Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Hydrology 320 (2006) 103–116
Journal of Hydrology
www.elsevier.com/locate/jhydrol

Using SSURGO data to improve Sacramento Model a priori parameter estimates

Richard M. Anderson[1,*], Victor I. Koren, Seann M. Reed

NOAA, National Weather Service, Office of Hydrologic Development, Hydrology Laboratory, 1325 East West Highway, Silver Spring, MD 20912, USA

Abstract

As it transitions to smaller scale, distributed hydrologic modeling approaches, the National Weather Service (NWS) is improving methods of estimating parameters for the Sacramento Soil Moisture Accounting model (SAC-SMA). This is the major hydrologic model used for flood forecasting at most of the 13 river forecasting centers throughout the United States. A physically based approach based on the nationally available State Soil Geographic Database (STATSGO) has been developed (Koren, V.I., Smith, M., Wang, D., Zhang, Z., 2000. Use of soil property data in the derivation of conceptual rainfall–runoff model parameters. Proceedings of the 15th Conference on Hydrology, AMS, Long Beach, CA, pp. 103–106; 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., Sorooshian, S., Gupta, H., Rosseau, H., Turcotte, H. (Eds.), Calibration of Watershed Models, Water Science and Applications 6, AGU, pp. 239–254), leading to objective, spatially consistent parameter estimates. This paper shows that a better representation of basin physical properties and potential improvements in hydrologic simulation performance can be obtained by basing parameter estimates on a finer-scale database of soils data, the Soil Survey Geographic Database (SSURGO), combined with high-resolution land use/land cover data. Results also suggest that an intermediate level of improvement may be obtained by combining detailed land cover data with STATSGO to refine current parameter estimates. This latter finding is significant because the SSURGO data are not yet available for the entire country.
Published by Elsevier B.V.

Keywords: SSURGO data; STATSGO data; A priori parameter estimation; Sacramento hydrologic model; Calibration; Distributed modeling; Flash flood forcasting

1. Introduction

The National Weather Service (NWS) is in the process of transitioning to smaller scale, distributed hydrologic modeling approaches to improve its river and flash flood forecasting capabilities for the nation. Part of this effort is improving a priori estimation of model parameters for the Sacramento Soil Moisture Accounting model (SAC-SMA; Burnash, 1995), the major hydrologic model used for flood forecasting at river forecasting centers throughout the United States. Initial research within the Hydrology Laboratory of the NWS Office of Hydrologic Development has developed a groundbreaking approach that uses

* Corresponding author.
E-mail address: randerso@oce.usda.gov (R.M. Anderson).
[1] US Department of Agriculture, Office of Risk Assessment and Cost Benefit Analysis, 1400 Independence Avenue, SW, Washington, DC 20250, USA.

0022-1694/$ - see front matter Published by Elsevier B.V.
doi:10.1016/j.jhydrol.2005.07.020

physical relationships to derive 11 major parameters of SAC-SMA based on an available, nationwide database of Geographic Information System (GIS)-compatible soils property data, the State Soil Geographic Database (STATSGO) (Koren et al., 2000, 2003).

This paper attempts to improve upon this initial parameter estimation methodology mainly by using the finer-scale soils data available in the Soil Survey Geographic Database (SSURGO). While not presently available as widely as STATSGO, generation of SSURGO data is ongoing rapidly. Both datasets are developed and maintained by the National Resource Conservation Service (NRCS) of the United States Department of Agriculture (USDA). Completion of SSURGO data digitizing is scheduled for 2008, and in the meantime many additional, 'unofficial', fine-scale soils data sets may be obtained, often from state natural resource agencies.

To our knowledge, only a few previous publications describe work comparing hydrologic simulations using SSURGO-based vs. STATSGO-based parameter estimates. Reed (1998) concluded that outlet streamflow simulations for the Little Washita watershed (600 km^2) in Oklahoma were not much different when using SSURGO-like data for parameter estimation when compared with STATSGO-based estimates. However, unlike some basins in this study, the overall surface, soil–texture distributions defined by SSURGO and STATSGO are similar in the Little Washita. More recently, Peschel et al. (2003) have added a SSURGO-based parameter estimation capability to the USDA's Soil and Water Assessment Tool (SWAT). Although they showed that the use of SSURGO data can produce significantly different output, they did not make any conclusions about the relative accuracy of the SSURGO-based vs. STATSGO-based results.

The present work is important because for a number of reasons there is an ongoing need to improve a priori parameter estimation procedures within NWS and in the larger hydrologic community. First, the state-of-the-art of parameter estimation within NWS and elsewhere still involves a large role of subjective, expert assessment which limits reproducibility of parameter estimates. Second, traditional parameter estimation methods that rely heavily on calibration with basin outlet streamflow data become impossible to apply as one moves to finer resolution, highly dimensioned, gridded models. Third, the availability of objectively derived parameter estimates over large areas can potentially improve our understanding of spatially variable hydrometeorological forcing errors, particularly those inherent in radar-based precipitation products.

2. Approach and methods

2.1. Strategies for improvement of SAC-SMA parameter estimation

This work is based on the premise that what is needed is an objective estimation procedure that can produce spatially consistent and physically reasonable parameter estimates. In the current NWS approach, initial estimates of model parameters are calculated based on the STATSGO soils database, thereby avoiding data quality problems of calibration based on rainfall–discharge data and resulting in improved spatial consistency (Koren et al., 2003). These improved starting point values may then be adjusted slightly, for example, by means of constrained calibration within reduced bounds to account for suboptimal magnitudes (e.g. Koren et al., 2003), or by using a manual calibration process (e.g. Smith et al., 2003). Adjustment of starting values was not performed in the studies reported in this paper.

STATSGO soils-based estimates of SAC-SMA 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 Modeling System, DHMS).

Even so, a number of important limitations of this current approach suggest that methodological improvements could be obtained. First, STATSGO dominant soils texture grids for eleven soil layers for the conterminous US are used (Miller et al., 1998). However, STATSGO data, at a typical scale of 1:250,000, is intended for multi-state and regional scale analysis. Soil polygons can be on the scale of 100–200 km^2, leading to limitations on the resolution of features (e.g. soil textures) below the scale of these polygons. This limitation becomes undesirable as one moves to smaller, distributed hydrologic modeling scales. Second, the initial national a priori parameter grids developed using the current approach did not account for the range of land cover/land uses that exist. Instead, throughout the US they assumed 'pasture or range land use' under 'fair' hydrologic conditions (Koren et al., 2003). Third, the STATSGO soil layers are defined only to a depth of 2.5 m, and thus are unable to account for areas of deep groundwater, which should affect estimates of lower soil layer storages. Finally, texture–hydrologic property relationships are subject to significant uncertainty, which has not been investigated.

In this work, the first two limitations have been directly addressed. The third limitation has been partially addressed. The fourth limitation is not addressed in this paper.

First, the soils data resolution issue is addressed by using SSURGO data, which is typically available at a scale of at least 1:24,000 (approximately 10 times the resolution of STATSGO). Second, the land cover issue is addressed by using the 1992 National Land Cover Dataset (NLCD) available from the United States Geological Survey (USGS), which contains land cover data for the conterminous US at a resolution of 30 m. Third, the issue of depth of data measurements is partially addressed from a data accuracy standpoint as follows: SSURGO data is more precise than STATSGO to begin with, and, in addition, available SSURGO data on the occurrence of various soil column restrictions (e.g. fragipan features, shallow bedrock) is used to supplement the basic soil column data.

A procedure has been developed that implements these improvements. The basic result is SSURGO-based estimates of eleven SAC-SMA parameters that also account for land cover based on the 30 m NLCD data. The first main question that was addressed in this work is the following: Does the use of SSURGO and NLCD data lead to improved hydrologic simulation performance? It was addressed based on a study of six basins within the Ohio River Forecasting Center (OHRFC) of the NWS, and included basins located in West Virginia, Ohio, Kentucky, and Virginia.

A second main question addressed is as follows. In the interim until a completed, national database of SSURGO soils data is available, can incremental improvement in hydrologic simulation performance be obtained by supplementing STATSGO data with the 30 m NLCD data? This was addressed based on analysis performed within the West Gulf River Forecasting Center (WGRFC) of the NWS. The basins studied in WGRFC were all in the state of Texas.

2.2. SAC-SMA structure and parameters

A detailed description of the SAC-SMA structure and parameters can be found in Burnash (1995). SAC-SMA represents the hydrologically active zone of the soil conceptually as two layers, a thin upper layer and usually much thicker lower layer. Each layer consists of tension and free water storages that interact to generate soil moisture states and a total of five components of runoff. The free water (fast) components are driven mostly by gravitational forces, while the tension water (slow) components are driven by evapotranspiration and diffusion. The free water storage of the lower layer is divided into two sub-storages: LZFSM, which controls supplemental (fast) baseflow, and LZFPM, which controls primary (slow) baseflow.

Partitioning of rainfall into surface runoff and infiltration is constrained by upper layer soil moisture conditions and the percolation potential of the lower layer. No surface runoff occurs before the tension water capacity of the upper layer, UZTWM, is filled. After that, surface runoff generation is controlled by the content of the upper layer free water storage, UZFWM, and the deficiency of lower layer tension water, LZTWM, and free water storages. Each free water reservoir can generate runoff depending on a depletion coefficient: UZK, for the upper layer, and LZSK and LZPK for the lower layer supplementary and primary storages, respectively. ZPERC is a ratio of the maximum and minimum percolation rates, and REXP is an exponent that governs the shape of the percolation curve. Parameter PFREE expresses the fractional split of percolated water between tension

Table 1
SAC-SMA parameters and their feasible ranges

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
12	*PCTIM*	Permanent impervious area fraction	
13	ADIMP	Maximum fraction of an additional impervious area due to saturation	
14	RIVA	Riparian vegetarian area fraction	
15	SIDE	Ratio of deep percolation from lower layer free water storages	
16	RSERV	Fraction of lower layer free water not transferable to lower layer tension water	

Italics indicate the 11 parameters estimated by our methodology.

and free water storages of the lower layer. Table 1 contains a list of all sixteen SAC-SMA parameters, and highlights the eleven mentioned here, which are estimated by our methodology.

2.3. Soil texture and SAC-SMA parameter relationships

Koren et al. (2000) assumed that tension water component storages of SAC-SMA are related to plant-extractable soil moisture, while free water components are related to gravitational soil moisture. Plant-extractable and gravitational soil moisture can be derived from soil properties such as saturated moisture content θ_s, field capacity θ_{fld}, and wilting point θ_{wp}.

Occurrence of these soil properties directly in SSURGO data is unpredictable (data records in these fields are frequently unpopulated), so the approach derived by Koren et al. (2000) is used here as well. They estimated these properties by using STATSGO dominant texture grids available for eleven soil layers (from ground surface to 2.5 m depth) for the conterminous US (Miller et al., 1998). To do so, the percentages of sand and clay were obtained from the midpoint values of each textural class using the USDA textural triangle, and these were related to soil properties using regression equations from Cosby et al. (1984). Experimental data reported by Clapp and Hornberger (1978) were used to estimate saturated hydraulic conductivity K_s, while an empirical relationship from Armstrong (1978) was used to estimate specific yield μ. These relationships are reported in Koren et al. (2003). Thus once soil texture is known, it becomes possible to determine associated physical properties. In our improved methodology, more precise and, we propose, more accurate texture data are obtained from the SSURGO database. The physical properties are reported for each texture class in Table 2.

To relate SAC-SMA parameters to soil-derived physical properties, model component storages expressed in water depth are converted to actual depths within the soil profile. Depth of the entire soil profile, $Z_{\max}$, is obtained from SSURGO data and estimated as the combined depth of both SAC-SMA soil layers. The split between upper and lower soil layers, Z_{up}, is determined by making use of the

Table 2
The physical properties corresponding to the 12 basic USDA soil textures

Texture	Sand (%)	Clay (%)	θ_s	θ_{fld}	θ_{wp}	K_s (mm/h)	μ
S	92	3	0.37	0.15	0.04	634.6	0.29
LS	82	6	0.39	0.19	0.05	562.6	0.23
SL	58	10	0.42	0.27	0.09	124.8	0.15
SIL	17	13	0.47	0.35	0.15	25.9	0.10
SI	9	5	0.48	0.34	0.11	20.0	0.12
L	43	18	0.44	0.30	0.14	25.0	0.13
SCL	58	27	0.42	0.29	0.16	22.7	0.12
SICL	10	34	0.48	0.41	0.24	6.1	0.04
CL	32	34	0.45	0.36	0.21	8.8	0.07
SC	52	42	0.42	0.33	0.21	7.8	0.07
SIC	6	47	0.48	0.43	0.28	3.7	0.02
C	22	58	0.46	0.40	0.28	4.6	0.03

concept of an initial rain abstraction from the curve number method developed by the Natural Resources Conservation Service (NRCS) (McCuen, 1998). In the curve number method, the amount of initial rainfall that does not reach the stream channel (the initial abstraction) is estimated as a function of soil and vegetation type, as well as hydrologic condition and antecedent moisture status. Koren et al. (2000) assumed an antecedent soil moisture condition in which the SAC-SMA upper layer tension water storage is full and the free water storage is empty. In this case, the initial rain abstraction should fill the upper layer free water storage, leading to an equation relating UZFWM to Z_{up}. Thus,

$$\mathrm{UZFWM} = (\theta_s - \theta_{fld})Z_{up} \tag{1}$$

in terms of millimeters of water depth, where, based on the curve number method

$$Z_{up} = 5.08 \times \frac{1000/\mathrm{CN} - 10}{\theta_s - \theta_{fld}} \tag{2}$$

in terms of millimeters of soil depth. In (2) the numerator of the expression for Z_{up} expresses the initial abstraction in water depth, while the denominator with volumetric units of mm water/mm soil converts the final expression into depth within the soil profile. Koren et al. (2000) showed how these assumptions lead to estimates for other SAC-SMA storages UZTWM, LZTWM, LZFSM, and LZFPM in terms of soil porosity, field capacity, and wilting point. Darcy's equation for an unconfined, homogenous aquifer was used to develop an expression for lower layer depletion rate, LZPK (Dingman, 2002). ZPERC was estimated from other known SAC-SMA parameters, assuming that maximum percolation occurs when the upper layer is fully saturated and the lower layer is dry. Estimates for UZK, LZSK, REXP, and PFREE relied upon using ratios of field capacity (θ_{fld}/θ_s) and wilting point (θ_{wp}/θ_s) as integrated indices of soil properties, modified by an exponent n that is estimated empirically. A summary of our parameter estimation methodology as well as the equations for all eleven estimated parameters are listed in Appendix A.

3. Description of tests performed and study areas

3.1. Improvement of hydrologic simulation performance using SSURGO data

Six basins were chosen within the Ohio River Forecasting Center domain (OHRFC), for which both coarse-scale (STATSGO) and fine-scale (SSURGO) data are available. The basins are in Kentucky, Ohio, West Virginia, and Virginia. See Table 3 for descriptive information on these six basins. Fig. 1 contains a map of the basin locations. Parameter values were estimated, averaged to the scale of the basin, and SAC-SMA was run in lumped mode. The analysis period was limited by the extent of the data record. Comparison of hydrograph statistics was conducted to determine if SSURGO data led to improved performance relative to STATSGO, and, if so, if this performance was qualitatively significant.

Table 3
List of basins analyzed in Ohio River Forecasting Center

No.	Name	USGS ID No.	Analysis period	Gage elevation (m) (NGVD, 1929)	Basin area (km^2)
1	Tygert Valley R. near Dailey, WV	03050000	10/1965–09/1975	560.8	479.1
2	Shavers Fork below Bowden, WV	03068800	09/1973–09/1981	646.2	391.1
3	Deer Creek at Mt Sterling, OH	03230800	10/1966–09/1981	254.9	590.5
4	Tug Fork at Welch, WV	03212750	10/1985–09/1993	386.5	450.7
5	Dry Fork at Beartown, WV	03212980	10/1985–09/1993	321.9	541.3
6	Johns Creek near Meta, KY	03210000	10/1983–09/1993	218.8	145.8

3.2. *Using 30 m NLCD data to improve STATSGO-based estimates*

Comparison of curve number estimates was conducted for NLCD+SSURGO, NLCD+STATSGO, and generic land use+STATSGO data for the West Gulf River Forecasting Center (WGRFC), which consists largely of the state of Texas. This region was chosen to perform additional analyses comparing STATSGO to SSURGO performance since it represents a very different region from that of OHRFC. Rather than an analysis of a particular basin, curve number maps were generated for a county or soil survey area. Three Texas counties were examined: Borden County (2346 km^2) in northwest Texas, and Hamilton (2167 km^2) and Williamson (2936 km^2) counties in southeast Texas. Madison County in Ohio (1210 km^2) was also examined. This anticipates potentially generating large regions of improved curve numbers, which could then in turn lead to improved SAC-SMA parameters for basins of variable size.

The assumption in supplementing STATSGO data with the NLCD land cover data is that SSURGO data with its fine-scaled information on hydrologic soil group is not available. STATSGO does contain

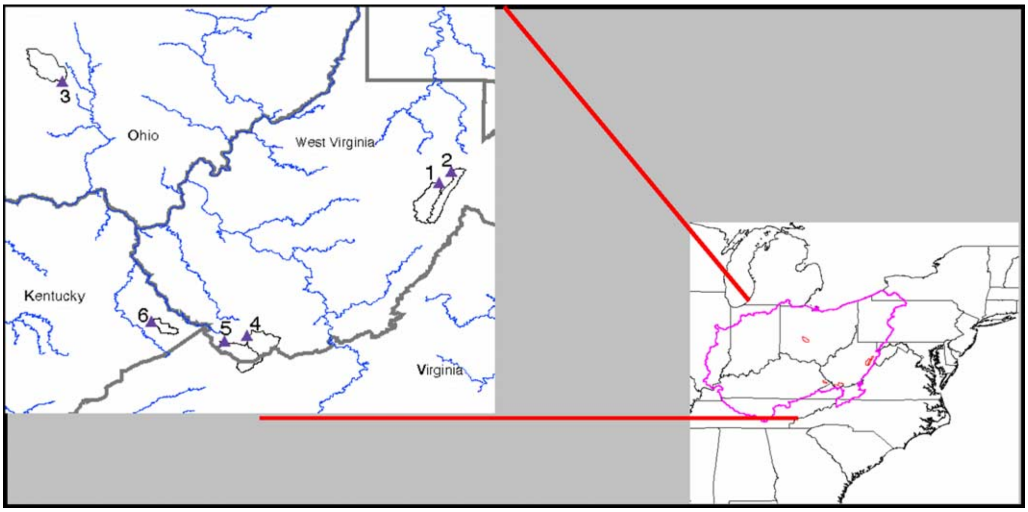

Fig. 1. Locations for analyzed basins within Ohio River Forecasting Centers. Basin outlets (triangles) and basin numbers are shown.

information on hydrologic soil group, but it is at a much coarser scale to begin with, and also only gives the percentage of each of four hydrologic groups (A, B, C, D) within a map unit. In contrast, SSURGO soil map units are assigned a single hydrologic soil group.

To generate the NLCD + STATSGO-based curve number estimates, four curve number grids were generated at 30 m resolution, one for each of the four hydrologic soil groups. These grids were then averaged to the scale of the STATSGO polygons and weighted by hydrologic soil group percentage to give the final curve number for each STATSGO polygon.

4. Results and discussion

4.1. Improvement of hydrologic simulation performance using SSURGO data

Improvement in estimation of SAC-SMA parameters was seen for two of the three OHRFC basins for which there was a significant difference in distribution of soil textures as measured by SSURGO vs. STATSGO data. This improvement is seen in terms of: (1) error statistics calculated based on comparing observed to simulated flows; and (2) recession behavior based on visual hydrograph comparisons of observed and simulated flows.

Depictions of the texture distributions based on the uppermost soil layer (thickness approximately 10–20 cm for SSURGO, 5 cm for the Miller and White (1998) gridded STATSGO) revealed that hydrologic simulation performance was noticeably different for (1) Deer Creek at Mt Sterling, where SSURGO clay textures were much more broadly distributed than measured by STATSGO, representing greater occurrence of a texture with a significant qualitative difference; (2) Shavers Fork below Bowden, where SSURGO loamy textures were more broadly distributed than measured by STATSGO, representing a significant quantitative difference of loamy soils; and (3) Johns Creek near Meta, where, unlike STATSGO, loamy and sandy loam SSURGO textures are interspersed in a regular pattern (Fig. 2). In contrast, for other basins where the overall distribution of surface textures for SSURGO vs. STATSGO was relatively similar little difference in hydrologic simulation performance was observed.

Statistical improvement in terms of hydrologic simulation performance was seen most clearly in terms of a statistic that calculates the root mean square error for hydrographs of the twelve largest storms within the simulation period—a flood root mean square error (FDRMS) (Table 4). For example, for three basins with little difference in texture distribution for STATSGO vs. SSURGO (Tygert Valley River, Tug Fork at Welch, Dry Fork at Beartown), FDRMS went from 46.7 to 48.0%, 31.5 to 33.6%, and from 35.7 to 35.3%, respectively, going from STATSGO to SSURGO. This is interpreted as being essentially no change. In contrast, for two of the three basins mentioned above with significant difference in texture distribution for STATSGO vs. SSURGO (Shavers Fork below Bowden, Deer Creek at Mt. Sterling), FDRMS went from 50.8 to 45.9% and from 66.3 to 52.9%, respectively, going from STATSGO to SSURGO. Thus FDRMS decreased when SSURGO data was used to estimate SAC-SMA parameters, indicating the greater accuracy of SSURGO-based parameter estimates. A similar pattern was seen for daily root mean square error (DRMS), monthly volume root mean square error (MVRMS), and correlation coefficient R.

Note that Johns Creek near Meta does not fit this pattern, since FDRMS increases from 49.8 to 59.0% going from STATSGO to SSURGO. This suggests that there may be additional factors besides surface texture (e.g. texture at lower depths) that account for SAC-SMA parameter differences. This is consistent with Table 5, which shows that differences in surface texture have translated into relatively large differences in parameters LZFSM and LZSK for Shavers Fork below Bowden and Deer Creek at Mt Sterling compared to Johns Creek near Meta. In addition to parameter differences, consequent differences in hydrologic simulation performance may also be driven by forcing errors. An issue with Johns Creek is that it is the smallest basin and thus the radar rainfall data used is the most error prone.

Improvement in terms of visual hydrograph comparisons was most clearly apparent when hydrographs were examined using a log-linear scale, since this depicts recession behavior better (Fig. 3). Here, it was again evident that there was improvement for Shavers Fork below Bowden and Deer Creek at Mt Sterling. For Deer Creek at Mt Sterling long term

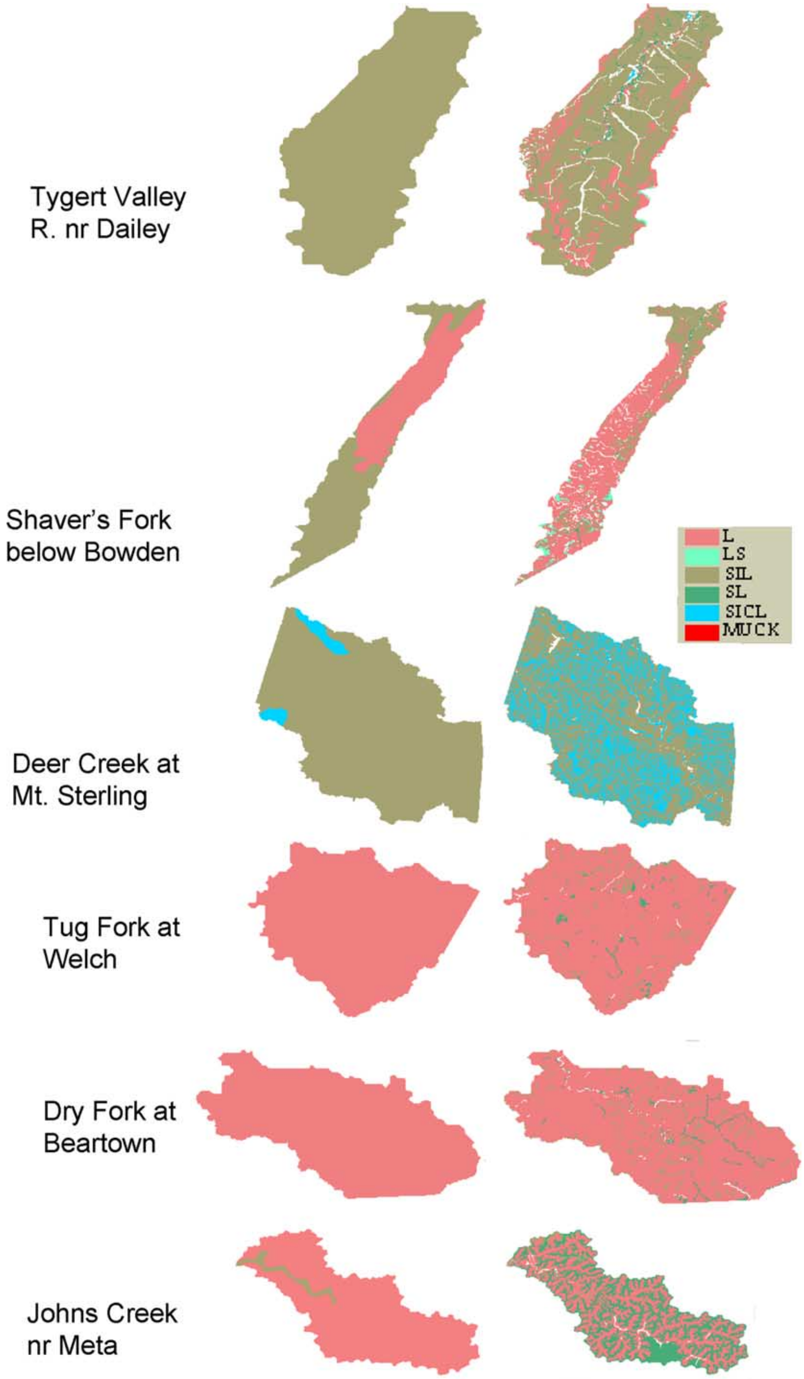

Fig. 2. Comparison of soil textures: STATSGO (left column) vs. SSURGO (right column).

Table 4
Comparison of accuracy statistics of hydrographs: SSURGO vs. STATSGO data

	Basin	DRMS (cm)	MVRMS (mm)	BIAS (%)	FDRMS (%)	R
SSURGO	Tygert Valley R.	9.1	14.8	−14.2	48.0	0.87
	Shavers Fork	9.0	21.5	−3.6	45.9	0.84
	Deer Creek	7.4	14.8	−18.8	52.9	0.84
	Tug Fork	4.5	11.2	−6.8	33.6	0.85
	Dry Fork	5.3	9.0	2.8	35.3	0.87
	Johns Creek	2.4	14.4	8.9	59.0	0.78
STATSGO	Tygert Valley R.	8.7	14.8	−14.1	46.7	0.88
	Shavers Fork	10.2	22.6	−3.4	50.8	0.81
	Deer Creek	8.7	18.7	−22.5	66.3	0.80
	Tug Fork	4.6	10.4	−10.3	31.5	0.83
	Dry Fork	5.2	8.1	−1.6	35.7	0.87
	Johns Creek	2.4	15.2	12.9	49.8	0.80

recession behavior was observed to be much better, while for Shavers Fork below Bowden short-term recession behavior was much better for SSURGO-based parameter estimates vs. STATSGO. This difference in performance could be attributed to the estimates for the lower layer supplemental free water storage parameter LZFSM, which for SSURGO vs. STATSGO are much more different for these two basins than for the others (Fig. 4). The depletion rate coefficient for the lower layer supplemental free water storage LZSK may also be making some contribution. In conclusion, this result again indicates that SSURGO-based parameters are more accurate.

Note that in making these SSURGO vs. STATSGO parameter and hydrologic simulation performance comparisons, SSURGO parameter values were available at the resolution of SSURGO soil polygons, which is much finer than STATSGO polygons (as discussed above.) To make this comparison, both SSURGO and STATSGO parameter values were averaged at the scale of the analyzed basins, and SAC-SMA runs were made using the lumped version of the model. Significant differences in lumped parameter values were often observed for SSURGO vs. STATSGO (Table 5). However, it is when the SSURGO-based approach is used in new, ongoing, NWS distributed modeling applications (Koren et al., 2004) that the greatest benefit of using the fine-scaled SSURGO data is expected to be obtained. As an example, Fig. 5 compares SAC-SMA parameter LZFSM derived at SSURGO and STATSGO resolutions for Shavers Fork below Bowden and shows significant differences in the spatially distributed values. Though the difference in basin-averaged mean is only approximately 5 mm, the root mean square error (RMSE) is more than two times larger, at 13.4 mm.

4.2. Using 30 m NLCD data to improve STATSGO-based estimates

Improvement in estimation of curve number was observed for all examined soil survey areas when 30 m NLCD data was used with STATSGO, as opposed to when generic 'pasture or range land use' was used. This improvement could be seen in that NLCD + STATSGO-based curve number estimates were more similar to NLCD + SSURGO-based curve number estimates than were those assuming generic land use. Table 6 reports curve number statistics for Borden, Hamilton, and Williamson counties, Texas (within WGRFC), and Madison County, Ohio (within OHRFC). NLCD + STATSGO-based curve number estimates should always be at least as accurate as generic land use-based estimates. However, they will have less accuracy relative to NLCD + SSURGO-based curve numbers, which are accurate to the scale of the 30 m land cover data or SSURGO polygons (which are mapped at a comparable scale to the land cover data).

In future applications where SSURGO soils data is not available, incremental improvement in SAC-SMA parameter estimates could be obtained by using the NLCD + STATSGO-based curve number estimates. This improvement would likely only be incremental

Table 5
Comparison of basin-averaged parameters: SSURGO vs. STATSGO data

Parameter	Tygert Valley R.	Shavers Fork	Deer Creek	Tug Fork	Dry Fork	Johns Creek
SSURGO						
Lztwm	125	144	202	119	117	166
Lzfsm	16	18	31	22	21	20
Lzfpm	74	95	101	73	72	110
Lzsk	0.16	0.17	0.15	0.14	0.14	0.16
Lzpk	0.005	0.010	0.005	0.007	0.007	0.009
Pfree	0.18	0.16	0.24	0.23	0.23	0.15
Uztwm	81	50	55	69	72	64
Uzfwm	57	37	29	54	54	52
Uzk	0.40	0.43	0.32	0.39	0.39	0.45
Zperc	78	70	64	59	62	71
Rexp	2.23	2.17	2.80	1.81	1.89	1.74
STATSGO						
Lztwm	134	118	234	160	167	157
Lzfsm	17	11	37	23	24	22
Lzfpm	67	61	148	120	118	116
Lzsk	0.09	0.10	0.09	0.12	0.11	0.12
Lzpk	0.004	0.006	0.007	0.005	0.005	0.005
Pfree	0.22	0.17	0.20	0.16	0.18	0.16
Uztwm	67	83	49	43	46	42
Uzfwm	40	58	28	39	39	37
Uzk	0.36	0.40	0.36	0.46	0.44	0.46
Zperc	118	127	94	90	92	92
Rexp	2.24	2.05	2.19	2.02	2.09	2.02
Percent difference by parameter						
Lztwm	−7	18	−16	−34	−43	6
Lzfsm	−7	37	−20	−6	−18	−12
Lzfpm	10	35	−46	−63	−65	−6
Lzsk	40	41	39	18	21	28
Lzpk	20	39	−30	32	29	47
Pfree	−19	−6	14	30	22	−8
Uztwm	17	−67	11	38	37	35
Uzfwm	30	−57	4	28	29	28
Uzk	10	7	−13	−18	−15	−1
Zperc	−52	−81	−46	−52	−47	−29
R_{exp}	0	6	22	−12	−10	−16
Average	91	106	114	109	113	85

since, as we have seen, the texture information in the STATSGO data may not be as accurate as in the SSURGO data. In the parameter estimation procedure, the curve number leads to the estimate of the depth of water to be allocated to the upper and lower zones, while the textures determine primarily how that water is split between tension and free water storages. Thus differences between SSURGO and STATSGO-based performance are driven by differences in curve numbers and initial rain abstraction and by accuracy in texture distributions.

The possibility of computing these improved, NLCD + STATSGO-based curve number estimates also suggests a useful future project might be to generate these estimates for the entire US, or for large or specially chosen regional studies within the US.

5. Summary

A number of conclusions may be drawn from this work. First, hydrologic simulation results suggest that

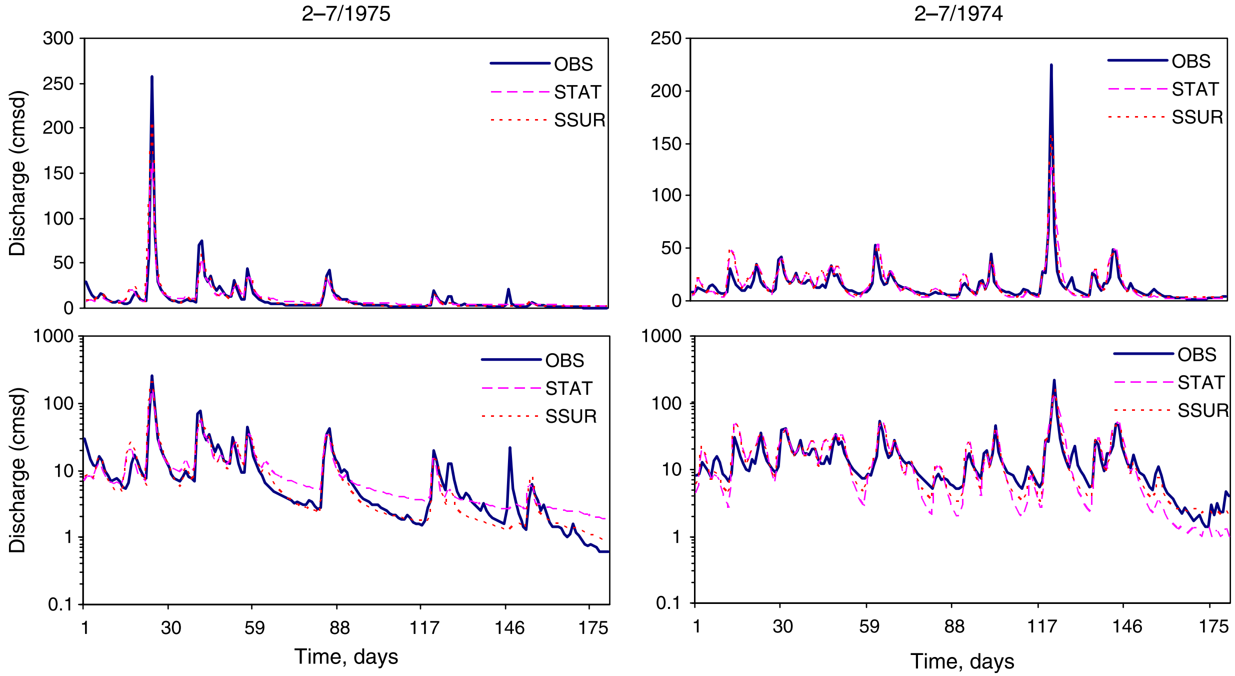

Fig. 3. Comparison of Hydrographs for STATSGO vs. SSURGO data: Deer Creek at Mt Sterling (left panel); Shavers Fork below Bowden (right panel). Note log-linear scale on lower plots in left and right panels.

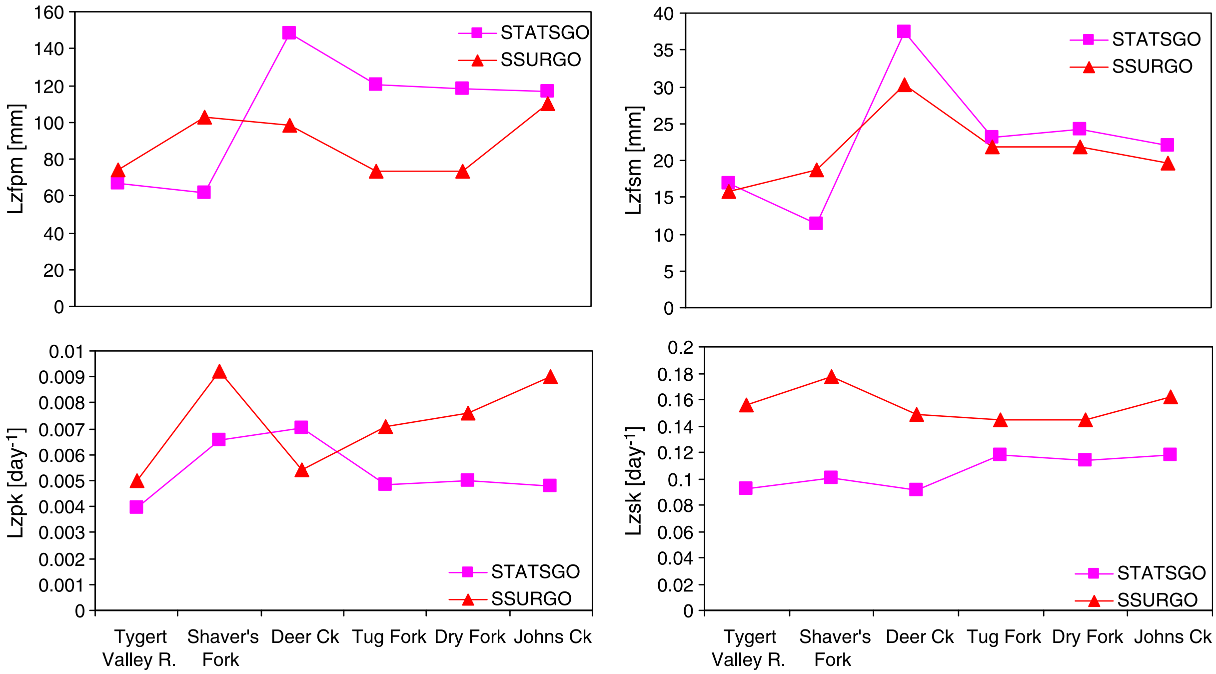

Fig. 4. Comparison of SAC-SMA parameters for STATSGO vs. SSURGO.

higher resolution, SSURGO-based parameter estimates can significantly improve flood prediction. Simulations generated using SSURGO-based parameter estimates compared favorably with simulations using STATSGO-based parameters in two basins, and, just as important, results in three of the other four basins tested were not significantly degraded by using the higher resolution data. Second, STATSGO data can lead to significant biases in model parameters, even for medium-sized basins such as the ones studied here. Smaller basins would be expected to have worse biases, since in those cases it becomes more likely that the correct textures are missed entirely. Third, combination of high resolution NLCD data and lower resolution STATSGO texture data can be used as a transitional option before SSURGO data become available throughout the entire US. Fourth, the fact that simulation improvements are seen when using more accurate, higher resolution data lends credence to the theory that is used to translate basic soil/land use data into hydrologic model parameters. Future work should include similar analyses on a larger sample of basins to validate and

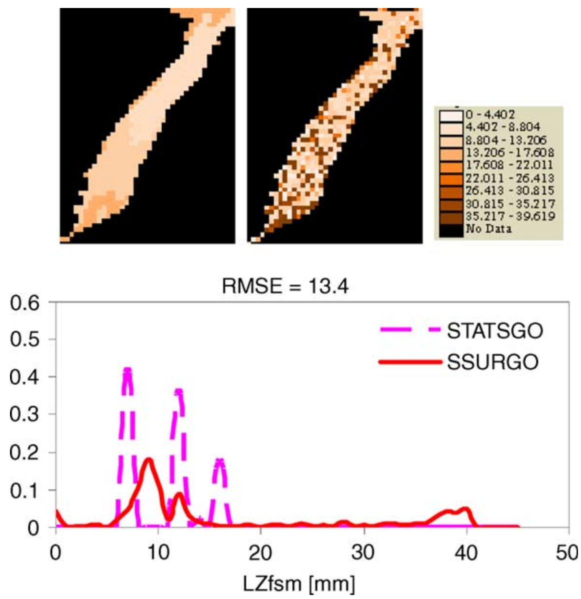

Fig. 5. Comparison of resolution of SAC-SMA parameter LZFSM for STATSGO (top left panel) vs. SSURGO (top right panel) for Shavers Fork below Bowden.

Table 6
Comparison of curve number estimates for NLCD+SSURGO, NLCD+STATSGO, and Generic land use+STATSGO estimates

	NLCD+ SSURGO	NLCD+ STATSGO	Generic land use+ STATSGO
Borden County, Texas			
Average (std dev.)	70.1 (4.32)	69.2 (5.24)	56.1 (4.97)
Difference of Avgs		0.9	14
Hamilton County, Texas			
Average (std dev.)	69.4 (3.21)	70.1 (4.27)	64.3 (4.53)
Difference of Avgs		-0.7	5.1
Williamson County, Texas			
Average (std dev.)	74.5 (3.02)	72.6 (8.10)	66.2 (7.34)
Difference of Avgs		1.9	8.3
Madison County, Ohio			
Average (std dev.)	68.9 (1.13)	71.6 (1.93)	64.4 (1.85)
Difference of Avgs		-2.7	4.5

strengthen our conclusions. Fifth, SSURGO data provide much finer spatial representation of a priori SAC-SMA parameters, which may be critical in high resolution, distributed modeling. SSURGO data spatially resolves information at flash flood basin scales, while STATGSO does not.

Acknowledgements

The authors thank two anonymous reviewers as well as M. Smith and F. Moreda of the Hydrology Laboratory, all of whose comments led to improvement of this paper.

Appendix A. Procedure to calculate SAC-SMA a priori parameters

SSURGO data containing information on soil texture, soil layer structure, and hydrologic soil group may be downloaded from the NRCS website of the USDA at http://www.ncgc.nrcs.usda.gov/branch/ssb/products/ssurgo/. The latest data format has a National Soil Information System (NASIS) attribute structure that can be imported into a Microsoft Access Template Database, available at http://nasis.nrcs.usda.gov/downloads/. Land cover data for the conterminous US at 30 m grid resolution are available from the US Geological Survey at http://landcover.usgs.gov/natllandcover.asp. Arcview GIS software (ESRI, 1996) is used to process the spatial and attribute data and calculate SAC-SMA parameter estimates, resulting in maps of soil polygons where a value for each model parameter for each polygon has been added to the associated attribute table. These polygons may then be readily transformed to grids (raster data) at any resolution for any SAC-SMA parameter of interest.

Below are the SAC-SMA parameter and soil property relationships as they appeared in Koren et al. (2003), except for (A8) which contains a slight correction: π has become π^2 and D_s has become D_s^2.

Upper layer parameters:

$$\mathrm{UZTWM} = (\theta_{\mathrm{fld}} - \theta_{\mathrm{wp}})Z_{\mathrm{up}} \tag{A1}$$

$$\mathrm{UZFWM} = (\theta_{\mathrm{s}} - \theta_{\mathrm{fld}})Z_{\mathrm{up}} \tag{A2}$$

$$\mathrm{UZK} = 1 - (\theta_{\mathrm{fld}}/\theta_{\mathrm{s}})^n \tag{A3}$$

Lower layer parameters:

$$\mathrm{LZTWM} = (\theta_{\mathrm{fld}} - \theta_{\mathrm{wp}})(Z_{\max} - Z_{\mathrm{up}}) \tag{A4}$$

$$\mathrm{LZFSM} = (\theta_{\mathrm{s}} - \theta_{\mathrm{fld}})(Z_{\max} - Z_{\mathrm{up}})(\theta_{\mathrm{wp}}/\theta_{\mathrm{s}})^n \tag{A5}$$

$$\mathrm{LZFPM} = (\theta_{\mathrm{s}} - \theta_{\mathrm{fld}})(Z_{\max} - Z_{\mathrm{up}}) \times [1 - (\theta_{\mathrm{wp}}/\theta_{\mathrm{s}})^n] \tag{A6}$$

$$\mathrm{LZSK} = \frac{1 - (\theta_{\mathrm{fld}}/\theta_{\mathrm{s}})^n}{1 + 2(1 - \theta_{\mathrm{wp}})} \tag{A7}$$

$$\mathrm{LZPK} = 1 - \exp\left\{ -\frac{\pi^2 K_{\mathrm{s}} D_{\mathrm{s}}^2 (Z_{\max} - Z_{\mathrm{up}})\Delta t}{\mu} \right\} \tag{A8}$$

$$\mathrm{PFREE} = (\theta_{\mathrm{wp}}/\theta_{\mathrm{s}})^n \tag{A9}$$

$$\text{ZPERC} = \frac{\text{LZTWM} + \text{LZFSM}(1 - \text{LZSK}) + \text{LZFPM}(1 - \text{LZPK})}{\text{LZFSM} \times \text{LZSK} + \text{LZFPM} \times \text{LZPK}} \tag{A10}$$

Percolation parameters:

$$\text{REXP} = [\theta_{\text{wp}}/(\theta_{\text{wp,sand}} - 0.001)]^{0.5} \tag{A11}$$

Upper layer thickness:

$$Z_{\text{up}} = 5.08 \times \frac{1000/\text{CN} - 10}{\theta_{\text{s}} - \theta_{\text{fld}}} \tag{A12}$$

References

Armstrong, B.L., 1978. Derivation of initial soil moisture accounting parameters from soil properties for the national weather service river forecast system, NOAA Technical Memorandum, NWS HYDRO 37, p. 53.

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, Colorado, pp. 311–366.

Clapp, R.B., Hornberger, G.M., 1978. Empirical equations for some soil hydraulic properties. Water Resources Research 14 (4), 601–604.

Cosby, B.J., Hornberger, G.M., Clapp, R.B., Ginn, T.R., 1984. A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soil. Water Resources Research 20 (6), 682–690.

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. Journal of Geophysical Research 108-D22, 12.

Dingman, S.L., 2002. Physical Hydrology, 2nd ed. Prentice Hall, Englewood Cliffs, New Jersey p. 600.

ESRI (Environmental Systems Research Institute), 1996.Anon., 1996. Using Arcview GIS. ESRI Press, Redlands, California p. 350.

Koren, V.I., Smith, M., Wang, D., Zhang, Z., 2000. Use of Soil Property Data in the Derivation of Conceptual Rainfall-Runoff Model Parameters Proceedings of the 15th Conference on Hydrology. AMS, Long Beach, CA, pp. 103–106.

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., Sorooshian, S., Gupta, H., Rosseau, H., Turcotte, H. (Eds.), Calibration of Watershed Models Water Science and Applications, vol. 6. AGU, pp. 239–254.

Koren, V.I., Reed, S., Smith, M., Zhang, Z., Seo, D.-J., 2004. Hydrology laboratory research modeling system (HL-RMS) of the US National Weather Service. Journal of Hydrology 291 (1–5), 297–318.

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-D07S9, 22.

McCuen, R., 1998. Hydrologic Analysis and Design, 2nd ed. Prentice Hall, New Jersey.

Miller, D.A., White, R.A., 1998. A conterminous United States multi-layer soil characteristics data set for regional climate and hydrology modeling, Earth Interactions 2, (available at http://EarthInteractions.org).

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. Journal of Geophysical Research 109-D07S90, 32.

Peschel, J.M., Haan, K.P., Lacey, R.E., 2003. A SSURGO pre-processing extension for the ArcView Soil and Water Assessment Tool, ASAE Meeting Presentation, Paper No. 032123, Las Vegas, NV, July 27–30, 2003.

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.

Reed, S.M., 1998. Use of Digital Soil Maps in a Rainfall-runoff Model, Doctoral Dissertation, Advisor: Dr. D.R. Maidment, The University of Texas at Austin, Austin, Texas, Dec. 1998.

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

Seo, D.-J., Koren V., Cajina N., Corby R., Finn B., Bell F., Howieson T., 2003. 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.B., Laurine, D., Koren, V., Reed, S., Zhang, Z., 2003. In: Duan, Q., Sorooshian, S., Gupta, H., Rosseau, H., Turcotte, H. (Eds.), Hydrologic model calibration in the National Weather Service Calibration of Watershed Models, Water Science and Applications, vol. 6, pp. 133–152 (AGU).

Smith, M.B., 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. Journal of Hydrology 298, 4–26.