



## From lumped to distributed via semi-distributed: Calibration strategies for semi-distributed hydrologic models

Behnaz Khakbaz \*, Bisher Imam, Kuolin Hsu, Soroosh Sorooshian

Center for Hydrometeorology and Remote Sensing, Department of Civil and Environmental Engineering, University of California, Irvine, E 4130 Engineering Gateway, Irvine, CA 92697, USA

### ARTICLE INFO

#### Article history:

Available online 20 February 2009

#### Keywords:

Distributed hydrologic modeling  
Calibration  
A priori parameters  
Multi-sensor precipitation  
Streamflow simulation

### SUMMARY

Modeling the effect of spatial variability of precipitation and basin characteristics on streamflow requires the use of distributed or semi-distributed hydrologic models. This paper addresses a DMIP 2 study that focuses on the advantages of using a semi-distributed modeling structure. We first present a revised semi-distributed structure of the NWS SACramento Soil Moisture Accounting (SAC-SMA) model that separates the routing of fast and slow response runoff components, and thus explicitly accounts for the differences between the two components. We then test four different calibration strategies that take advantage of the strengths of existing optimization algorithms (SCE-UA) and schemes (MACS). These strategies include: (1) lumped parameters and basin averaged precipitation, (2) semi-lumped parameters and distributed precipitation forcing, (3) semi-distributed parameters and distributed precipitation forcing and (4) lumped parameters and basin averaged precipitation, modified using a priori parameters of the SAC-SMA model. Finally, we explore the value of using discharge observations at interior points in model calibration by assessing gains/losses in hydrograph simulations at the basin outlet. Our investigation focuses on two key DMIP 2 science questions. Specifically, we investigate (a) the ability of the semi-distributed model structure to improve stream flow simulations at the basin outlet and (b) to provide reasonably good simulations at interior points.

The semi-distributed model is calibrated for the Illinois River Basin at Siloam Springs, Arkansas using streamflow observations at the basin outlet only. The results indicate that lumped to distributed calibration strategies (1 and 4) both improve simulation at the outlet and provide meaningful streamflow predictions at interior points. In addition, the results of the complementary study, which uses interior points during the model calibration, suggest that model performance at the outlet can be further improved by using a semi-distributed structure calibrated at both interior points and the outlet, even when only a few years of historical record are available.

© 2009 Elsevier B.V. All rights reserved.

### Introduction

Spatial variabilities of precipitation and basin properties have significant impacts on the hydrologic response of basins. Characterizing and modeling the relationship between the spatial distribution of rainfall, basin characteristics, and runoff generation has been the subject of many studies for more than two decades now. Using a semi-distributed hydrologic model, Wilson et al. (1979) showed that accounting for the spatial variation of precipitation significantly influences the volume, time to peak, and peak flow of predicted hydrograph. Studies conducted by Troutman (1983), Beven (1985), Krajewski et al. (1991), Ogden and Julien (1994), and Shah et al. (1996) reached similar conclusion.

Modeling the effect of spatial variability of precipitation and basin characteristics on streamflow requires the use of distributed or

semi-distributed hydrologic models. Similar to lumped models, distributed models are conceptual or physically based (Kampf and Burges, 2007). Physically based models solve the equations expressing the conservation of mass, momentum and energy (Kampf and Burges, 2007) and therefore, require a significant amount of information. Conceptual models, on the other hand, approximate the general physical mechanisms governing the hydrologic processes (Duan et al., 1992), and may be less demanding in terms of model input.

While in theory distributed models are expected to outperform their lumped counterparts, reality however, has produced mixed results. Beven (1989) and Grayson et al. (1992) concluded that physically based distributed models, when compared to lumped models, often provide only slightly better, if not equal or even worse simulated flows. A similar observation was made by Reed et al. (2004), who reported the results of a comprehensive inter-comparison study of several physically based and conceptual distributed models and concluded that "... lumped model outper-

\* Corresponding author. Tel.: +1 949 8249350.

E-mail address: [bkhakbaz@uci.edu](mailto:bkhakbaz@uci.edu) (B. Khakbaz).

formed distributed models in more cases than distributed models outperformed the lumped model...” Given these conclusions and the magnitude of efforts required to parameterize and validate distributed hydrologic models, it is reasonable to question whether distributed models can effectively be used in operational hydrologic forecasting?

Beside the potential to improve streamflow prediction at the basin outlet, another benefit of distributed models is their ability to produce streamflow predictions at interior locations where streamflow measurements may not be available (Koren et al., 2004). Michaud and Sorooshian (1994) showed that a complex distributed model calibrated at the basin outlet had the ability to generate interior point streamflow simulations that were comparable in accuracy to outlet predictions. They also demonstrated that a simple distributed model was as accurate as a complex distributed model, and noted that model complexity does not necessarily improve the accuracy of the simulations. Reed et al. (2004) reported that when calibrated at the outlet of larger parent basins, distributed models, which participated in the first phase of the Distributed Model Inter-comparison Project (DMIP), produced reasonable performance at interior locations where no explicit calibration was performed. However, when these models were calibrated at the outlet of relatively smaller parent basins, a degradation of the performance at interior points was observed. Arguing that the lack of explicit calibration at interior points does not fully explain poor model performance at such points, they concluded that further studies are required to discern the causes of the above-described performance degradation.

An important aspect of distributed models is their highly parameterized nature. The multi-dimensional optimization problem, commonly associated with lumped models, becomes hyper-dimensional in the case of distributed models. With the successful application of any hydrologic model being dependent on the quality of its calibration (Duan et al., 1992), developing calibration strategies for distributed models is naturally a requirement for their proper application in hydrologic forecasting. Recent hydrologic literature provides a number of examples of strategies to tackle the parameterization of distributed models, mostly by reducing the dimensionality of the calibration problem and therefore making it solvable by existing optimization algorithms. One approach relies on adjusting the parameters of each individual cell of a distributed model from their a priori values using calibrated adjustment factors, which are applied uniformly to all cells (Leavesley et al., 2003; Koren et al., 2003, 2004; Eckhardt et al., 2005; Giertz et al., 2006). More specifically, for example, Koren et al. (2004), used a priori SAC-SMA parameter grids developed by Koren et al. (2000) as initial parameter values for the National Weather Service (NWS) Hydrologic Lab Research Distributed Hydrologic Model (HL-RDHM), to account for spatial variability of soil and land use within their study basin. A lumped model calibration was then performed and the ratios between lumped calibrated parameters and spatially averaged a priori values were subsequently used to adjust individual cell parameters in the distributed model. Frances et al. (2007) divided the effective parameters at each grid of their distributed model (TETIS) into two components: (1) parameters representing the hydrological characteristic at point scale, (2) a correction factor for each parameter applied identically to all the grid cells to account for the combined effects of modeling errors such as temporal and spatial scale impacts. The dimension of the calibration problem is thus reduced from  $(n_{\text{cells}} \times n_{\text{par}})$  to  $n_{\text{par}}$  of correction factors, which were calibrated using the Shuffled Complex Evolution (SCE-UA, Duan et al., 1992) resulting in what they described as very satisfactory results.

A different approach to reducing the calibration problem dimensionality was proposed by Ajami et al. (2004). Their ap-

proach includes three different calibration strategies: lumped, semi-lumped, and semi-distributed, which were used to optimize the parameters of a semi-distributed version of the SAC-SMA model. The semi-lumped calibration strategy, which assigns identical parameter values and spatially varied Mean Areal Precipitation (MAP) at sub-basins, outperformed lumped and semi-distributed calibration strategies. Ajami et al. (2004) argued that their results are consistent with the uniformity of the physical characteristics of their study basin (Illinois River basin at Watts, Oklahoma).

Substantial reduction in the number of calibration parameters can be accomplished by describing the spatial variability within each watershed/grid element in terms of probability distributions. For example, Cole and Moore (2008) developed a topography based probability distribution runoff production scheme, which utilizes a suite of empirical equations (Bell and Moore, 1998) to describe the relationship between saturation, moisture capacity, and topographic gradient within each grid cell of the distributed model.

From an operational point of view, distributed models provide an opportunity to expand operational forecasts beyond traditional streamflow forecasting. As indicated by Smith et al. (this issue), the NWS is interested in infusing advanced hydrologic modeling tools, including distributed models, into its operational forecasting system. This need is motivated by the increasing demands for complementary water resources relevant forecasts. The NWS has adopted model inter-comparison experiments (DMIP) as the venue for model developers, including those in academia, to test new advances and improvements using operational quality data (See Smith et al., this issue). DMIP phase 1 (2001–2004) was successful in providing the hydrologic research and operational communities with useful results. However, gaps resulting from (a) short verification data period, (b) concerns about the quality of radar precipitation estimates used in the experiment, and (c) the limited geographic domain of the experiment (Smith et al., 2004; Reed et al., 2004) motivated the NWS to initiate a second DMIP phase (DMIP 2). As mentioned in Smith et al. (this issue), DMIP 2 was designed to coordinate among other research issues, community efforts to determine whether distributed models can reliably produce basin response at interior points, and research to develop and refine calibration strategies that are suitable for distributed models. In this paper, we report our DMIP 2 study, in which we investigate the potential to improve interior point simulations using semi-distributed modeling framework and also test several distributed model calibration strategies. We first expand the contribution of Ajami et al. (2004) by extending the realism of their semi-distributed version of SAC-SMA, through our separate routing of fast and slow response runoff components, which explicitly accounts for the differences between the two components. We then test different calibration strategies that are suitable for the calibration of distributed hydrologic models by investigating calibration scenarios that take advantage of the strengths of existing optimization algorithms, while at the same time allow for the calibration of distributed models. Finally, we explore the value of using discharge observations at interior points in model calibration by assessing improvements in hydrograph simulations at the basin outlet.

## Model description

The SAC-SMA model (Burnash et al., 1973; Burnash, 1995), which is one of the major components of the NWSRFS, is used as the core component of our semi-distributed modeling structure to generate hourly streamflow simulations. The model uses sub-basins as the computational elements of rainfall-runoff modeling with each sub-basin consisting of a lumped SAC-SMA. Mean Areal Precipitation (MAP) and potential evapotranspiration provide forcing data for each sub-basin. The model generates runoff response

components for each sub-basin. The distributed configuration used in this study separates fast response components, which are routed over the hillslopes using the unit hydrograph of the sub-basin, from slow response components, which bypass the overland flow routing and are introduced directly to the sub-basin outlet. The kinematic wave routing method provides the mechanism for sub-basin-to-sub-basin channel routing. In this approach, the main stream in each sub-basin is divided into several reaches depending on the slope homogeneity and the length of the reach. The generated discharge from the contributing area in each sub-basin is added to the routed streamflow at the end of each channel reach. Therefore, the lateral flow in the kinematic wave routing is assumed to be negligible. To some extent, the proposed structure is built on the contribution of Ajami et al. (2004), and is similar in many aspects to the approach implemented in the USACE HEC-HMS package (USACE, 2000). However, this study extends the realism of the semi-distributed structure, particularly with respect to the SAC-SMA, by separating the fast and slow response components of the SAC-SMA at each sub-basin, and by routing them in manners more consistent with distributed modeling framework. It is also worth noting that this approach is compatible with the NWS distributed model (Koren et al., 2004; Smith et al., 2007-DMIP workshop), which is based on a gridded SAC-SMA model with kinematic hillslope and channel routing. The distributed configuration enables the model to simulate streamflow at the basin outlet as well as pre-specified interior locations along the channel.

#### *Sacramento Soil Moisture Accounting Model (SAC-SMA) as the water balance component*

The SAC-SMA is essentially a conceptual lumped-input and lumped-parameter model (Peck, 1976). It utilizes precipitation

and evapotranspiration averaged over the whole basin, as the inputs to the model. The model considers two zone layers: an upper zone representing the uppermost layer, and a lower zone representing the deeper portion of the soil layer.

Each zone consists of tension and free water storages. The model generates five runoff response components: (1) direct runoff produced from falling precipitation on permanent and temporary impervious areas, (2) surface runoff generated when precipitation occurs at a rate faster than percolation, (3) interflow, which is the lateral outflow from the upper-zone free water storage, (4) supplementary base flow, which is the lateral drainage from lower-zone supplementary free water storage, and (5) primary base flow, which is the lateral drainage from the lower-zone primary free water storage.

In our distributed model structure, direct runoff, surface runoff and interflow are considered as fast response components, and primary and supplementary base flows as slow response components of runoff. Fig. 1 shows a schematic of the Sacramento model (SAC-SMA) structure along with the routing approach used herein. Readers interested in more details about the SAC-SMA are referred to Burnash et al. (1973) and Burnash (1995).

#### *Overland flow routing*

In lumped applications of the SAC-SMA, a Unit Hydrograph (UH) is used to transform excess rainfall to discharge at the basin outlet. In our distributed configuration of SAC-SMA, the fast runoff response components are combined and routed over land using the UH of each sub-basin outlet. The UH, which is defined as the discharge produced by a unit volume of effective rainfall of a given duration applied uniformly over the basin (Bras, 1990), assumes that the basin responds linearly to effective rainfall. In general,

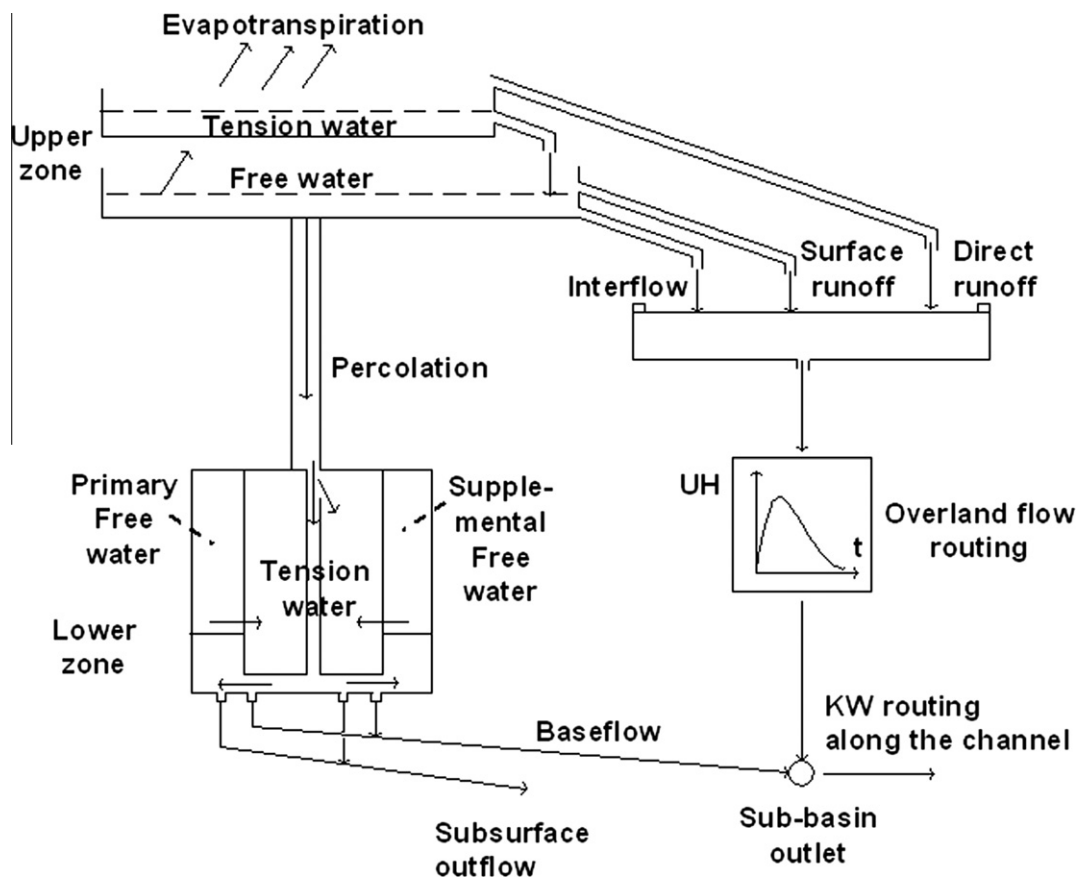


Fig. 1. NWS SAC-SMA model in the semi-distributed model structure.

for a given basin, a UH is derived either directly using streamflow and precipitation data for selected storms, or synthetically using methods such as the Snyder UH (Snyder, 1938), the SCS dimensionless UH (Soil Conservation Service, 1972), or the time-area histogram method (Clark, 1943). In this study, the synthetic SCS dimensionless UH method as described by Chow et al. (1988) is used to derive the UHs of sub-basins. The dimensionless SCS hydrograph is shown in Fig. 2. In this method, the time separating the start of the excess rainfall and peak discharge,  $t_p$  (h), is obtained by

$$t_p = \frac{D}{2} + t_l \quad (1)$$

where  $D$  is the duration of excess rainfall (herein 1 h) and  $t_l$  is the lag time from excess rainfall centroid to peak discharge (h). The lag time  $t_l$  can be estimated by

$$t_l = C_t(L.L_c)^{0.3} \quad (2)$$

where  $C_t$  is a factor representing the average main channel slope, ranging from 1.4 (mountainous regions) to 1.6 (flat areas) (Alizadeh, 2003);  $L$  is the longest flow path distance (m) and  $L_c$  is the distance between the basin outlet and the nearest point to the basin centroid along the river channel (m). The peak discharge  $Q_p$  ( $\text{m}^3/\text{s}$ ) is then computed as

$$Q_p = \frac{0.208A}{t_p} \quad (3)$$

where  $A$  is basin area in  $\text{km}^2$ . Having computed  $t_p$  and  $Q_p$ , the UH of each sub-basin is derived from the SCS dimensionless unit hydrograph.

#### Channel routing

The kinematic wave approximation is used to route the flow through the channels connecting the sub-basins. The approxima-

tion simplifies the 1-D formulation of the Saint-Venant equations by assuming that inertial and pressure forces are negligible and that gravity and friction forces are balanced (Chow et al., 1988). This reduces the momentum equation to:

$$S_o = S_f \quad (4)$$

Using the Manning equation to represent flow resistance, the momentum equation can be written as

$$A = \alpha Q^\beta \quad (5)$$

with  $\alpha = \left( \frac{nP^2}{\sqrt{S_o}} \right)^{0.6}$  and  $\beta = 0.6$ , where  $n$  is the Manning roughness coefficient,  $P$  is the wetted perimeter of the channel (m) and  $S_o$  is the slope of the channel.

Combining the continuity and momentum equations, one can obtain an equation with  $Q$  as the only dependent variable:

$$\frac{\partial Q}{\partial x} + \alpha \beta Q^{\beta-1} \frac{\partial Q}{\partial t} = q \quad (6)$$

In this study, we use the nonlinear scheme, which is unconditionally stable (Li et al., 1975), as described by Chow et al. (1988) to solve the finite difference equations of the kinematic wave approximation.

#### Model application to the Illinois River basin

The semi-distributed configuration of the SAC-SMA is applied here to the Illinois River basin upstream of USGS gauging station (07195430) located south of Siloam Springs, Arkansas. This DMIP 2 test watershed, hereafter referred to as the Siloam basin, occupies 1489  $\text{km}^2$ , which is typical of the size used as an operational forecasting unit by NWS (Smith et al., 2004). Elevation changes are mild and range from 285 m at the outlet to 590 m at the highest point in the watershed. The basin's mild topography is evident by the low average slope of 0.35% along its 76 km longest flow path. In general, the basin's land cover can be described as uniform with approximately 90% of the basin area being covered by deciduous broadleaf forest with the remainder being mostly woody savannahs and croplands. According to Smith et al. (2004), the dominant soil types in the Siloam basin are silty clay (SIC), silty clay loam (SICL), and silty loam (SIL). The average annual rainfall and runoff are 1160 and 302 mm/year, respectively. The average annual free water evaporation within the larger basin (above Tahl-equah) is 1066 mm/year, with the maximum monthly averages (147 and 155 mm) occurring during the months of June and July, respectively. DMIP 2 modeling instructions for this basin, call for participants to generate the basin outlet hydrograph as well as hydrographs at three interior points (Table 1).

#### Model configuration

Dividing the watershed into sub-basins linked with channel reaches is the first step of constructing a semi-distributed model configuration. The ArcView HEC-GEOHMS extension (USACE, 2003) was used to perform basin delineation using the USGS 30 m Digital Elevation Model data (DEM). The initial stream confluence-derived set of sub-basins was modified to allow for

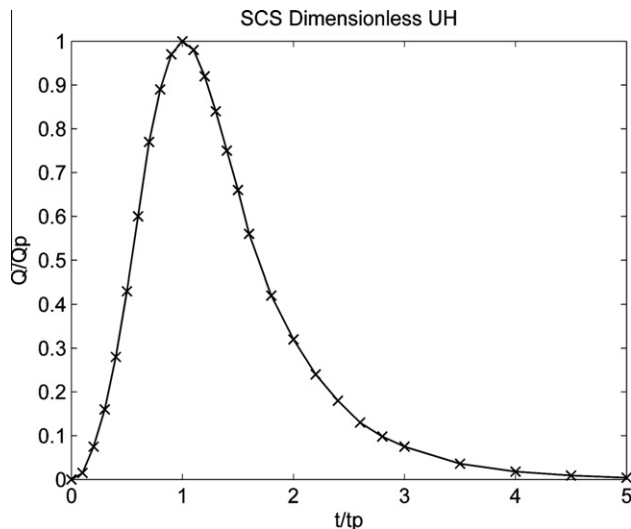
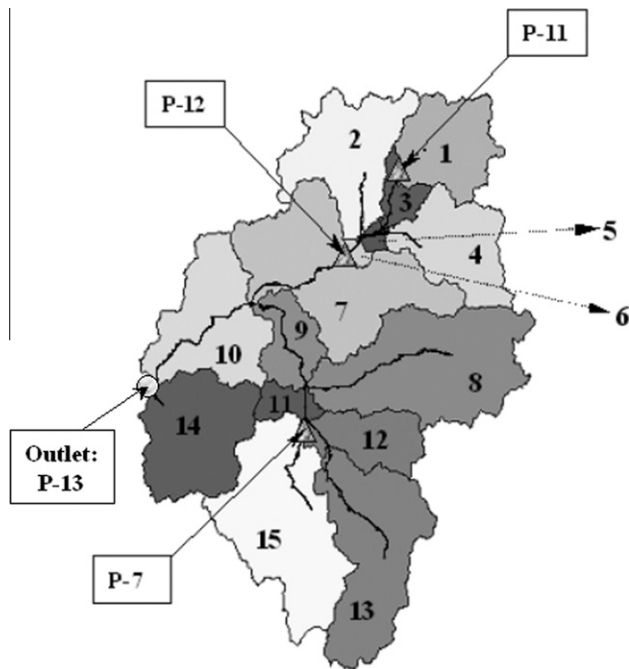


Fig. 2. SCS dimensionless UH.

Table 1  
Stream gauges in the Siloam basin.

DMIP 2 gauge#	Type	USGS ID	Gauge Name	Abbreviation	Contributing area ( $\text{km}^2$ )
7	Interior	07194800	Illinois River at Savoy, Arkansas	Savoy	433
11	Interior	07194880	Osage Creek at Cave Springs, Arkansas	Caves	90
12	Interior	07195000	Osage Creek at Elm Springs, Arkansas	Elmsp	337
13	Outlet	07195430	Illinois River South of Siloam Springs, Arkansas	Siloam	1489





**Fig. 3.** Sub-basin delineation of the Illinois River basin at south of Siloam Spring, AR (three interior points are shown by triangles).

streamflow simulation at each of the three interior points leading to a total of fifteen sub-basins with areas ranging from 5 to 200 km<sup>2</sup> (Fig. 3). Then, the SCS dimensionless UH parameters were determined by applying Eqs. (1)–(3) using GIS extracted watershed area ( $A$ ), longest flow path ( $L$ ), and centroidal distance along the longest flow path ( $L_c$ ).

While channel reach length and slope were also determined using the GIS, the 30 m resolution of the DEM prevented accurate extraction of channel cross-section information. Therefore, we assume wide rectangular channel geometry with constant width as well as constant Manning roughness coefficient for the entire river network. This assumption is consistent with the findings of Ajami et al. (2004), who noted the relatively homogeneous physical properties of the basin.

#### Forcing data

For each sub-basin in the distributed structure, the lumped SAC-SMA is forced by hourly Mean Areal Precipitation (MAP) and Potential Evapotranspiration (PET). Eleven years of hourly MAP time-series were obtained from NOAA's multi-sensor (NEXRAD and gauge) data set. These data were made available to DMIP 2 participants in Hydrologic Rainfall Analysis Project (HRAP) grid format at 4 km × 4 km spatial resolution. A sub-basin grid mask of identical spatial resolution to that of the precipitation data was used to extract precipitation grids over each sub-basin, which was averaged to compute the sub-basin's MAP.

Climatological monthly mean Free Water Surface (FWS) Evaporation estimates (mm/day) representing the parent 2484 km<sup>2</sup> basin (above USGS Gauge 07196500) were obtained from DMIP 2 website (Table 2). Because of the lack of information regarding the spatial and diurnal variability of PET, we assume uniform spatial and diurnal patterns of PET over all sub-basins and during the day. As

such, dividing FWS estimates by 24 yields the average hourly PET for each month.

#### Model parameterization and calibration

##### Model parameters

Given the above-described model structure, parameters of the semi-distributed SAC-SMA are divided into two classes: (1) non-calibrated parameters and (2) calibrated parameters. The non-calibrated class represents parameters that were not calibrated in this study because they represent geometric properties of the sub-basins and river channel network. As mentioned earlier, sub-basin delineation using GIS processing provides the drainage area, river channel slope, and river reach lengths as geometric parameters, which are also used in computing the parameters of the UH. Similarly, channel network parameters such as Manning roughness coefficient and channel cross sections, obtained from previous studies conducted on this basin, fall within the non-calibrated class. Calibrated parameters include parameters of the SAC-SMA water balance component. In lumped implementation, SAC-SMA has 13 major parameters that cannot be measured directly and need to be defined through calibration (Table 3).

##### Optimization algorithm and scheme

A difficult and important task, which always accompanies the use of hydrologic models, is calibration or parameter optimization. It is now well known that Conceptual Rainfall–Runoff (CRR) models pose significant calibration challenges (e.g., Sorooshian and Gupta, 1983; Hendrickson et al., 1988; Sorooshian et al., 1993; Gan and Biftu, 1996). One of the most difficult of these challenges arises as a direct result of the multimodal nature of the objective function's response surface. Consequently, local search methods such as simplex method (Nelder and Mead, 1965) have a very low success probability in finding the global optimum parameter set in CRR models (Duan et al., 1992).

Development of global search algorithms has been an active area of hydrologic research for nearly two decades now. Two common global search methods are the population-evolution-based Shuffled Complex Evolution-Univ. of Arizona (SCE-UA) (Duan et al., 1992) and Genetic Algorithms (GA) (Wang, 1991). A number of studies have compared SCE-UA, GA and other global and local search algorithms to calibrate the CRR models (Duan et al., 1992; Cooper et al., 1997; Kuczera, 1997; Thyer et al., 1999). These studies showed that the SCE-UA is an effective and efficient search algorithm that can be applied to calibrate complex conceptual hydrologic models. SCE-UA combines the nonlinear simplex method of Nelder and Mead (1965), a random search procedure, and complex shuffling (Duan et al., 1992) to direct the evolution of the parameter space towards the global optima. The SCE-UA global optimization algorithm is used to calibrate the parameters of the semi-distributed SAC-SMA model in this study.

The traditional approach to calibrate CRR models has relied on using a single objective function such as the Root Mean Square Error (RMSE) or Percent Bias, among others (Hogue et al., 2000). While this allows the modeler to employ automatic calibration, which is fast in comparison with manual calibration, experience shows that it can result in simulated hydrographs that are hydrologically acceptable but may be biased towards certain aspects of the watershed response (Gupta et al., 1998; Ajami et al., 2004). Hogue et al. (2000) proposed a Multi-step Automatic Calibration

**Table 2**

Free water surface evaporation data (PE) of Illinois River at Tahlequah, Oklahoma, which is the parent basin of Siloam basin (values are for total area including sub-basins).

Month	J	F	M	A	M	J	J	A	S	O	N	D
PE (mm/day)	0.84	1.36	2.23	3.34	4.32	4.92	5.00	4.49	3.58	2.48	1.49	0.90

**Table 3**  
SAC-SMA model major parameters and their feasible ranges (Koren et al., 2004).

Parameter	Definition	Feasible range
UZWWM	Upper-zone tension water capacity	10–300
UZFWM	Upper-zone free water capacity	5–150
UZK	Lateral drainage rate of upper-zone free water	0.1–0.75
ZPERC	Factor to define the proportional increase in percolation from saturated to dry lower-zone condition	5–350
REXP	Parameter related to the shape of percolation curve	1–5
LZWWM	Lower-zone tension water capacity	10–500
LZFWM	Lower-zone supplemental free water capacity	5–400
LZFSM	Lower-zone primary free water capacity	10–1000
LZSK	Lateral drainage rate of lower-zone supplemental free water	0.01–0.35
LZPK	Lateral drainage rate of lower-zone primary free water	0.001–0.05
PFREE	Percentage of percolation goes directly to lower-zone free water storages	0.0–0.8
PCTIM	Fraction of permanent impervious area	
ADIMP	Maximum fraction of additional impervious area due to saturation	

Scheme (MACS) to resolve this problem. The MACS procedure, which combines the strength of manual and automatic calibration, consists of three steps as following (Hogue et al., 2000):

**Step 1.** Calibrate all parameters of the SAC-SMA model using the LOG objective function (Eq. (7)). By reducing the influence of high flows, the LOG transformation prevents the calibrated parameters from being biased towards the simulation of high flows at the expense of low flows. As such, the LOG transformation would be suitable to estimate the lower-zone parameters associated with base flow runoff components. Hogue et al. (2000) noted that by computing the criterion over the entire parameter set, this step provides good estimates of lower-zone parameters and an approximate fitting of the hydrograph peaks.

$$\text{LOG} = \left( \frac{1}{n} \sum_{t=1}^n (\log Q_{\text{sim},t} - \log Q_{\text{obs},t})^2 \right)^{0.5} \quad (7)$$

where  $Q_{\text{sim},t}$  is simulated,  $Q_{\text{obs},t}$  is observed flow at time step  $t$ , and  $n$  is the total number of values within the time period of analysis.

**Step 2.** Fix the lower-zone parameters at the values estimated from step 1 and optimize the SAC-SMA upper-zone and percolation parameters using the RMSE as the objective function (Eq. (8)). This places more weight on reproduction of the peak flows,

$$\text{RMSE} = \left( \frac{1}{n} \sum_{t=1}^n (Q_{\text{sim},t} - Q_{\text{obs},t})^2 \right)^{0.5} \quad (8)$$

**Step 3.** Refine the lower-zone parameters by optimizing the lower-zone parameters using the LOG objective function while maintaining the upper-zone parameters at the values estimated in the previous step.

In the above discussion, we distinguish between calibration algorithms and calibration schemes. Calibration algorithms (e.g., SCE-UA) are the set of tools used to search for the optimal value of the objective functions and to identify the parameter set associated with such optima. Calibration schemes (e.g., MACS), on the other hand, define the sequential application of various objective functions (e.g., RMSE, LOG and %Bias) and methods (manual/automatic/both) to calibrate sub-sets of model parameters with the objective of improving the model's ability to capture key characteristics of the observed hydrograph. Following this, we introduce the notion of calibration strategy, which, in this case, pertains to the approach we apply to address the spatial variability of model input and parameters.

#### Calibration strategies

While the SCE-UA and MACS provide the calibration algorithm and scheme, respectively, this paper focuses on utilizing these “tools” to test various model calibration strategies that are suitable for distributed modeling. This is accomplished by conducting cali-

bration under four distinct strategies, which are then compared to the un-calibrated baseline simulation using the a priori parameters developed by Koren et al. (2003):

**Strategy (1): lumped parameters and basin averaged precipitation (L2D).** In this strategy, precipitation data are averaged over the entire basin. The optimal parameter set is estimated, through calibration of the lumped model over the entire watershed. Then, the calibrated parameter set is applied uniformly to the sub-basins constituting the semi-distributed model structure to simulate discharge at the basin outlet as well as interior points. This strategy, which was introduced by Ajami et al. (2004), is called Lumped parameters applied to Distributed structure (L2D) and it does not account for the distributed structure in the calibration process.

**Strategy (2): semi-lumped parameters and distributed precipitation forcing (SL).** Herein, and to consider the distributed model structure in the calibration, distributed precipitation forcing is averaged over each sub-basin. However, identical SAC-SMA model parameters are used at all sub-basins and model calibration, which is carried for the distributed structure, optimizes a single parameter set. This calibration strategy was also presented in Ajami et al. (2004) and hereafter is called Semi-Lumped (SL). Clearly the parameters obtained using L2D strategy do not account for the effect of the spatial variability of precipitation and routing (overland/channel). On the other hand, the parameters obtained using SL, do account for these factors, and are therefore different.

**Strategy (3): semi-distributed parameters and distributed precipitation forcing (SD).** Using the NRCS-STATSGO soil data set, which provides estimates of soil properties for 11 layers from ground surface to 2.5 m depth, the NWS developed a gridded SAC-SMA parameter data set for the United States based on Koren et al. (2003). These SAC-SMA “a priori” parameter grids are utilized in this calibration strategy. First, the a priori SAC-SMA parameters for each sub-basin as well as the entire basin are calculated by averaging the grid-scale parameters over the sub-basin and the whole basin domains. Assuming that these parameters are spatially adjustable, each adjusted parameter for each sub-basin can be defined as following:

$$P_{ij} = \frac{AP_{ij}}{AP_{bj}} \times P_j \quad (9)$$

where

- $P_{ij}$ : adjusted Parameter  $j$  for sub-basin  $i$ ;
- $AP_{ij}$ : averaged a priori parameter  $j$  for sub-basin  $i$ ;
- $AP_{bj}$ : averaged a priori parameter  $j$  for entire basin;
- $P_j$ : common parameter  $j$  for all of the sub-basins;
- $i = 1, 2, \dots, N$  sub-basin;
- $j = 1, 2, \dots, M$  parameter.

For the  $j$ th parameter, both  $AP_{ij}$  (sub-basin average) and  $AP_{bj}$  (basin average) are defined a priori and the ratio is constant for each sub-basin ( $i$ ). As such, only the common parameter  $P_j$  needs to be calibrated. The parameter estimation is accomplished by applying  $P_{ij}$ s, which are functions of  $P_j$ , to the relevant sub-basins and calibrating the distributed modeling structure with respect to the vector of  $P_j$ s. This strategy, hereafter termed Semi-Distributed (SD), reduces the dimension of the calibration problem from  $M$  times  $N$  to  $M$  parameters. Frances et al. (2007) used a similar approach to calibrate a fully distributed rainfall-runoff model (TETIS).

*Strategy (4): lumped parameters and basin averaged precipitation, modified using a priori parameters of SAC-SMA (L2D-M).* We now combine the frameworks of strategies 1 and 3. First, the calibration step is performed using lumped model to obtain optimal parameter set ( $P_j^l$ ) for the entire watershed. Then, and without further calibration, Eq. (9) is used to compute the “post-calibration” distributed parameters with the lumped optimal parameter set ( $P_j^l$ ) replacing the common parameter ( $P_j$ ) used during the calibration of scenario 3. This scenario is similar to the calibration strategy performed by Koren et al. (2004) to calibrate the NWS distributed model (HL-RDHM). Hereafter this strategy is referred to as Lumped parameter applied to Distributed structure with some Modifications (L2D-M).

*Baseline simulation: a priori parameters of SAC-SMA and distributed precipitation forcing (AP).* The objectives of the baseline simulation are: (1) to examine the applicability and performance of initial parameters based on soil properties in distributed hydrologic modeling, which is suitable for un-gauged catchments and (2) to identify gains from calibration in comparison to the best-available a priori parameterization technique. In the baseline simulation hereafter referred to as (AP), the a priori SAC-SMA parameter grids of Koren et al. (2003) are averaged over each sub-basin and used in the distributed model structure along with the distributed precipitation forcing without any calibration.

#### Calibration rule

DMIP explicitly requires the participants to calibrate their distributed models using only observed discharge at the basin outlet (Smith et al., 2004) and to disregard the observed streamflow at any interior locations during model calibration. The objective of this constraint is to assess distributed models' ability to reproduce streamflow at interior points when observations are only available at the outlet. We abide by this requirement. However, in ‘Integrating the interior point observed streamflow in model calibration’ section of this manuscript, we present a complementary study in which the observed streamflow at selected interior locations are utilized for model calibration as additional information along with the observed discharge at the basin outlet. Our goal is to explore the value of using discharge observations at interior points in model calibration by assessing gains/losses in hydrograph simulations at the basin outlet.

#### Performance measures

The four calibration strategies are used to calibrate the semi-distributed model configuration as applied to the Illinois River basin, located south of Siloam Spring, Arkansas. The historical record between October 1995 and September 2005 was divided into three periods: a warm up period from October 1995 to September 1996, a calibration period, which encompasses October 1996 through September 2002, and a validation period covering the remainder of the record (i.e. October 2002 to September 2005). The perfor-

mance of the semi-distributed model is evaluated through visual and statistical inspections with the latter relying on the following three statistical goodness-of-fit indices:

#### 1. Percent Bias (%Bias):

$$\%Bias = \frac{\sum_{i=1}^N (Q_{sim,i} - Q_{obs,i})}{\sum_{i=1}^N Q_{obs,i}} \times 100 \quad (10)$$

where  $Q_{sim,i}$  and  $Q_{obs,i}$  represent simulated and observed streamflow at time step  $i$ , respectively. As described by Smith et al. (2004), %Bias represents the total volume difference between simulated and observed fluxes. As such, negative/positive biases correspond to model under-estimation/over-estimation.

#### 2. Percent Root Mean Square Error (%RMSE):

$$\%RMSE = \frac{\left( \frac{1}{N} \sum_{i=1}^N (Q_{sim,i} - Q_{obs,i})^2 \right)^{0.5}}{\overline{Q_{obs}}} \times 100 \quad (11)$$

$\overline{Q_{obs}}$  is the mean observed discharge over the entire time period of analysis. %RMSE is a measure, which emphasizes high flow simulations.

#### 3. Modified Correlation Coefficient ( $r_{mod}$ ):

McCuen and Snyder (1975) pointed to two deficiencies in using the correlation coefficient as a goodness-of-fit measure for hydrologic models: (a) sensitivity to outliers and (b) insensitivity to the differences in the size of hydrographs. To overcome these concerns, they introduced a modified correlation coefficient  $r_{mod}$  as:

$$r_{mod} = r \times \frac{\min\{\sigma_{sim}, \sigma_{obs}\}}{\max\{\sigma_{sim}, \sigma_{obs}\}} \quad (12)$$

where  $r$  is correlation coefficient,  $\sigma_{obs}$  and  $\sigma_{sim}$  are standard deviations of observed and simulated hydrographs, respectively.

#### Results and discussion

In this study, we adopt the viewpoint of the NWS (Smith et al., 2004) that a distributed model meets operational forecasting requirements when (1) it produces reasonable simulations at interior points, (2) it produces simulations at the outlet that are generally comparable to or better than those produced by the operational lumped model, and (3) the distributed model generates better simulations than the operational lumped model in cases of highly variable forcing and/or basin characteristics.

From an operational perspective, the objective of any model calibration study is to identify a parameter set that generates the “best” streamflow hydrograph under various conditions. We believe that when multiple calibration approaches are considered, such as the case in this study, the first step is then to identify, among the strategies, a single “best performer” before the distributed model is then compared with its lumped counterpart. In the following sections, we present the results of this study in a manner consistent with this point of view. First, cross comparisons of the above-described calibration strategies and the baseline simulation are presented for both the outlet and the interior points. These comparisons will identify the calibration strategy that best meets requirement 1. Once the “best performer” strategy is identified, we proceed to compare its simulations with those produced by the operational lumped NWS SAC-SMA model at the basin outlet to assess requirement 2. The third requirement is addressed by comparing model results for selected storms with both uniform and spatially variable patterns. Finally, the results

of the complementary study, which integrates interior point information into the calibration process are presented and discussed.

#### Cross comparison of calibration strategies

In this section we follow the DMIP 2 calibration rule by calibrating the distributed model using observations only from the basin outlet gauge (DMIP gauge 13) on the Illinois River basin (USGS gauging station – 07195430) south of Siloam Spring (Siloam), Arkansas. It is important to re-iterate that the results described below represent hourly simulation runs for both calibration and validation periods.

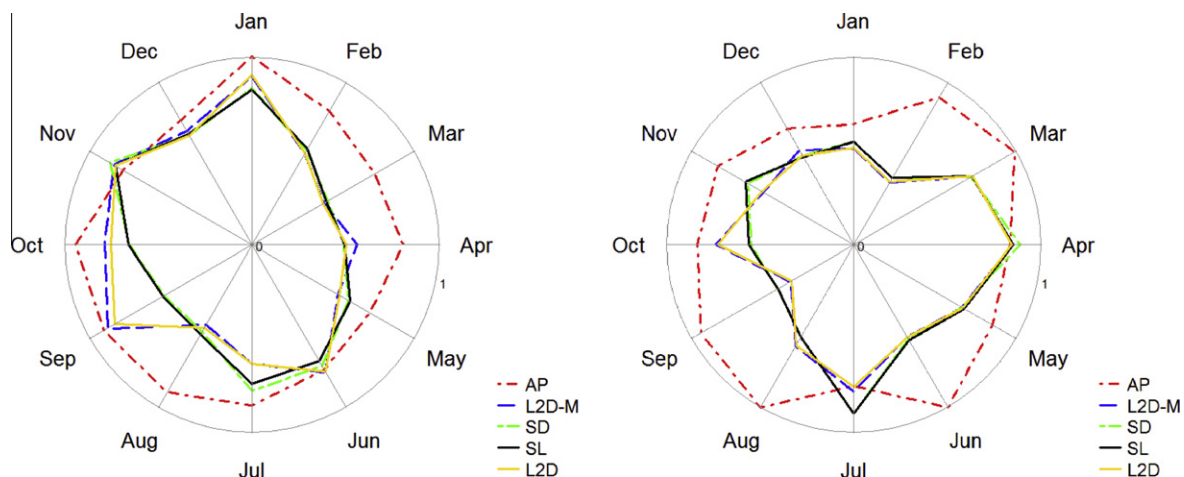
**Results at the basin outlet.** Summary statistics of the four calibration strategies (i.e. L2D, SL, SD, L2D-M) and the un-calibrated baseline simulation (AP) are presented in Table 4 for both calibration and validation periods. All four calibration strategies yielded reasonably good and rather similar performance measures as indicated by the relatively low Bias, high  $r_{\text{mod}}$ , and reasonable %RMSE. When ranked based on individual statistics (superscript with parenthesis), SL consistently outperforms all other simulations, albeit slightly, in the calibration period leading to an average rank of 1. Based on average ranking, the five strategies and baseline are ranked in the following order: SL, SD, L2D, L2D-M and AP. Noticeably, strategies based on calibrating distributed structure outranked those based on calibrating the lumped structure. However, both quantitative (%Bias, RMSE, and  $r_{\text{mod}}$ ) and qualitative (individual and averaged ranks) performance measures tell a different story during the validation period. As seen from the table, during validation, L2D-M and L2D outranked distributed calibration strategies for the semi-distributed model.

Another noteworthy observation is the phase shift in %Bias values between the calibration and validation periods. All four calibration strategies consistently over-estimate streamflow at the outlet during the calibration and under-estimate it during the validation period. The phase shift was also accompanied by an increase in the magnitude of %Bias during the validation period for all four calibration strategies as well as for the baseline simulation. With the exception of  $r_{\text{mod}}$  during the validation period, model calibration resulted in improved model performance when compared to the baseline a priori-based simulation (AP). However, the reasonable performance of the un-calibrated AP indicates that using the a priori SAC-SMA parameter grids, developed by Koren et al. (2003) is a viable parameterization scheme for un-gauged basins.

Fig. 4 shows a wind-rose comparison of a composite Euclidean distance-based performance index calculated for each month. For each calibration strategy, the absolute values of %Bias and RMSE are calculated for each individual month, and then scaled by their respective maximum monthly values from both calibration and validation periods to obtain a comparable scale for the two different measures. The composite index is then calculated as the Euclidean distance from the (0,0) point in the [0–1] space and the month of the year provides the angle required for the wind-rose plot. Ideally, a smaller area confined by such plot represents a “better” strategy. Furthermore, the plots can indicate for a given calibration strategy, whether or not the model performs consistently during the year as its plot conforms to or departs from a circular shape. It is difficult to discern from Fig. 4 which calibration strategy dominates. Although the patterns are different during the calibration and validation periods, the four calibration strategies display rather similar behaviors for each period. However, during the calibration period, distributed-based strategies (SL: solid black, and

**Table 4**  
Performance assessment of streamflow simulation at the outlet for different calibration strategies and baseline (AP) in calibration and validation periods.

Outlet	Performance measure	L2D	SL	SD	L2D-M	AP
Calibration	%Bias	7.32 <sup>(4)</sup>	5.72 <sup>(1)</sup>	5.82 <sup>(2)</sup>	7.55 <sup>(5)</sup>	–6.81 <sup>(3)</sup>
	%RMSE	114 <sup>(3)</sup>	107 <sup>(1)</sup>	110.0 <sup>(2)</sup>	115 <sup>(4)</sup>	166 <sup>(5)</sup>
	$r_{\text{mod}}$	0.77 <sup>(3)</sup>	0.84 <sup>(1)</sup>	0.83 <sup>(2)</sup>	0.76 <sup>(4)</sup>	0.61 <sup>(5)</sup>
	Average rank	3.66	1	2	4.33	4.33
Validation	%Bias	–13.29 <sup>(2)</sup>	–15.7 <sup>(4)</sup>	–15.54 <sup>(3)</sup>	–12.81 <sup>(1)</sup>	–31.36 <sup>(5)</sup>
	%RMSE	114 <sup>(1)</sup>	121 <sup>(2)</sup>	128 <sup>(3)</sup>	114 <sup>(1)</sup>	132 <sup>(4)</sup>
	$r_{\text{mod}}$	0.77 <sup>(2)</sup>	0.66 <sup>(3)</sup>	0.64 <sup>(4)</sup>	0.77 <sup>(2)</sup>	0.84 <sup>(1)</sup>
	Average rank	1.66	3	3.33	1.33	3.33



**Fig. 4.** Wind-rose comparison of a composite Euclidean distance-based performance index calculated based on monthly scaled %Bias and RMSE for calibration strategies (left panel: calibration period; right panel: validation period).



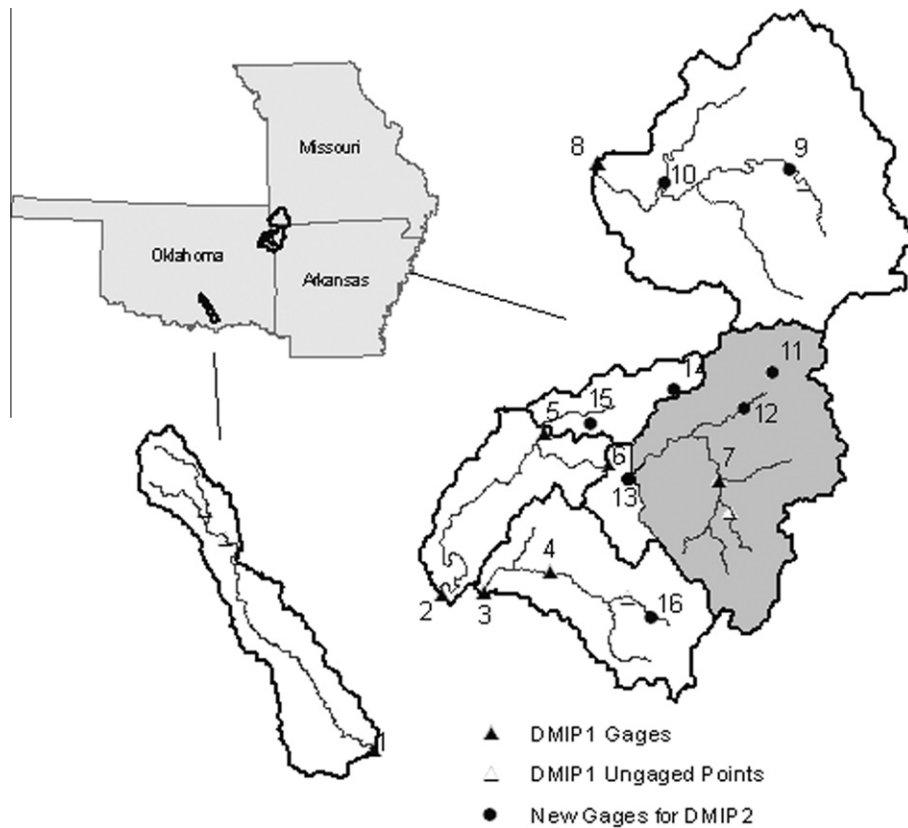


Fig. 5. DMIP 2 test basins and specified points for Oklahoma experiment (Illinois River basin at south of Siloam Spring is shown as colored area).

SD: dashed green) seem to outperform the lumped based strategies (L2D and L2D-M) albeit with minor differences. During the validation period, all four strategies perform similarly but detailed visual inspection shows that lumped strategies were better in 10 out of the 12 months than distributed calibration. This confirms the results from the overall comparison presented in Table 4. When compared against the un-calibrated “best available” a priori parameterization (AP: dashed thick red line), the wind-rose plots demonstrate that, with very few exceptions, all strategies yielded better performance.

In interpreting these results, one must recall that the study basin is reasonably uniform in physical properties. In such a case, calibrating the lumped model, and then applying the calibrated parameters to a distributed structure, may be sufficient to capture the effects of spatial variability in precipitation on the basin’s hydrologic response at the outlet. Because of the relatively short calibration and validation periods (6 and 3 years, respectively), one must also be cautious before generalizing the results of monthly performance.

Comparisons of the non-scaled %Bias and RMSE (not shown) indicate that in general and regardless of the calibration strategy, the calibrated model tends to over-estimate the streamflow at the outlet in the calibration period and under-estimate in the validation period. Again, this is consistent with the results from the overall comparison shown in Table 4.

**Simulation results at the interior locations.** One of the key science questions of both phases of DMIP is testing the hypothesis that improved simulations at the basin outlet are direct measures of the distributed model’s ability to better capture the hydrologic conditions upstream. To test this hypothesis, participants in the DMIP 2 Oklahoma experiment were asked to compare the results of their

models, calibrated at the basin outlet, with observations at three interior points (Savoy, Caves, and Elmsp, see Table 1 and Fig. 5 map if needed). As noted before, we designed the semi-distributed model structure to provide simulated hydrographs at the basin outlet as well as the three interior points. Figs. 6 and 7 show the statistical summary of simulations at Savoy, Elmsp, and Caves for the calibration and validation periods, respectively.

At Savoy, which has the largest contributing area, during the calibration period, the semi-lumped (SL) strategy, the modified lumped to distributed L2D-M, and the lumped to distributed (L2D) show the best performance with respect to %Bias, RMSE, and  $r_{mod}$ , respectively. During validation, the L2D strategy performs better with respect to both RMSE and  $r_{mod}$  but L2D-M provides the best results based on %Bias.

At Elmsp, which has the second largest contributing area, L2D shows the best performance when considering RMSE and %Bias during both calibration and validation periods. L2D yield the best  $r_{mod}$  performance during calibration, and comes second to L2D-M during validation.

Finally, for streamflow simulations at Caves, the best results are obtained across the board using L2D except for  $r_{mod}$  during validation, where L2D closely follows L2D-M.

The above discussion excludes the baseline (AP) from the comparison and focuses on the four calibration strategies to identify the best overall calibration framework. When AP is considered, the results show that AP has in general resulted in reasonably good simulations, which were at times as good as or better than those obtained using calibration at the basin outlet.

Again, and similar to the comparison at the basin outlet, all calibration strategies, as well as the baseline simulation were associated with negative bias during the validation period. The consistent underestimation of discharge at all three interior points

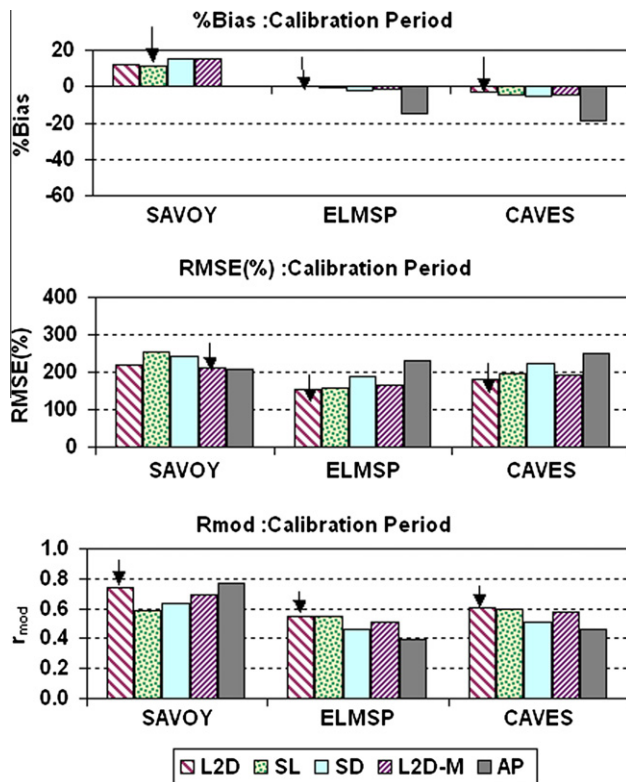


Fig. 6. Performance assessment of simulations at Savoy, Elmsp, and Caves for different calibration strategies in calibration period (the best strategy for each performance measure and interior point is marked).

explains the similar consistency of underestimation at the basin's outlet. On the other hand, unlike the comparison at the basin outlet, where distributed strategies showed better performance than lumped during calibration, and lumped calibration strategies were better performers during validation, interior point simulations point to lumped to distributed-based calibration strategies as being better more frequently than distributed strategies. This is particularly the case for L2D, which leads to better performance in 12 out of the 18 possible performance measurements (three interior points  $\times$  3 indices  $\times$  2 time periods). Given that L2D has also scored as one of the best calibration alternative in overall and monthly comparisons, the strategy will be selected as the best candidate for detailed comparison with the NWS lumped simulations, which is the subject of the next section.

It is important to note that the above analyses aiming to identify a best calibration strategy, benefits from data made available only after the submission of un-validated simulation results to DMIP 2. As seen in Table 4, a decision based only on calibration results at the outlet, would have pointed the semi-lumped strategy (SL) as being the best. This is in fact what was submitted to DMIP 2 earlier and is likely to appear in the "overall result paper" (Smith et al., this issue). By this, we note the reason for the discrepancies between the results presented in this manuscript and those appearing in Smith et al. (this issue). An important observation, however, is that our interior point comparison study, reported in the previous section, confirms the initial identification of L2D as one of the two best choices based on monthly and overall model performances in the validation period at the basin outlet (also seen in Table 4 and Fig. 4 rose diagram). Arguably, all of the calibration strategies generated reasonable-to-good simulations at interior points, therefore satisfying the first requirement of an operational distributed model forth in the beginning of this section.

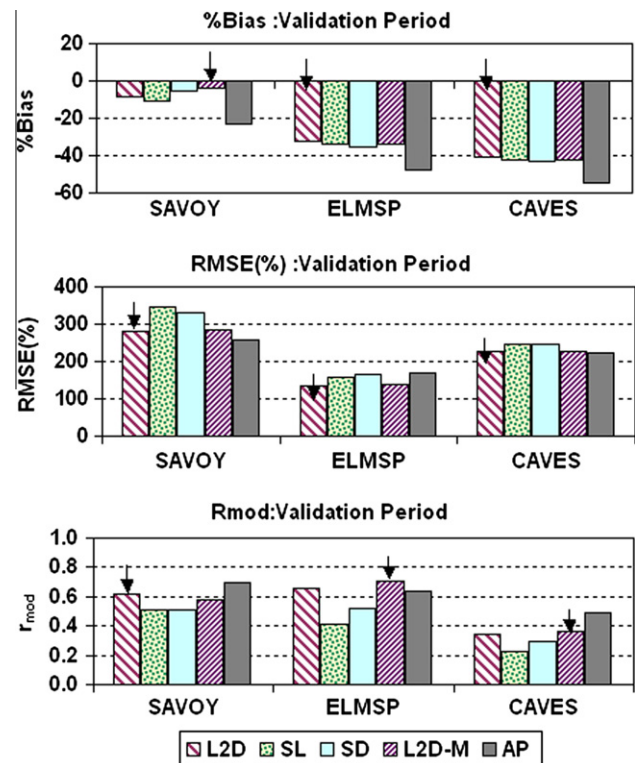


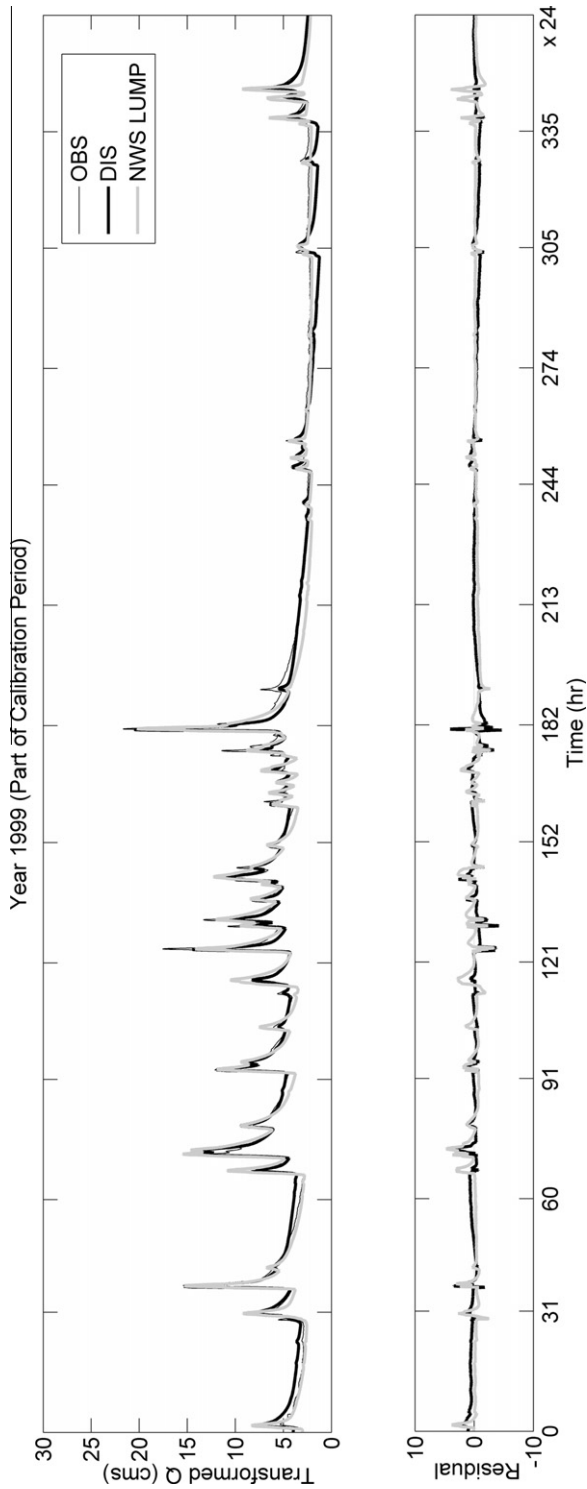
Fig. 7. Performance assessment of simulations at Savoy, Elmsp, and Caves for different calibration strategies in validation period (the best strategy for each performance measure and interior point is marked).

#### Comparison of the lumped and semi-distributed simulations at the basin outlet

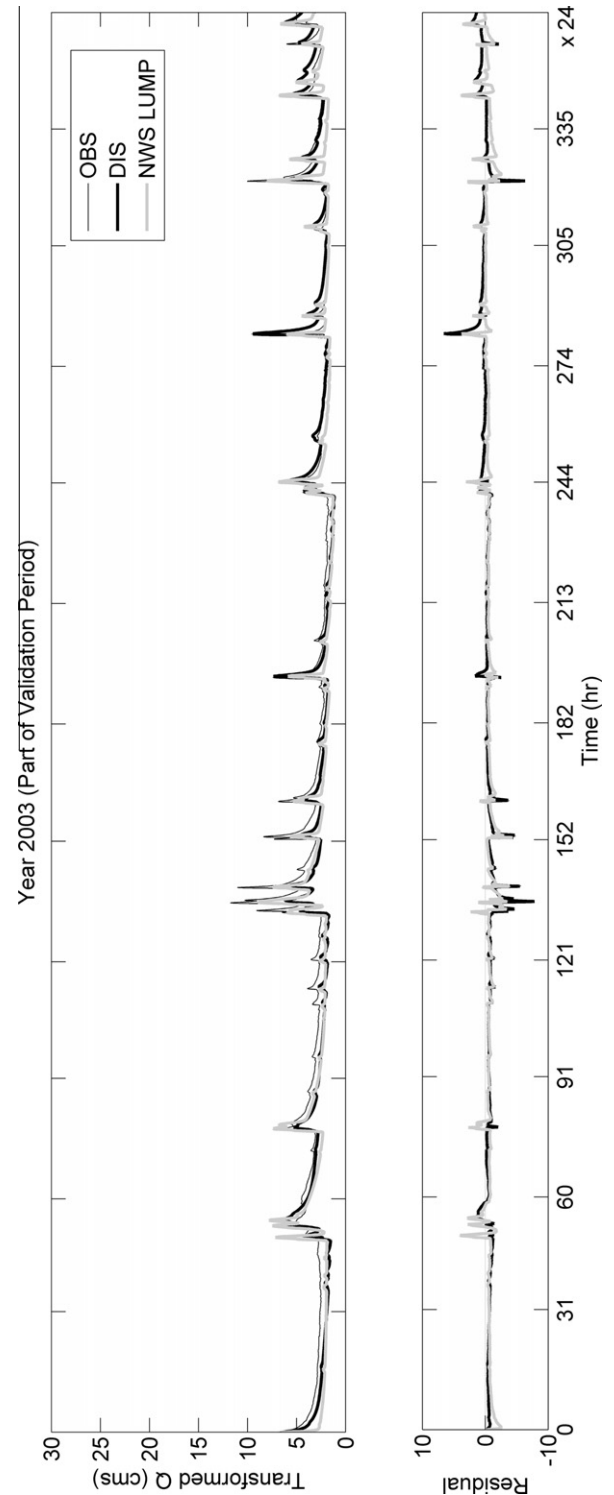
Having identified the best calibration strategy for the semi-distributed model, we now proceed to test whether the proposed model, when calibrated accordingly, satisfies the second operational requirements of a "good" distributed model as defined by Smith et al. (2004). First, an overall comparison between the best calibration strategy (L2D) and the current NWS implementation of the SAC-SMA model for the study basin is conducted to determine whether the semi-distributed model produces results at the basin's outlet that are comparable or better than the lumped model. In Figs. 8 and 9, comparisons between observed hydrograph, L2D strategy simulation, and results from the lumped SAC-SMA are presented. Fig. 8 shows the comparison for a representative year (1999) from the calibration period, and Fig. 9 shows another representative year (2003) from the validation period. To better represent the recession parts of the hydrograph while preserving a reasonable visual representation of high flows, the hydrographs in Figs. 8 and 9 were transformed using the transformation proposed by (Hogue et al., 2000) as following:

$$Q_{trans} = \frac{(Q + 1)^{0.3} - 1}{0.3} \quad (13)$$

The lower panel of each figure shows the residual of the transformed flows. In Fig. 8, both models demonstrate good performance in simulating the watershed hydrologic response during the shown calibration year as indicated by the similarity of patterns of all three hydrographs. However, closer visual inspection of the residuals of the transformed flow reveals that both models, although closely following the observations, do occasionally over/under-estimate discharge across the entire spectrum of flows. While the two models continue to preserve the general patterns



**Fig. 8.** NWS lumped and semi-distributed simulations versus observed discharge in 1999 (part of calibration period).



**Fig. 9.** NWS lumped and semi-distributed simulations versus observed discharge in 2003 (part of validation period).

of the observed hydrograph during the sample validation year (Fig. 9), there seems to be some deterioration in model performance for both models.

Fig. 10 shows four scatter plots of the observed and simulated discharge for both NWS lumped and semi-distributed L2D at the basin outlet during calibration and validation periods. At first glance, the figure shows the NWS lumped model as outperforming the semi-distributed model during the calibration period. How-

ever, the semi-distributed L2D in general performs better than NWS lumped during validation period. The quantitative analysis shown in Table 5 confirms this result with the NWS lumped model showing better RMSE, %Bias, and  $r_{\text{mod}}$  during calibration, and the semi-distributed model performing better in all three measures during the validation. Further insight into the performance of the two models can be attained by classifying the streamflow at the outlet into three categories: (1) low flows ( $0 < Q_{\text{obs}} < 10 \text{ m}^3/\text{s}$ ); (2)



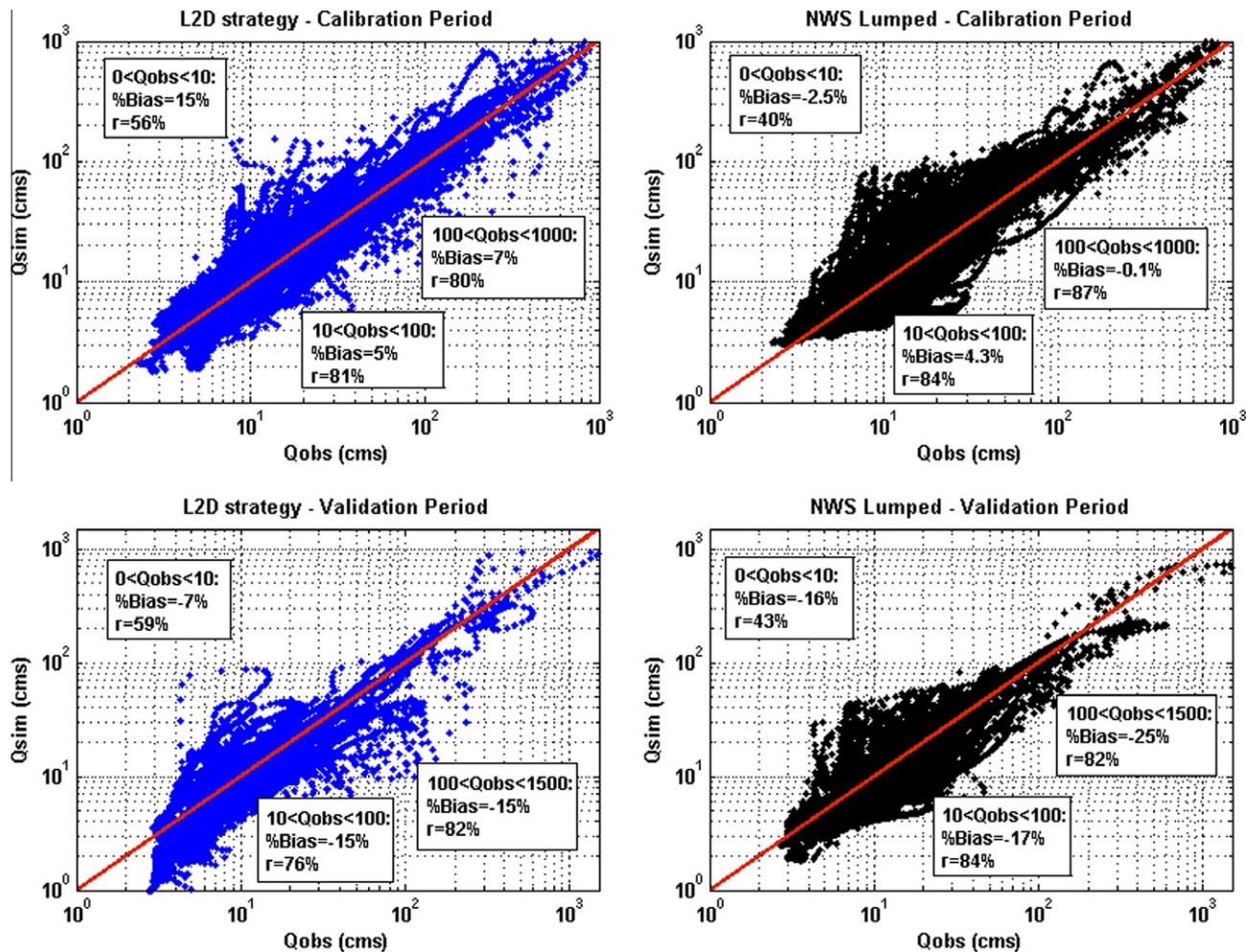


Fig. 10. Scatter plots of observed versus simulated streamflow for semi-distributed and NWS lumped during calibration and validation periods.

Table 5

Performance assessment of streamflow simulation at the outlet for the semi-distributed model with the best calibration strategy and NWS lumped SAC-SMA in calibration and validation periods.

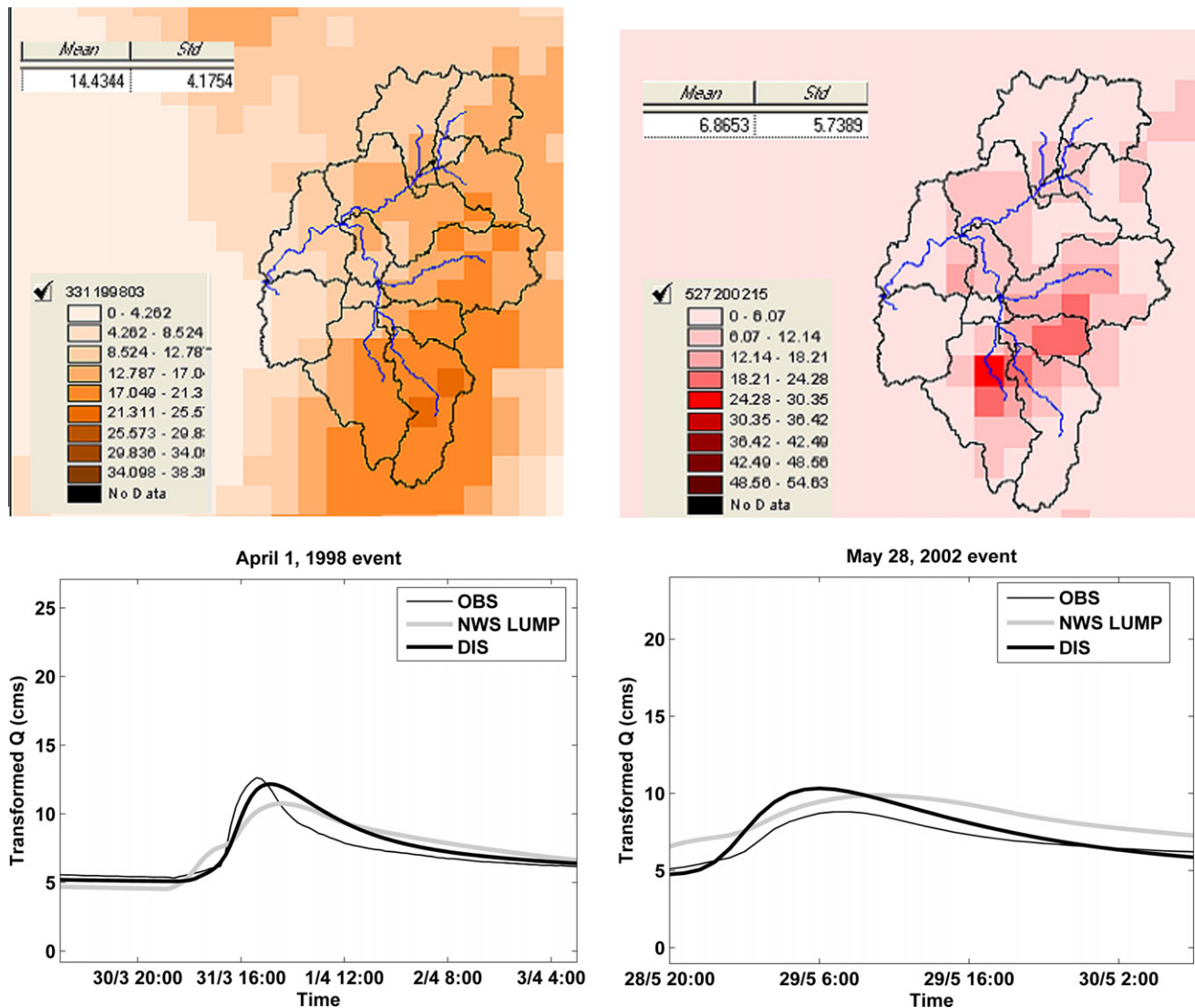
Outlet	Performance measure	Semi-distributed (L2D)	NWS lumped
Calibration	%Bias	7.32	1.92
	%RMSE	114	91
	$r_{mod}$	0.77	0.84
Validation	%Bias	-13.29	-18.33
	%RMSE	114	118
	$r_{mod}$	0.77	0.70

mid flows ( $10 \text{ m}^3/\text{s} < Q_{obs} < 100 \text{ m}^3/\text{s}$ ); and (3) high flows ( $100 \text{ m}^3/\text{s} < Q_{obs} < 1000 \text{ m}^3/\text{s}$  ( $1500 \text{ m}^3/\text{s}$  during validation)) and by computing the summary statistics of each flow category (boxes in Fig. 10). Notwithstanding the site specificity of the above classification, the computed statistics show that in general, the semi-distributed L2D tends to consistently over-estimate the streamflow during the calibration period and under-estimate it during validation period. During calibration, the NWS lumped model under-estimates the low and high flows but over-estimates the mid flows. This may explain the low %Bias result during the calibration period as the overestimation of mid-flows balances the underestimation of high and low flows. During the validation period, NWS lumped model generally under-estimates the streamflow. Noticeably, the Table 5 shows no differences between RMSE and  $r_{mod}$  of the

semi-distributed model's calibration and validation results. To summarize, results shown in Figs. 8–10, as well as the results from Table 5 demonstrate that the semi-distributed SAC-SMA model, calibrated using the L2D strategy, satisfies the first NWS requirement of operational distributed models.

The third requirement of an operational distributed model is that such a model performs as well as or better than the lumped model not only in general cases, but particularly for highly variable forcing and/or basin characteristics. Given that the study basin has fairly uniform topography, soil, and land cover, the conformity of our model to the aforementioned criterion can be tested only under conditions imposed by spatial variability of precipitation forcing. Several rainfall events with different spatial patterns ranging from fairly uniform to highly variable are selected for comparison between our semi-distributed and the NWS lumped models. Fig. 11 represents two spatially variable rainfall events. The first event (left panels) occurred on April 1, 1998, and had an average rainfall of 14 mm and standard deviation of 4 mm over the basin. The second event (right panels), which occurred on May 28, 2002, had an average rainfall of 7 mm and a standard deviation of 6 mm over the basin. The hydrographs of the NWS lumped and semi-distributed L2D along with the observed streamflow are shown in the lower panel of Fig. 11 for both storms. As seen in the figure, for the April 1, 1998 event, the semi-distributed model improves both the magnitude and timing of the peak discharge. With respect to the May 28, 2002 event, the semi-distributed model retained similar peak flow magnitude to that produced by the lumped model, but with





**Fig. 11.** Two examples of spatially variable precipitation events, occurring on April 1, 1998 (top left panel) and May 28, 2002 (top right panel). The two bottom panels show the hydrographs of observed discharge along with the NWS lumped and the semi-distributed simulations for those events.

marked improvement in the timing of peak discharge. In Fig. 12, two cases of fairly uniform rainfall events occurring on January 29, 2001, and March 19, 2002 are illustrated. The average precipitation values of these two events are 12 mm/h and 6 mm/h, and their standard deviations are approximately 3 and 2 mm/h, respectively. Under uniform rainfall conditions, the NWS lumped and the semi-distributed L2D perform equally well. Again, it is reasonable to conclude that the semi-distributed model has satisfied the second requirement for a good distributed model, yet with the caveat imposed by the basin's uniform characteristics.

#### *Integrating the interior point observed streamflow in model calibration*

In this section, we present results of the complementary study to integrate additional observed streamflow at interior points besides the observed discharge at the outlet for model calibration. While this section is outside the DMIP 2 context, assessing the value of interior point information in improving model performance at the basin outlet is an important component of the overall distributed modeling framework.

In the previous section, the L2D calibration strategy, in which parameters of the lumped model calibration are applied identi-

cally to all sub-basins in the distributed structure, was identified as the best performer. Herein, we utilize the same strategy. The following four lumped model calibrations are first conducted: (1) lumped calibration of sub-basin 1 using observed discharge at point 11 (Caves); (2) lumped calibration of the area upstream of point 12 (Elmsp), created by merging sub-basins 1–6; (3) lumped calibration of the area upstream of point 7 (Savoy), created by merging sub-basins 11, 12, 13 and 15; (4) lumped calibration of entire watershed at the outlet (Siloam), which is already made available from the previous section. Fig. 13 shows the normalized values of the four optimal parameter sets  $\theta_{11}$ ,  $\theta_{12}$ ,  $\theta_7$ , and  $\theta_o$  resulting from the above-mentioned calibration studies. As seen in the figure, the parameter set  $\theta_7$  is very close to  $\theta_o$ , indicating that  $\theta_o$ , which is obtained by lumped calibration at the outlet, can sufficiently represent the behavior of part of the basin above point 7. However, the parameter sets  $\theta_{11}$  and  $\theta_{12}$  show clear differences between each others and with  $\theta_o$  and  $\theta_7$ , thus suggesting the need for different parameterization in those areas. Consequently, only observed streamflow at points 11 and 12 are utilized by using parameter sets  $\theta_{11}$  and  $\theta_{12}$  for their respective areas and  $\theta_o$  for the remainder of the basin. Because point 11 is upstream of point 12, the parameter sets  $\theta_{11}$  and  $\theta_{12}$  can be applied using three different scenarios:

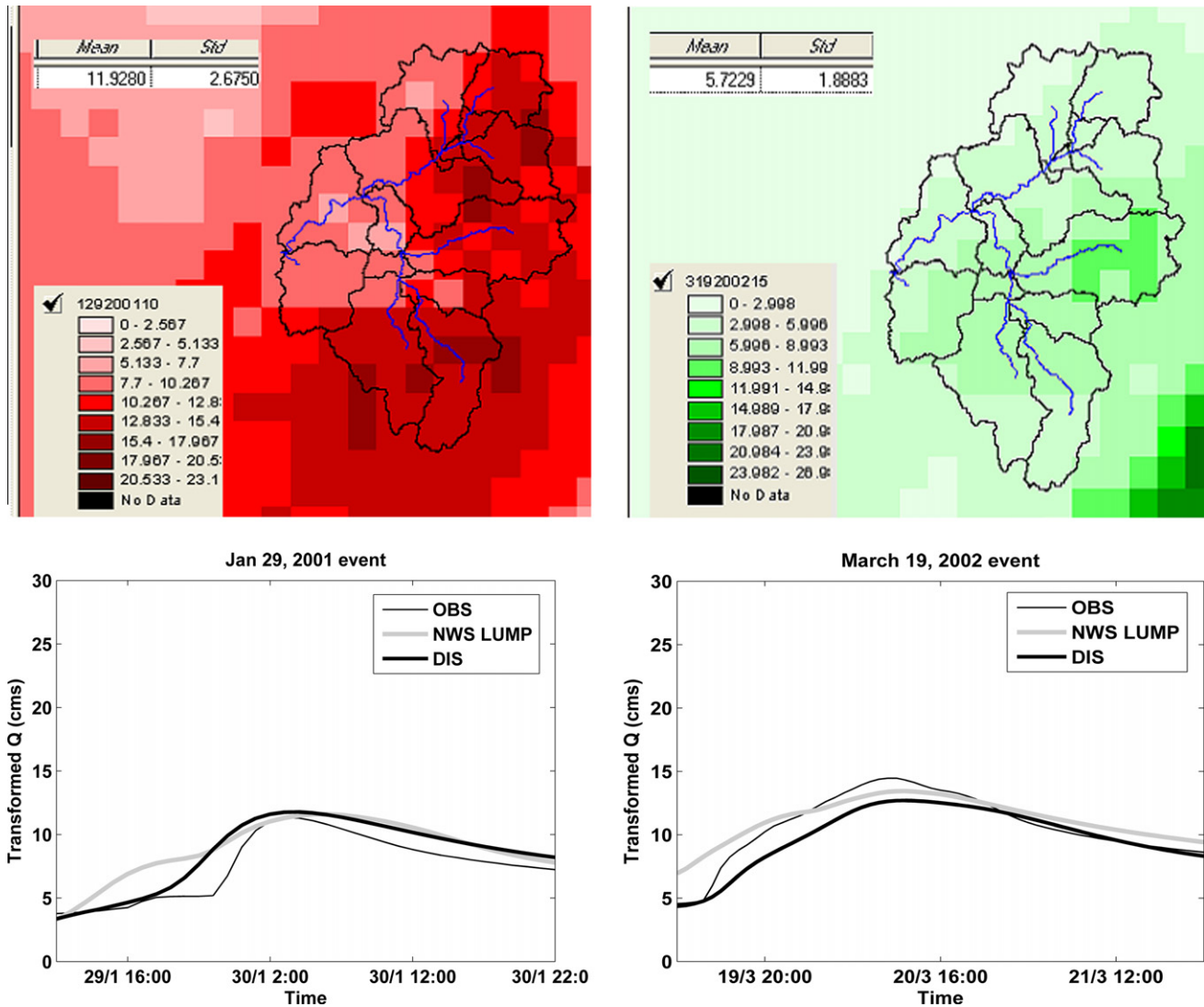


Fig. 12. Two examples of fairly uniform precipitation events, occurring on January 29, 2001 (top left panel) and March 19, 2002 (top right panel). The two bottom panels show the hydrographs of observed discharge, NWS lumped and the semi-distributed simulations for those events.

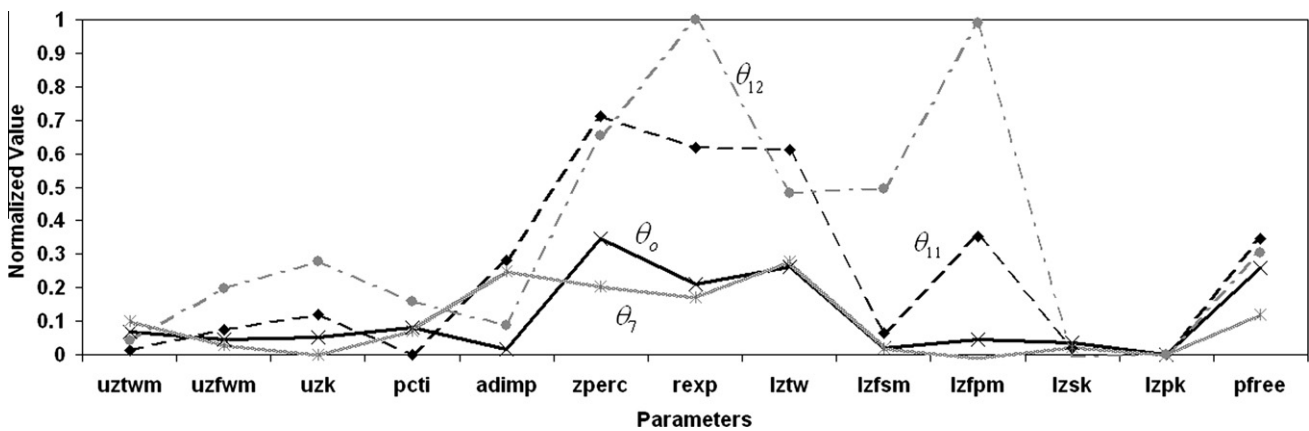


Fig. 13. Normalized parameter sets resulting from lumped calibration at the basin outlet ( $\theta_o$ ) and interior points 7 ( $\theta_7$ ), 11 ( $\theta_{11}$ ) and 12 ( $\theta_{12}$ ).

- (1)  $\theta_{11}$  is applied at sub-basin 1 and  $\theta_o$  for the remaining of sub-basins;
- (2)  $\theta_{12}$  is applied at sub-basins upstream of the point 12 including the one upstream of point 11 and  $\theta_o$  for the rest of the sub-basins; and

- (3)  $\theta_{11}$  is applied for sub-basin 1,  $\theta_{12}$  for sub-basins 2–6, and  $\theta_o$  for the rest of the sub-basins.

Statistical performance measures (%Bias, RMSE, and  $r_{mod}$ ) are obtained for each of the above-described simulations. Perfor-

mance gain or loss at the basin outlet is then computed as the relative difference between the computed measures and those associated with the best-performing (L2D) semi-distributed simulation from previous sections. Fig. 14 shows the relative differences in model performance. As seen in the figure, scenario 1 results in small performance gain in terms of %Bias and RMSE during both calibration and validation, with similarly minor gain extending to  $r_{mod}$  during the validation only. Scenario 2 is associated with significant improvements in all three measures during calibration, but not during the validation in which only RMSE improved, albeit, slightly, while %Bias and  $r_{mod}$  show small degradations. In scenario 3, which better accounts for sub-basin heterogeneities, improvements are observed during both calibration and validation for all performance measures with validation  $r_{mod}$  being the only exception. Fig. 15 shows the hydrographs of the spatially variable events that occurred in April 1 1998 and May 28 2002 for which the NWS lumped and semi-distributed L2D, using scenario 3 were compared with the observed streamflow at the outlet. In comparison with Fig. 11, which depicts the same events but without the benefit of using interior points' information, Fig. 15 shows small improvement in hydrograph peak flow for the April event and substantial improvement in simulating the May event. In both events, however, the semi-distributed model performs better than the lumped model.

## Summary and conclusions

We presented a DMIP 2 based investigation of a semi-distributed modeling framework that accounts for the spatial variability of basin characteristics and forcing. The approach, which utilizes the SAC-SMA model at sub-basins and the Kinematic Wave method for channel routing, is implemented using four different calibration strategies that consider input forcings and basin characteristics having various degrees of spatial heterogeneity. Among these calibration strategies, those based on lumped calibrations applied to semi-distributed model structure performed better than distributed calibration strategies, and were able to sufficiently account for the effects of spatial variability in precipitation on streamflow predictions both at the basin outlet and selected interior points. Arguably, these results may be influenced, to one degree or another by the apparent uniformity of the basin. However, results from the complementary study indicate that some of the sub-basins do have markedly different hydrologic behavior as indicated by the differences in the parameter sets obtained by lumped calibration using interior points. The fact that the “lumped to distributed” (L2D) calibration strategy performed well despite these differences, demonstrates the feasibility of our approach. Furthermore, factors such as over fitting and uncertainties in channel characteristics may have contributed to the apparent discrepancies in model performance between the calibration and validation periods.

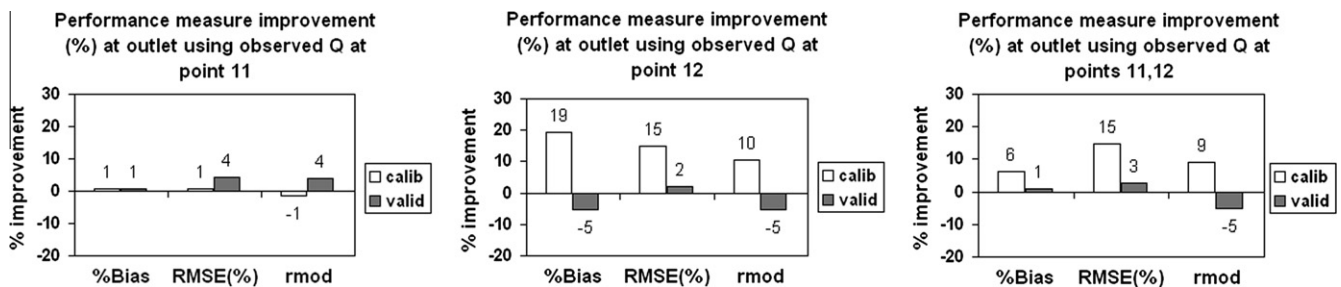


Fig. 14. Performance measure improvement at the outlet by integrating the observed discharge at interior points.

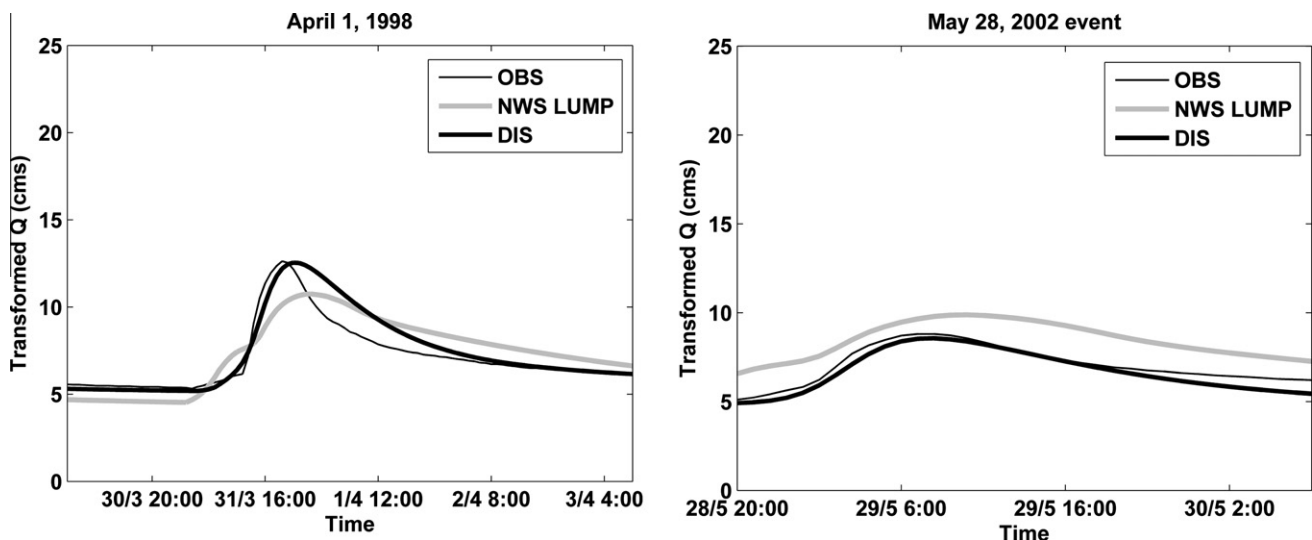


Fig. 15. Simulated hydrographs of NWS lumped and new semi-distributed (after integrating the information of interior points) versus observed discharge for two spatially variable precipitation events on April 1, 1998 and May 28, 2002.

Under operational conditions, an important consideration in selecting an appropriate distributed modeling structure is the ease of transition from a lumped model to a fully distributed (e.g. grid-based model) structure. The semi-distributed model structure presented herein, coupled with L2D calibration strategy provides exactly the type of smooth transition required to gain forecasters' support. First, it utilizes existing parameter sets, obtained from lumped calibration, while allowing the forecasters to add interior forecast points with relative ease. Second, for relatively uniform basins, the approach meets the requirements of operational distributed models.

Another finding of this study is related to the a priori SAC-SMA parameterization proposed by Koren et al. (2003). In general, when applied in a semi-distributed modeling structure, the approach resulted in relatively good performance in comparison with the various calibration strategies. We believe that the approach provides an alternative parameterization for both un-gauged and gauged catchment. In the former, it provides the "best" available estimates of the SAC-SMA parameters. In the latter case, the approach provides a good starting point for model calibration.

Finally, the results from the complementary study ('Integrating the interior point observed streamflow in model calibration') suggest that semi-distributed models can be constructed by dividing the larger basin at locations where even a short historical record may be available. The improvements in model performance, while not very large in terms of statistical measures, were significant in terms of producing better simulations at the outlet for spatially variable storms.

## Acknowledgments

Partial support for this research was provided by the National Oceanic and Atmospheric Administration and National Weather Service (contracts # DG133W-03-CQ-0021, NA04OAR4310086) and by the National Aeronautic and Space Administration (contract # NNS06AA78G). Additional support was also provided to the first author from the University of California, Irvine, Henry Samueli School of Engineering graduate student block funds. The authors would like to thank Dr. Newsha K. Ajami for her assistance with the semi-distributed model and her collaboration on implementation of the model. We also would like to thank Drs. Michael Smith and Victor Koren for their input on DMIP 2 and their comments on the study as well as the anonymous reviewers for their valuable comments. The authors also thank Mr. Dan Braithwaite for his assistance in data processing.

## References

- Ajami, N.K., Gupta, H., Wagener, T., Sorooshian, S., 2004. Calibration of a semi-distributed hydrologic model for streamflow estimation along a river system. *Journal of Hydrology* 298 (1–4), 112–135.
- Alizadeh, A., 2003. Principles of Applied Hydrology, sixteenth ed. Astan Ghods Razavi (in Persian).
- Bell, V.A., Moore, R.J., 1998. A grid-based distributed flood forecasting model for use with weather radar data 2: case studies. *Hydrology and Earth System Sciences* 2 (2–3), 283–298.
- Beven, K.J., 1985. Distributed models. In: Anderson, M.G., Burt, T.P. (Eds.), *Hydrological Forecasting*. Wiley, New York (Chapter 13).
- Beven, K., 1989. Changing ideas in hydrology – the case of physically-based models. *Journal of Hydrology* 105, 157–172.
- Bras, R.L., 1990. *Hydrology: an introduction to hydrologic science*. Addison-Wesley Series in Civil Engineering.
- Burnash, R.J.C., 1995. The NWS river forecasting-catchment modeling. In: Singh, V.J. (Ed.), *Computer Models of Watershed Hydrology*. Water Resources Publication, Highlands Ranch, Colorado, pp. 311–366.
- Burnash, R.J.C., Ferral, R.L., McGuire, R.A. Joint Federal-State River Forecast Center, 1973. A Generalized Streamflow Simulation System; Conceptual Modeling for Digital Computers. US Dept. of Commerce National Weather Service and State of California Dept. of Water Resources, Sacramento, p. 204.
- Chow, V.T., Maidment, D.R., Mays, L.W., 1988. *Applied Hydrology*. Water Resources and Environmental Engineering Series. McGraw-Hill.
- Clark, C.O., 1943. Storage and the unit hydrograph. *Proceedings of the American Society of Civil Engineers* 9, 1333–1360.
- Cole, S.J., Moore, R.J., 2008. Hydrological modelling using raingauge- and radar-based estimators of areal rainfall. *Journal of Hydrology* 358, 159–181.
- Cooper, V.A., Nguyen, V.T.V., Nicell, J.A., 1997. Evaluation of global optimization methods for conceptual rainfall-runoff model calibration. *Water Science and Technology* 36 (5), 53–60.
- Duan, Q.Y., Sorooshian, S., Gupta, V., 1992. Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resources Research* 28 (4), 1015–1031.
- Eckhardt, K., Fohrer, N., Frede, H.G., 2005. Automatic model calibration. *Hydrological Processes* 19, 651–658.
- Frances, F., Velez, J.L., Velez, J.J., 2007. Split-parameter structure for the automatic calibration of distributed hydrological models. *Journal of Hydrology* 332 (1–2), 226–240.
- Gan, T.Y., Biftu, G.F., 1996. Automatic calibration of conceptual rainfall-runoff models: optimization algorithms, catchment conditions, and model structure. *Water Resources Research* 32 (12), 3513–3524.
- Giertz, S., Diekkruiger, B., Steup, G., 2006. Physically-based modeling of hydrological processes in a tropical headwater catchment (West Africa – process representation and multi-criteria validation). *Hydrology and Earth System Sciences* 10, 829–847.
- Grayson, R.B., Moore, I.D., McMahon, T.A., 1992. Physically based hydrologic modeling: 1. A terrain-based model for investigative purposes. *Water Resources Research* 28, 2639–2658.
- Gupta, H.V., Sorooshian, S., Yapo, P.O., 1998. Toward improved calibration of hydrologic models: multiple and noncommensurable measures of information. *Water Resources Research* 34 (4), 751–763.
- Hendrickson, J., Sorooshian, S., Brazil, L.E., 1988. Comparison of Newton-type and direct search algorithms for calibration of conceptual rainfall-runoff models. *Water Resources Research* 24 (5), 691–700.
- Hogue, T.S., Sorooshian, S., Gupta, H., Holz, A., Braatz, D., 2000. A multi-step automatic calibration scheme for river forecasting models. *Journal of Hydrometeorology* 1 (6), 524–542.
- Kampf, S.K., Burges, S.J., 2007. Parameter estimation for a physics-based distributed hydrologic model using measured outflow fluxes and internal moisture states. *Water Resources Research* 43 (12), W12414. doi:10.1029/2006WR005605.
- Koren, V.I., Smith, M., Wang, D., Zhang, Z., 2000. Use of soil property data in the derivation of conceptual rainfall-runoff model parameters. In: 15th Conference on Hydrology. American Meteorological Society, Long Beach, CA, pp. 103–106 (Preprints).
- 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., Sorooshian, S., Rousseau, A., Turcotte, R. (Eds.), *Calibration of Watershed Models: Water Science and Application Series*, vol. 6. American Geophysical Union, Washington, DC.
- 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. *Journal of Hydrology* 291, 297–318.
- Krajewski, W.F., Lakshmi, V., Georgakakos, K.P., Jain, S.C., 1991. A Monte Carlo study of rainfall sampling effect on a distributed catchment model. *Water Resources Research* 27 (1), 119–128.
- Kuczera, G., 1997. Efficient subspace probabilistic parameter optimization for catchment models. *Water Resources Research* 33 (1), 177–185.
- 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. In: Duan et al. (Eds.), *Calibration of Watershed Models*. Water Science and Application, vol. 6. pp. 255–266.
- Li, R.M., Simons, D.B., Stevens, M.A., 1975. Nonlinear kinematic wave approximation for water routing. *Water Resources Research* 11 (2), 245–252.
- McCuen, R.H., Snyder, W.M., 1975. A proposed index for computing hydrographs. *Water Resources Research* 11 (6), 1021–1024.
- Michaud, J., Sorooshian, S., 1994. Comparison of simple versus complex distributed runoff models on a mid-sized semiarid watershed. *Water Resources Research* 30, 593–605.
- Nelder, J.A., Mead, R., 1965. A simplex-method for function minimization. *Computer Journal* 7 (4), 308–313.
- Ogden, F.L., Julien, P.Y., 1994. Runoff model sensitivity to radar rainfall resolution. *Journal of Hydrology* 158, 1–18.
- Peck, E.L., 1976. Catchment modeling and initial parameter estimation for the national weather service river forecast system. NOAA Technical Memorandum NWS Hydro-31.
- Reed, S., Koren, V., Smith, M., Zhang, Z., Morela, F., Seo, D.J., 2004. Overall distributed model intercomparison project results. *Journal of Hydrology* 298 (1–4), 27–60.
- Shah, S.M.S., O'Connell, P.E., Hosking, J.R.M., 1996. Modeling the effects of spatial variability in rainfall on catchment response. 2. Experiments with distributed and lumped models. *Journal of Hydrology* 175, 89–111.
- Smith, M.B., Seo, D., Koren, V.I., Reed, S.M., Zhang, Z., Duan, Q., Morela, F., Cong, S., 2004. The distributed model intercomparison project (DMIP): motivation and experiment design. *Journal of Hydrology* 298, 4–26.
- Smith, M.B., Koren, V.I., Reed, S.M., Cui, Z., Zhang, Z., Morela, F., Zhang, Y., 2007. Overview of NWS distributed model (HL-RDHM). DMIP 2 Workshop. NWS Headquarter Silver Spring, MD, September 10–12.



- Smith, M.B., Koren, V.I., Reed, S.M., Zhang, Z., Zhang, Y., Morel, F., Cui, Z., Mizukami, N., Sheldon, S., this issue. The distributed model intercomparison project: phase 2 motivation and experiment design.
- Snyder, F.F., 1938. Synthetic unit-graphs. *Transactions – American Geophysical Union* 19, 447–454.
- Soil Conservation Service, 1972. *Hydrology. National Eng. Handbook*. Springer, Washington, DC (Sec. 4).
- Sorooshian, S., Gupta, V.K., 1983. Automatic calibration of conceptual rainfall-runoff models: the question of parameter observability and uniqueness. *Water Resources Research* 19 (1), 260–268.
- Sorooshian, S., Duan, Q., Gupta, V.K., 1993. Calibration of rainfall-runoff models: application of global optimization to the Sacramento soil moisture accounting model. *Water Resources Research* 29 (4), 1185–1194.
- Thyer, M., Kuczera, G., Bates, B.C., 1999. Probabilistic optimization for conceptual rainfall-runoff models: a comparison of the shuffled complex evolution and simulated annealing algorithms. *Water Resources Research* 35 (3), 767–773.
- Troutman, B.M., 1983. Runoff prediction errors and bias in parameter estimation induced by spatial variability of precipitation. *Water Resources Research* 19 (3), 791–810.
- US Army Corps of Engineers (USACE), Hydrologic Engineering Center, 2000. *Hydrologic Modeling System, HEC-HMS. Technical Reference Manual*.
- US Army Corps of Engineers (USACE), Hydrologic Engineering Center, 2003. *Geospatial Hydrologic Modeling Extension, HEC-GeoHMS. User's Manual, Version 1.1*.
- Wang, Q.J., 1991. The genetic algorithm and its application to calibrating conceptual rainfall-runoff models. *Water Resources Research* 27 (9), 2467–2471.
- Wilson, C.B., Valdes, J.B., Rodriguez-Iturbe, I., 1979. On the Influence of the spatial distribution of rainfall on storm runoff. *Water Resources Research* 15 (2), 321–328.