a large area. In: Prediction in Ungauged Basins: Promise and Progress (Proceedings of Symposium S7at Foz do Iguacu, Brazil, April 2005), vol. 303. IAHS Publishers, pp. 47–56.

Kuczera, G., Mroczkowski, M., 1998. Assessment of hydrological parameter uncertainty and the worth of multi-response data. Water Resources Research 34, 1481–1489.

Kull, D.W., Feldman, A.D., 1998. Evolution of Clark's unit graph method to spatially distributed runoff. Journal of Hydrologic Engineering 3 (1), 9–19.

Kuzmin, V., Seo, D.J., Koren, V., in preparation. Fast and efficient optimization of hydrologic model parameters using *a priori* estimates and stepwise line search. Q3

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. Journal of Geophysical Research 109 (D07S91), 22.

Moreda, F., Koren, V., Zhang, Z., Reed, S., Smith, M., 2006. Parameterization of distributed hydrological models: learning from the experiences of lumped modeling. Journal of Hydrology 320, 218–237.

Reed, S., Koren, V., Smith, M., Zhang, Z., Moreda, F., Seo, D.J., 2004. Overall distributed model intercomparison project results. Journal of Hydrology 298, 27–60.

Please cite this article in press as: Koren, V. et al., Use of soil moisture observations to improve parameter consistency ..., J. Phys. Chem. Earth (2008), doi:10.1016/j.pce.2008.01.003

Refsgaard, J.C., 1997. Parameterization, calibration and validation of distributed hydrological models. Journal of Hydrology 198, 69–97.

Seibert, J., McDonnell, J.J., 2003. The quest for an improved dialog between modeler and experimentalist. In: Duan, Q., Sorooshian, S., Gupta, H., Rosseau, H., Turcotte, R. (Eds.), Calibration of Watershed Models, Water Science and Applications, vol. 6. AGU, pp. 301–316.

Seo, D.J., Koren, V., Cajina, N., 2003. Real-time variational assimilation of hydrologic and hydrometeorological data into operational hydrologic forecasting. Journal of Hydrometeorology 4, 627–641.

Smith, M.B., Laurine, D., Koren, V., Reed, S., Zhang, Z., 2003. Hydrologic model calibration in the national weather service. In: Duan, Q., Gupta, H.V., Sorooshian, S., Rousseau, A.N., Turcotte, R. (Eds.), Calibration of Watershed Models, Water Science and Applications, vol. 6. AGU, pp. 133–152.

Young, C.B., Bradley, A.A., Krajewski, W.F., Kruger, A., 2000. Evaluating NEXRAD multisensor precipitation estimates for operational hydrologic forecasting. Journal of Hydrometeorology 1, 241–254.

Please cite this article in press as: Koren, V. et al., Use of soil moisture observations to improve parameter consistency ..., J. Phys. Chem. Earth (2008), doi:10.1016/j.pce.2008.01.003