New Multisite Cascading Calibration Approach for Hydrological Models: Case Study in the Red River Basin Using the VIC Model

Xianwu Xue, Ph.D.[1]; Ke Zhang, Ph.D.[2]; Yang Hong, Ph.D.[3]; Jonathan J. Gourley, Ph.D.[4]; Wayne Kellogg[5]; Renee A. McPherson, Ph.D.[6]; Zhanming Wan[7]; and Barney N. Austin, Ph.D.[8]

Abstract: A novel multisite cascading calibration (MSCC) approach using the shuffled complex evolution–University of Arizona (SCE-UA) optimization method, developed at the University of Arizona, was employed to calibrate the variable infiltration capacity (VIC) model in the Red River Basin. Model simulations were conducted at 35 nested gauging stations. Compared with simulated results using a priori parameters, single-site calibration can improve VIC model performance at specific calibration sites; however, improvement is still limited in upstream locations. The newly developed MSCC approach overcomes this limitation. Simulations using MSCC not only utilize all of the available streamflow observations but also better represent spatial heterogeneities in the model parameters. Results indicate that MSCC largely improves model performance by decreasing the number of stations with negative Nash-Sutcliffe coefficient of efficiency (NSCE) values from 69% (66%) for a priori parameters to 37% (34%) for single-site calibration to 3% (3%) for MSCC, and by increasing the number of stations with NSCE values larger than 0.5 from 9% (9%), to 23% (23%) to 34% (29%) during calibration (and validation) periods across all sites. **DOI: 10.1061/(ASCE)HE.1943-5584.0001282.**

Author keywords: Hydrological model; Multisite calibration; Optimization; Shuffled complex evolution method–University of Arizona (SCE-UA).

Introduction

Hydrological models are a useful tool for flood and drought prediction, water resources management, and climate change assessment. However, the accuracy and robustness of hydrological models are usually limited by large uncertainties in forcing data, model parameters, and sometimes model structures. Parametric uncertainties are largely caused by considerable spatial variabilities in geomorphologies, land cover, and soil characteristics (Zak and Beven 1999; Beven and Freer 2001). To model the effect of spatial variability in forcing data and basin characteristics on streamflow, distributed or semidistributed hydrological models, which are capable of using spatially distributed forcing and parameter values, are preferred in operational forecasts of river and stream hydrographs (Cosgrove et al. 2009; Yao et al. 2009; Sharif et al. 2010; Khakbaz et al. 2012; Smith et al. 2012; Furusho et al. 2013; Park et al. 2013; Xue et al. 2013; Formetta et al. 2014).

Distributed or semidistributed hydrological models contain parameters whose values depend on topography, land cover, and soil properties (e.g., texture and depth). A priori parameters derive their values from these topographic and soil characteristics in watersheds, with spatial variability dictated by in situ survey or details in remote-sensing observations (Anderson et al. 2006; Zhang et al. 2011, 2012; Yao et al. 2012). Because accurate a priori parameter estimation assigns values to the spatially variable fields (Yao et al. 2012), it offers the potential to reduce uncertainty and guide efforts in both manual and automated calibration approaches (Anderson et al. 2006; Zhang et al. 2012). Most important, a priori parameters provide valuable spatial information, sometimes completely based on remote-sensing observations, for applying hydrological models to ungauged basins (Zhang et al. 2012).

However, although the a priori parameter approach can improve model performance in some parts of a basin (Anderson et al. 2006; Zhang et al. 2012), derivation of physically based effective parameters at large scales is subject to scaling issues because of nonlinearities in processes and spatial heterogeneity (Bashford et al. 2002). Hydrological model skill can be further improved through calibration, resulting in better predictive capabilities for floods, flash floods, and water resources management (Duan et al. 1992; Zhijia et al. 2013; Choi et al. 2014; Guo et al. 2014).

[1]Postdoctoral Research Fellow, School of Civil Engineering and Environmental Sciences, Univ. of Oklahoma, Norman, OK 73019; and Hydrometeorology and Remote Sensing Laboratory and Advanced Radar Research Center, National Weather Center, Norman, OK 73072 (corresponding author). E-mail: xuexianwuqd@gmail.com

[2]Research Scientist, Cooperative Institute for Mesoscale Meteorological Studies, Univ. of Oklahoma, Norman, OK 73072.

[3]Professor, School of Civil Engineering and Environmental Sciences, Univ. of Oklahoma, Norman, OK 73019; and Dept. of Hydraulic Engineering, Tsinghua Univ., Beijing 100084, China. E-mail: yanghong@ou.edu

[4]Research Hydrometeorologist, NOAA/National Severe Storms Laboratory, Norman, OK 73072.

[5]Environmental Engineer, Chickasaw Nation Division of Commerce, Ada, OK 74820.

[6]Associate Professor, Dept. of Geography and Environmental Sustainability, Univ. of Oklahoma, Norman, OK 73019; and University Co-director, South Central Climate Science Center, Norman, OK 73072.

[7]Ph.D. Candidate, School of Civil Engineering and Environmental Sciences, Univ. of Oklahoma, Norman, OK 73019; and Hydrometeorology and Remote Sensing Laboratory and Advanced Radar Research Center, National Weather Center, Norman, OK 73072.

[8]Aqua Strategies, Dripping Springs, TX 78620.

Note. This manuscript was submitted on December 23, 2014; approved on June 26, 2015; published online on August 17, 2015. Discussion period open until January 17, 2016; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Hydrologic Engineering***, © ASCE, ISSN 1084-0699/05015019(9)/$25.00.

Conventionally, hydrologic models are calibrated only against the observed discharge at a single site, usually the watershed outlet (hereafter single-site calibration). Then the calibrated parameter values are directly applied to the whole basin. Single-site calibration can improve model performance at the outlet of the parent basin, but it may give unsatisfactory results in interior locations because it aims to achieve the best results at the reference site, which may not properly reflect the spatial variation of the geomorphological characteristics of the subbasins (Ricard et al. 2013; Choi et al. 2014). Simultaneously improving the simulations at both the outlet and the inner locations is thus a major challenge to hydrologic modelers, which is why multisite calibration is needed to utilize observations from as many stations as possible to improve the accuracy of the simulated results for any location in a watershed (Ajami et al. 2004; Choi et al. 2014).

Xue (2010) applied a multisite calibration method to a hydrological model over a karstic aquifer, resulting in more stable and accurate results than single-site calibration. Khakbaz et al. (2012) suggested that model simulations at the outlet could be further improved by being calibrated at both interior points and the outlet. Smith et al. (2012) applied multisite calibration to nested subbasins using the MIKE SHE distributed hydrologic model for a large watershed. Their results showed that multisite calibration slightly decreased model performance at two out of three stations, but greatly improved results at the third station. Thus, they argued that multisite calibration generally has advantages over single-site calibration. Choi et al. (2014) also showed that multisite calibration improved model simulations at the calibrated sites and even at some noncalibrated sites.

All of the studies just mentioned applied calibration at multiple sites independently. In other words, they did not account for the topographical and geomorphological connections between these stations and basins. To further advance the multisite calibration approach, a new multisite cascading calibration approach using the SCE-UA (shuffled complex evolution–University of Arizona) optimization method was developed to take advantage of observations from all nearby gauging stations, to account for hydrological connections between stations, and to maximize the representation of spatial heterogeneities of model parameters in a watershed by optimizing those parameters from upstream to downstream in a cascading sequence.

The newly developed calibration approach was applied to calibrate the VIC model over the Red River Basin in the south-central United States. The Red River Basin has its headwaters in the mountains of New Mexico and drains generally from west to east into the Mississippi River. Its water serves as a major source for ecosystems, tourism and recreation, drinking water supply, agriculture, and cultural ceremonies in Oklahoma and Texas. Major metropolitan areas in Oklahoma and Texas plan to further use the surface waters of the Red River Basin for increasing water supplies in the future.

The objectives of this research were (1) to develop a novel multisite cascading calibration approach to improve VIC model performance in the Red River Basin, including its subbasins, and (2) to evaluate simulations by comparing them with the results using a priori parameter values and single-site calibrated parameter values. This paper is organized as follows: the study area and methodology are introduced and described next; simulation results are then presented and discussed. Finally, major findings are summarized and conclusions are drawn.

Study Area and Methods

Study Area

The Red River Basin in the south-central United States encompasses approximately 239,361 km^2 across portions of New Mexico, Texas, Oklahoma, Arkansas, and Louisiana. It overlaps with three U.S. Department of the Interior landscape conservation cooperatives (LCCs)—the Great Plains, the Gulf Coast Prairie, and the Gulf Coastal Plains and Ozarks—and is within the purview of the U.S. Department of the Interior South Central Climate Science Center. The Red River stretches 2,189 km from the New Mexico Mountains to the Mississippi River and has five reaches (Fig. 1). Elevations range from 5 m at the outlet to 1,505 m at the headwater origin to the west (Fig. 1). The Red River originates in the High Plains of New Mexico near the Texas border and flows through the Gulf Coast Plains before draining into the Mississippi River.

Water resources in the Red River Basin have been stressed in recent years because of frequent drought, extreme flooding, species' needs for environmental flows, and increasing demands from consumptive users (Famiglietti and Rodell 2013). In addition, major metropolitan areas in Oklahoma and Texas view the surface waters in the basin as viable sources for future water supplies. For these reasons, it is of great importance to understand the basin's hydrological conditions and enable hydrologic models to accurately simulate the hydrological processes in this region.

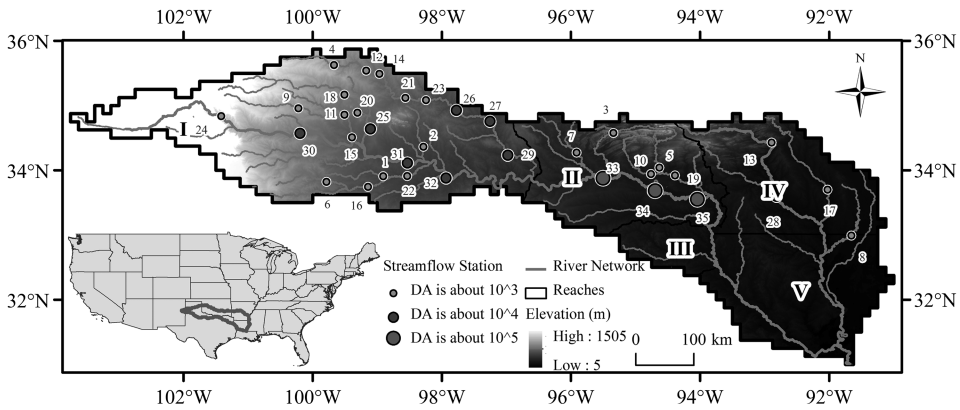

Fig. 1. Study domain showing the Red River Basin in the south-central United States (the thick outline), its river network and reaches, and the locations of its 35 stream gauges

Hydrological Model and Data

The VIC (Liang et al. 1994, 1996) model is a macroscale, semidistributed hydrologic model that has been successfully applied in many regional and global hydrologic studies (Nijssen et al. 2001; Maurer et al. 2002; Mitchell et al. 2004; Rodell et al. 2004; Yong et al. 2010, 2014; Xia et al. 2012; Zhang et al. 2014). It is also a process-based model that simulates snow pack, canopy interception, evapotranspiration, surface runoff, base flow, and other hydrological processes at daily or subdaily time steps. One of the model's distinguishing features is that it solves full water and energy balances within each grid cell at each time step for multiple elevation bands and vegetation types, in this way capturing variability at the subgrid scale. It computes runoff generation components using both saturation and infiltration excess runoff processes in a model grid cell through a statistical parameterization of subgrid heterogeneity. A nonlinear semidistributed conceptual rainfall-runoff model (ARNO) (Franchini and Pacciani 1991; Todini 1996) is used to compute the base flow. The impulse response function (IRF) method, which predicts the lumped response of the contributing area (watershed) at arbitrary locations of interest, is used as a postprocessor for river routing (Li et al. 2013). Additional details of the VIC model can be found in Wood et al. (1992), Liang et al. (1994, 1996), Ferguson et al. (2012), and Li et al. (2013).

The VIC model is forced by either daily or subdaily surface meteorological data, which include precipitation, temperature (maximum and minimum), wind, vapor pressure, incoming long-wave and shortwave radiation, and air pressure. The latest version (4.1.2.h) was used in this study.

The precipitation and temperature forcing data are based on the daily parameter elevation regressions on independent slopes model (PRISM) precipitation and temperature data produced by the PRISM group at Oregon State University (Daly et al. 2008). The PRISM data set consists of 4-km gridded estimates of precipitation and temperature for the continental United States (CONUS) based on observations from a wide range of monitoring networks with sophisticated quality control and bias and orographic corrections. PRISM's interpolation method calculates climate elevation regressions for each grid cell; stations entering a regression are assigned weights based primarily on the physiographic similarity of the station to the cell. Factors considered are location, elevation, coastal proximity, topographic facet orientation, vertical atmospheric layer, topographic position, and orographic enhancement caused by the underlying terrain. The heart of the model is its extensive spatial climate knowledge base that calculates station weights on entering the regression function. In this study, the gridded daily precipitation and temperature (minimum and maximum) were aggregated to 1/8°to match the study's modeling spatial resolution using an area-weighted method (hereafter the aggregated PRISM data are denoted Upscaled-PRISM).

Using an area-weighted approach, the 1/8°digital elevation model (DEM) was derived from the 30-arcsec DEM from hydrological data and maps based on shuttle elevation derivatives at multiple scales (HydroSHEDS). Flow direction data were from the global river-routing network data set produced by Wu et al. (2011).

To calibrate and validate the VIC model, streamflow data in the Red River Basin were obtained from the U.S. geological survey (USGS) (http://nwis.waterdata.usgs.gov/nwis). The available USGS stream-gauging stations in this study region were screened by four strict quality control criteria: (1) observed daily streamflow data from all selected stations had to be available from January 1981 to December 2012; (2) the drainage areas of these selected stations had to contain at least 10 grid cells; (3) the absolute relative difference between drainage areas derived from the DEM and flow direction data and reported by USGS had to be less than 20% for each station; and (4) if more than one station was located within the same $0.125 \times 0.125°$ grid cell, the station with the larger drainage area was selected for further analysis. A total of 35 stations met the aforementioned criteria (Table 1). These stations captured streamflow covering approximately 65% of the entire Red River Basin.

Model Calibration Strategies

Selection of Model Parameters for Calibration

In addition to the elevation and meteorological forcing data, the VIC model requires a large number of parameters that are related to soil and vegetation types or conceptual terms. Most of the parameters can be derived from soil physical properties and vegetation types using the a priori approach developed by Maurer et al. (2002). Instead of calibrating all VIC model parameters, the six most sensitive were selected for calibration following the sensitivity analysis proposed by Demaria et al. (2007) and the online documentation for model calibration (http://www.hydro.washington.edu/Lettenmaier/Models/VIC/Documentation/Calibration.shtml). These parameters were the shape of the variable infiltration capacity curve (b), the thickness of the second soil layer [$d2$ (cm)], the thickness of the third soil layer [$d3$ (cm)], the maximum base flow velocity from the lowest soil layer [$Ds_{\max}$ (mm day^{-1})], the fraction of $Ds_{\max}$ where nonlinear base flow occurs (Ds), and the fraction of the maximum soil moisture of the lowest soil layer (Ws) where nonlinear base flow occurs. Parameter b controls the amount of water that can infiltrate into the soil. A higher value of b results in a lower infiltration rate and higher surface runoff, whereas the $d2$ and $d3$ parameters affect water availability for transpiration and base flow. Usually, less runoff (both surface and base flow) is generated in the thicker soil depths. To represent soil moisture dynamics near the surface, the first soil layer ($d1$) is defined as a constant ($d1 = 10$ cm) for all grid cells, following Liang et al. (1996). The valid ranges of $Ds_{\max}$, Ds, and Ws are (0,30] mm day^{-1}, (0, 1], and (0, 1], respectively (http://www.hydro.washington.edu/Lettenmaier/Models/VIC/Documentation/CalibrateSoil.shtml).

Calibration Methods

To effectively and automatically calibrate the VIC model and maximize its performance, a novel multisite calibration approach called MSCC was developed. MSCC sorts the subbasins from upstream to downstream based on drainage area, and for each one prepares the parameter ranges and default values. For each subbasin, SCE-UA generates the parameter files as inputs for *VIC.exe* and *Routing.exe*. *Routing.exe* routes the grid-simulated runoff and base flow from *VIC.exe* to output the simulated streamflow at the outlet. The 1.0-NSCE values are calculated and then passed to SCE-UA until the optimization criteria are met. Finally, MSCC outputs a calibrated parameter file (Fig. 2).

In this research, the VIC model was calibrated at all 35 stations in sequence, from upstream to downstream, following the topographic relationships of these subbasins, as shown in Fig. 3. For example, Stations 4, 12, 14, 21, 23, 26, 27, and 29 were all situated on the same stream (Figs. 1 and 2). Model parameters in the areas draining to Station 4 were first calibrated and optimized by comparing the simulated values with the observations at that station. These parameters above Station 4 were then fixed. Subsequently, model parameters in the areas between Station 4 and Station 12 were calibrated to achieve the best result by comparing the modeled streamflow values against the observations at Station 12. Following

Table 1. Characteristics of the 35 USGS Streamflow Gauging Stations Used in This Study and the Peak Flow Qualification Codes of the USGS Streamflow Stations

Number	USGS_ID	Longitude	Latitude	Drainage area (km^2)	Delineated area (km^2)	Area relative error (%)	Peak flow qualification code
1	**07312200**	**98.91 W**	**33.91 N**	**1,689**	**1,923**	**13.86**	**6**
2	**07311000**	**98.28 W**	**34.36 N**	**1,797**	**1,905**	**6.00**	**6**
3	07335790	95.34 W	34.57 N	1,810	2,065	14.08	5
4	**07316500**	**99.67 W**	**35.63 N**	**2,007**	**1,727**	**−13.98**	**5**
5	07339000	94.63 W	34.04 N	2,072	1,912	−7.74	6
6	**07311700**	**99.79 W**	**33.82 N**	**2,427**	**2,564**	**5.67**	**0**
7	07334000	95.91 W	34.27 N	2,820	2,863	1.50	5
8	07364200	91.66 W	32.99 N	3,074	2,739	−10.90	0
9	**07300000**	**100.22 W**	**34.96 N**	**3,165**	**2,847**	**−10.04**	**6**
10	07338500	94.76 W	33.94 N	3,181	3,352	5.38	6
11	**07300500**	**99.51 W**	**34.86 N**	**3,766**	**3,957**	**5.08**	**5**
12	**07324400**	**99.17 W**	**35.54 N**	**3,952**	**3,298**	**−16.57**	**6**
13	07359002	92.89 W	34.43 N	4,014	3,816	−4.93	6
14	**07325000**	**98.96 W**	**35.49 N**	**5,079**	**5,025**	**−1.05**	**6**
15	**07301110**	**99.39 W**	**34.51 N**	**5,136**	**4,751**	**−7.49**	**5**
16	**07312100**	**99.14 W**	**33.74 N**	**5,403**	**5,779**	**6.96**	**6**
17	07363500	92.03 W	33.70 N	5,439	6,383	17.37	0
18	**07301500**	**99.51 W**	**35.17 N**	**6,869**	**7,248**	**5.53**	**5**
19	07340000	94.39 W	33.92 N	6,889	7,184	4.27	6
20	**07303000**	**99.31 W**	**34.89 N**	**7,335**	**7,723**	**5.29**	**6**
21	**07325500**	**98.56 W**	**35.12 N**	**8,070**	**7,237**	**−10.33**	**6**
22	**07312500**	**98.53 W**	**33.91 N**	**8,133**	**8,984**	**10.47**	**6**
23	**07326500**	**98.24 W**	**35.08 N**	**9,428**	**9,603**	**1.86**	**5**
24	**07297910**	**101.41 W**	**34.84 N**	**9,723**	**10,618**	**9.21**	**5**
25	**07305000**	**99.10 W**	**34.64 N**	**11,810**	**12,940**	**9.57**	**6**
26	**07328100**	**97.77 W**	**34.93 N**	**12,349**	**11,502**	**−6.86**	**6**
27	**07328500**	**97.25 W**	**34.75 N**	**13,742**	**12,929**	**−5.92**	**6**
28	07362000	92.82 W	33.60 N	13,882	14,702	5.91	6
29	**07331000**	**96.98 W**	**34.23 N**	**18,575**	**18,340**	**−1.27**	**6**
30	**07299540**	**100.19 W**	**34.57 N**	**20,008**	**19,359**	**−3.24**	**6**
31	**07308500**	**98.53 W**	**34.11 N**	**53,276**	**54,446**	**2.20**	**5**
32	**07315500**	**97.93 W**	**33.88 N**	**74,392**	**76,534**	**2.88**	**6**
33	07335500	95.50 W	33.88 N	115,112	116,467	1.18	6
34	07336820	94.69 W	33.68 N	122,424	124,616	1.79	5
35	07337000	94.04 W	33.55 N	124,319	126,386	1.66	6

Note: For the last column, 0 = natural or unregulated or undefined; 5 = some regulation; 6 = regulation or diversion. Bold stations are located in upstream, Reach I; the other stations are located in the downstream, Reaches II–V.

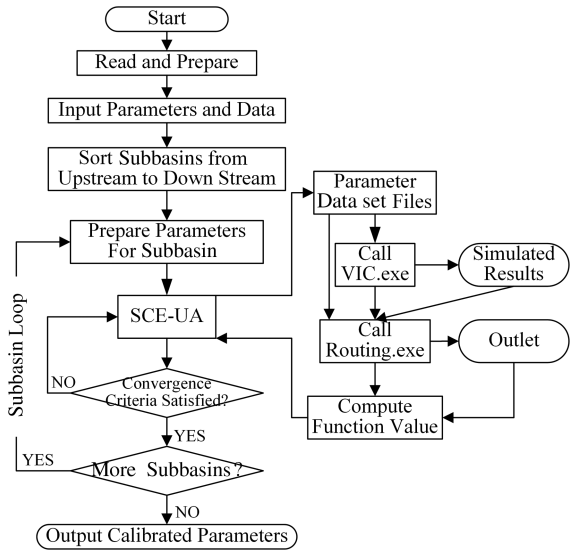

Fig. 2. MSCC Flowchart using SCE-UA and the VIC model

this, model parameters in the areas between Station 12 and Station 14 were calibrated and determined. By following the upstream-to-downstream topographic relationships, model parameters in the regions between Station 27 and Station 29 were finally calibrated. The calibrations from Station 4 to Station 29 formed a cascading calibration chain: $4 \rightarrow 12 \rightarrow 14 \rightarrow 21 \rightarrow 23 \rightarrow 26 \rightarrow 27 \rightarrow 29$ (Fig. 3). in other words, before model calibration was conducted at a given Station i, it had been performed at all upstream stations; at Station i it was only used to calibrate and optimize model parameters for these areas, excluding all drainage areas of the upstream "children" stations. Similarly, Stations 6, 16, 1, and 22 formed another calibration chain, which was composed of two chains of children: $6 \rightarrow 16$ and 1 (Fig. 3). during each step of the calibration chain, the SCE-UA method was applied to calibrate the model parameters in the specific drainage areas as described previously (Sorooshian and Gupta 1983; Duan et al. 1993, 1994; Xue et al. 2013).

The MSCC method has several advantages: (1) it can sufficiently utilize observations from all gauging stations in a large basin; (2) it accounts for the spatial heterogeneities of model parameters across a basin; (3) it is automated and objective; (4) it achieves better model performance than conventional single-site calibration and nonnested multisite calibration in theory; and (5) it is straightforward to implement using parallel-computing algorithms.

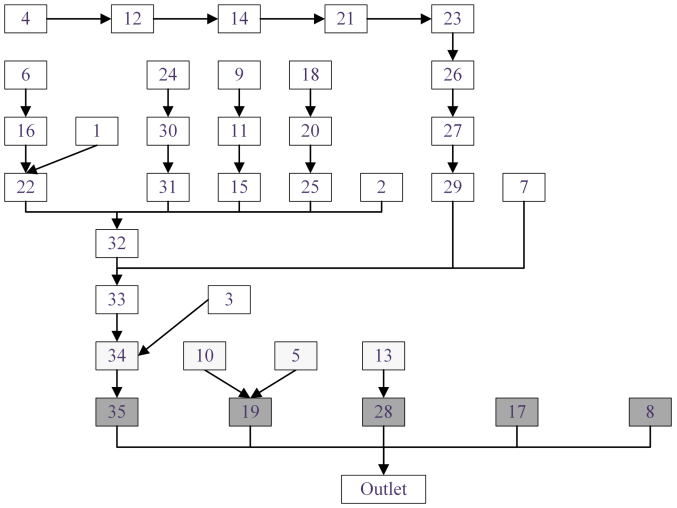

Fig. 3. Topological relationships of the 35 stream gauges and their nested subbasins; the shaded boxes correspond to the stations at basin outlets used for single-site calibration

The study period, from 1981 to 2012, was divided into three parts: warm-up (1981–1982), calibration (1983–1990), and validation (1991–2012). The model was run during the warm-up period to minimize the impact of uncertain initial conditions in the region and to bring the soil states to equilibrium. To evaluate the effectiveness of the newly developed MSCC method, the simulations using MSCC calibration were compared with results from two other calibration strategies: a priori model parameters (i.e., no calibration) and single-site calibration.

A priori parameter values were obtained directly from Maurer et al. (2002). Although there was no calibration for simulations using them, the study period was divided into calibration and validation periods to be consistent with the simulations using the other calibration strategies. In single-site calibration, the VIC model was calibrated against observed streamflow data at Stations 8, 17, 19, 28, and 35 (Figs. 1 and 3), which control five independent subbasins of the Red River Basin (Fig. 1). Three of these stations (19, 28, and 35) have additional upstream gauging stations, but the observations from them were withheld.

Model Performance Metrics

To quantify the performance of the model simulations, three widely used statistical metrics were used in this study: Nash-Sutcliffe coefficient of efficiency (NSCE), correlation coefficient (CC), and percent bias (PB). The NSCE was used to determine the relative magnitude of residual variance in the simulations compared with observed variance, and to assess the predictive power of a hydrological model. It is defined as (Nash and Sutcliffe 1970):

$$\text{NSCE} = 1 - \frac{\sum_{i=1}^{n} (\text{OBS}_i - \text{SIM}_i)^2}{\sum_{i=1}^{n} (\text{OBS}_i - \overline{\text{SIM}}_i)^2} \quad (1)$$

where n = total number of pairs of simulated and observed streamflow; SIM and OBS = simulated and observed streamflow, respectively; $i = i$th value of simulated and observed streamflow; and $\overline{\text{SIM}}$ and $\overline{\text{OBS}}$ = mean values of simulated and observed streamflow, respectively.

The NSCE can range from $-\infty$ to 1. A NSCE of 1 indicates perfect agreement between simulated and observed streamflow. Positive NSCE values are generally viewed as acceptable levels of performance, whereas negative values indicate that the model is a worse predictor than the observed mean, which is often deemed unacceptable (Moriasi et al. 2007).

The CC is used to assess the degree of linear association between simulated and observed streamflow values and is defined as

$$\text{CC} = \frac{\sum_{i=1}^{n} (\text{OBS}_i - \overline{\text{OBS}})(\text{SIM}_i - \overline{\text{SIM}})}{\sqrt{\sum_{i=1}^{n} (\text{OBS}_i - \overline{\text{OBS}})^2 \sum_{i=1}^{n} (\text{SIM}_i - \overline{\text{SIM}})^2}} \quad (2)$$

The PB measures the relative difference between modeled and observed streamflow:

$$\text{PB} = \frac{\sum_{i=1}^{n} \text{SIM}_i - \sum_{i=1}^{n} \text{OBS}_i}{\sum_{i=1}^{n} \text{OBS}_i} \times 100 \quad (3)$$

The optimal PB value is 0.0%. In this study, 1.0-NSCE was selected as the objective function for the SCE-UA calibration procedure during the calibration period.

Results and Discussion

Daily Model Results for the Calibration Period

Fig. 4 shows the values of the model performance metrics of the calibrated daily simulations at 35 stations for the three calibration strategies. In the case of no calibration (i.e., using a priori parameter values), 69% of the 35 stations show negative NSCE values whereas only 9% show NSCE values larger than 0.5 [Fig. 4(a)]. This indicates that the VIC model generally performed poorly in the Red River Basin using default, a priori parameter values, and highlights a need to calibrate the model before applying it to hydrological simulations. Once single-site calibration was conducted at 5 stations (Stations 8, 17, 19, 28, and 35), model performance improved in terms of NSCE values: 23% of the stations show NSCE values larger than 0.5 whereas the percentage showing negative NSCE values decreased to 37% [Fig. 4(d)]. These 5 stations are located in five independent subbasins and are not connected with each other; therefore, their drainage areas can be regarded as parallel basins. In the case of single-site calibration at the basin outlets, calibration in each station or subbasin was conducted independently and separately. Once the new MSCC calibration method was applied, the NSCE values across the basin substantially improved relative to the results of the single-site calibrations: 34% of the 35 stations show a NSCE value of more than 0.5 whereas only 1 station shows a negative NSCE value (Fig. 4).

Similar findings in terms of model simulation improvements from no calibration to single-site calibration to multisite cascading calibration are observed in the performance metrics of CC [Figs. 4(b, e, and h)] and PB [Figs. 4(c, f, and i)]. Only 46% of the 35 stations show a CC value larger than 0.5 for no calibration [Fig. 4(b)]. In contrast, the percentages increase to 57% for single-site calibration [Fig. 4(e)] and 60% for MSCC [Fig. 4(h)]. The daily model simulations show PB values between −15 and 15% only at 6 of the 35 stations [Fig. 4(c)], whereas 12 and 29 stations show PB values of [−15%, 15%] for single-site [Fig. 4(f)] and MSCC calibration [Fig. 4(i)], respectively.

The model performance metrics in all three calibration cases show a similar clear spatial pattern in which model performance increases from upstream to downstream stations (Fig. 4). This may be due to the mountainous and complex terrain in the upstream areas (Reach I), uniform parameter values in subbasins, uncertainty in the forcing data, and the impact of regulation or diversion (Table 1). The downstream gradient in improved simulations

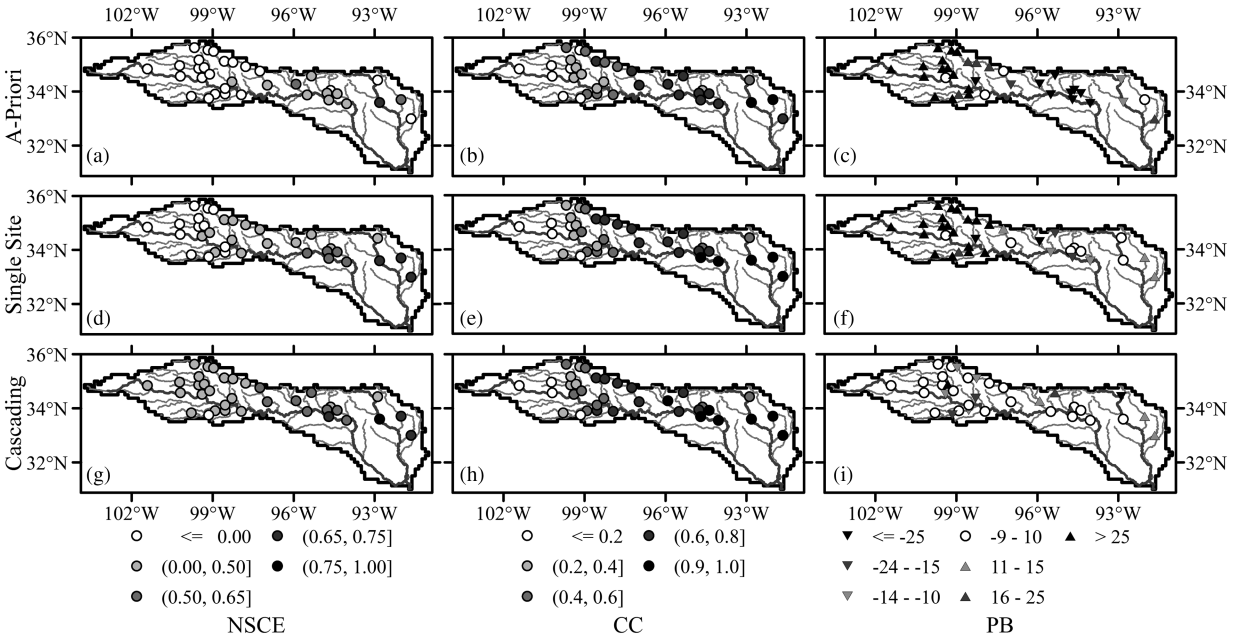

Fig. 4. Spatial maps of skill metrics for daily simulations during the calibration period using (a–c) a priori parameters, (d–f) single-site calibration, and (g–i) MSCC; the three metrics used were Nash-Sutcliffe coefficient of efficiency (NSCE), correlation coefficient (CC), and percent bias (PB)

was more apparent for no calibration and single-site calibration, partly because these upstream stations are well displaced from the selected stations for calibrations, suggesting that a well-calibrated model at one station cannot warrant good or similar skill in the upstream or neighboring basins. This finding is similar to that of Choi et al. (2014).

Daily Model Results for Validation Period

To examine the predictive capabilities of the calibrated VIC model and the effectiveness of the calibration methods, model skill was evaluated during the validation period using the parameter values determined by a priori parameter estimation and the different calibration strategies. All three daily simulation metrics show similar spatial patterns and magnitudes during both the validation period (Fig. 5) and the calibration period (Fig. 4). That is, statistical performance generally increases from west to east or from upstream to downstream locations or from Reach I to Reaches II, IV, and V for all methods. Also, the greatest improvements occur with MSCC at the upstream, or Reach I, western stations (Figs. 4 and 5).

Of the 35 stations, 12 had positive NSCE values using a priori parameters during the validation period [Fig. 5(a)]. This number increased to 23 stations for the single-site calibration method [Fig. 5(d)] and to 34 stations for the MSCC method [Fig. 5(g)].

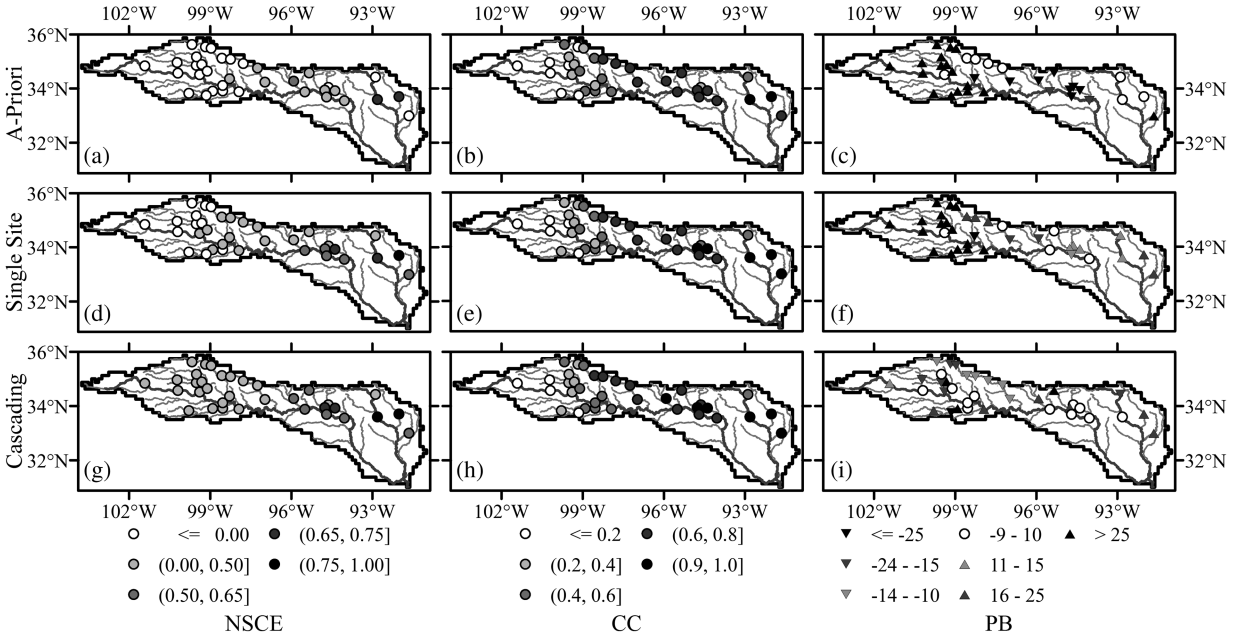

Fig. 5. Spatial maps of skill metrics for daily simulations during the validation period using (a–c) a priori parameters, (d–f) single-site calibration, and (g–i) MSCC; the three metrics used were Nash–Sutcliffe coefficient of efficiency (NSCE), correlation coefficient (CC), and percent bias (PB)

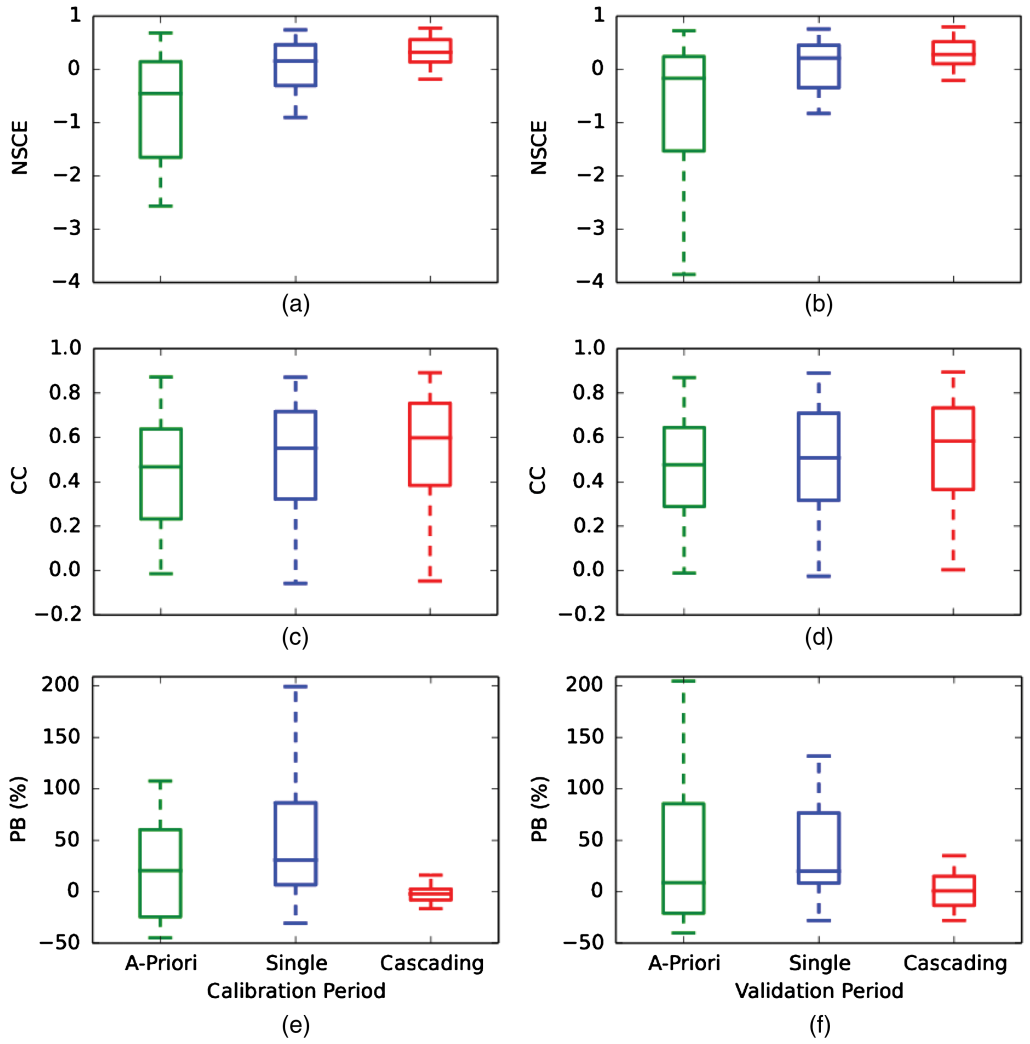

Fig. 6. Box-and-whisker plot of the three model skill metrics of daily simulations at the 35 stations using a priori parameter values, single-site calibration, and MSCC; the three metrics evaluated were (a and b) Nash–Sutcliffe coefficient of efficiency (NSCE), (c and d) correlation coefficient (CC), and (e and f) percent bias (PB); the panels on the left correspond to the calibration period and those on the right correspond to the validation period

The statistics were compared during the validation period at the five stations situated at the independent subbasins that were used for the single-site calibration. Two of these stations (8 and 17) show no upstream, interior gauges (Figs. 1 and 3), so the statistical results from the single-site method were the same as from MSCC [Figs. 5(d–i)]. Comparisons at the other three stations used for single-site calibration (19, 28, and 35) show that the MSCC method yields slight improvements over the single-site method at the basin outlets according to all statistical measures [Figs. 5(d–i)]. This result indicates that the cascading approach of optimizing parameter values in the upstream subbasins has value.

Application of MSCC to the VIC model produced reliable simulations of surface hydrological processes in the Red River Basin over a period of years. Fig. 6 shows the distributions of the NSCE, CC, and PB statistics for all simulations (including the calibration and validation periods). The top, middle, and bottom of the box represent the 75th percentile, median, and 25th percentile, respectively; the top and bottom lines show the maximum and minimum value, respectively. It is clear that the distribution of the NSCE metric had shifted to higher values from simulations using a priori parameters to those using single-site calibration and to those using MSCC, suggesting that MSCC leads to better results [Figs. 6(a and b)]. Similarly, the distribution of the CC statistic in the MSCC results had shifted to higher values than those in the a priori parameter and single-site calibration results [Figs. 6(c and d)]. At the same time, the mean and spread of the PB in the MSCC results largely decreased in the a priori parameter and single-site calibration results [Figs. 6(e and f)], indicating that the MSCC method shows a substantial improvement over the other methods in the PB—the most important statistic for water balance studies. This implies that the MSCC-calibrated VIC model can be used in future and past long-term simulation and drought analysis to account for the impacts of climate change. These results can be evaluated at interior points in the Red River Basin, enabling reliable simulations of anthropogenic changes (both direct and inadvertent) to the basin's water balance, such as adding reservoirs, diverting water out of the basin, and increasing evapotranspiration rates.

Summary and Conclusion

In this study, the MSCC method was introduced and applied to a semidistributed hydrological model to achieve optimum skill simultaneously at the basin outlet and at interior, nested locations.

The MSCC method works by calibrating model parameters in the most upstream basin, fixing those parameter values, and then cascading downstream to calibrate model parameters in the areas that drain to the next gauged location, and so on. The method yielded systematically improved simulations compared with when it was run with a priori parameters and when it was calibrated at the basin outlet alone. Validation statistics at three basin outlets with upstream stations showed improvement with MSCC when compared with single-site calibration at the basin outlets. This indicates that the method succeeds in capturing subbasin-scale variability in model parameters.

The results of this study indicate that spatial variations in model parameters are critical to the simulation of hydrological processes. The MSCC method takes advantage of observations from multiple sites and uses them to constrain the model in a nested, cascading, and unified way. Simulations using MSCC can not only utilize all of the available observations but also better represent the spatial heterogeneities in model parameters, and it can improve simulations of spatially variable hydrological processes. In this way, it provides scientifically sound data and information to natural and cultural resource managers for analyzing the impacts of climate change on streamflow in the Red River and making informed decisions regarding future use and diversion of both stream flow and reservoir water storage in the basin.

The MSCC method is most applicable to instrument-rich areas and large river basins that are more typically gauged. It was successfully applied in this study on the 239, 361-km^2 Red River Basin in the south-central United States. There is no clear relationship between model performance and regulation because the USGS only uses binary codes to describe the degree of regulation. Future studies will use the MSCC-calibrated hydrologic model to evaluate long-term impacts on water resources from proposed diversions, reservoirs, and climate change.

Acknowledgments

This study was supported by the USGS South Central Climate Science Center at the University of Oklahoma through Grant #G13AC00386, "Impacts of Climate Change on Flows in the Red River Basin." Partial support for the second author was provided by the Disaster Relief Appropriations Act of 2013 (P.L. 113-2), which funded NOAA research grant NA14OAR4830100. The authors also appreciate Mr. Cody Hudson of INTERA for sharing the script to download the USGS observed streamflow data and Mr. Manabendra Saharia and Mr. Humberto Vergara for peak flow qualification of USGS streamflow stations. The authors gratefully acknowledge the editor, associate editor, and anonymous reviewers for their valuable and constructive comments on this manuscript.

References

Ajami, N. K., Gupta, H., Wagener, T., and Sorooshian, S. (2004). "Calibration of a semi-distributed hydrologic model for streamflow estimation along a river system." *J. Hydrol.*, 298(1–4), 112–135.

Anderson, R. M., Koren, V. I., and Reed, S. M. (2006). "Using SSURGO data to improve Sacramento model a priori parameter estimates." *J. Hydrol.*, 320(1–2), 103–116.

Bashford, K. E., Beven, K. J., and Young, P. C. (2002). "Observational data and scale-dependent parameterizations: Explorations using a virtual hydrological reality." *Hydrol. Process.*, 16(2), 293–312.

Beven, K. J., and Freer, J. (2001). "Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology." *J. Hydrol.*, 249(1–4), 11–29.

Choi, Y. S., Choi, C. K., Kim, H. S., Kim, K. T., and Kim, S. (2014). "Multi-site calibration using a grid-based event rainfall-runoff model: A case study of the upstream areas of the Nakdong River basin in Korea." *Hydrol. Process.*, 29(9), 2089–2099.

Cosgrove, B., Reed, S., Ding, F., Zhang, Y., Cui, Z., and Zhang, Z. (2009). "Flash flood forecasting for ungauged locations with NEXRAD precipitation data, threshold frequencies, and a distributed hydrologic model." *Proc.,World Environmental and Water Resources Congress*, ASCE, Reston, VA, 1–10.

Daly, C., et al. (2008). "Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States." *Int. J. Climatol.*, 28(15), 2031–2064.

Demaria, E. M., Nijssen, B., and Wagener, T. (2007). "Monte Carlo sensitivity analysis of land surface parameters using the variable infiltration capacity model." *J. Geophys. Res.: Atmos.*, 112(D11), D11113.

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

Duan, Q., Sorooshian, S., and Gupta, V. K. (1994). "Optimal use of the SCE-UA global optimization method for calibrating watershed models." *J. Hydrol.*, 158(3–4), 265–284.

Duan, Q. Y., Gupta, V. K., and Sorooshian, S. (1993). "Shuffled complex evolution approach for effective and efficient global minimization." *J. Optim. Theory Appl.*, 76(3), 501–521.

Famiglietti, J. S., and Rodell, M. (2013). "Environmental science. Water in the balance." *Science*, 340(6138), 1300–1301.

Ferguson, C. R., Wood, E. F., and Vinukollu, R. K. (2012). "A global intercomparison of modeled and observed land-atmosphere coupling." *J. Hydrometeorol.*, 13(3), 749–784.

Formetta, G., Antonello, A., Franceschi, S., David, O., and Rigon, R. (2014). "Hydrological modelling with components: A GIS-based open-source framework." *Environ. Model. Software*, 55, 190–200.

Franchini, M., and Pacciani, M. (1991). "Comparative analysis of several conceptual rainfall-runoff models." *J. Hydrol.*, 122(1–4), 161–219.

Furusho, C., Chancibault, K., and Andrieu, H. (2013). "Adapting the coupled hydrological model ISBA-TOPMODEL to the long-term hydrological cycles of suburban rivers: Evaluation and sensitivity analysis." *J. Hydrol.*, 485, 139–147.

Guo, J., Zhou, J. Z., Lu, J. Z., Zou, Q., Zhang, H. J., and Bi, S. (2014). "Multi-objective optimization of empirical hydrological model for streamflow prediction." *J. Hydrol.*, 511, 242–253.

Khakbaz, B., Imam, B., Hsu, K., and Sorooshian, S. (2012). "From lumped to distributed via semi-distributed: Calibration strategies for semi-distributed hydrologic models." *J. Hydrol.*, 418–419, 61–77.

Li, H., et al. (2013). "A physically based runoff routing model for land surface and earth system models." *J. Hydrometeorol.*, 14(3), 808–828.

Liang, X., Lettenmaier, D. P., Wood, E. F., and Burges, S. J. (1994). "A simple hydrologically based model of land surface water and energy fluxes for general circulation models." *J. Geophys. Res.*, 99(D7), 14415–14428.

Liang, X., Wood, E. F., and Lettenmaier, D. P. (1996). "Surface soil moisture parameterization of the VIC-2L model: Evaluation and modification." *Global Planet. Change*, 13(1–4), 195–206.

Maurer, E. P., Wood, A. W., Adam, J. C., Lettenmaier, D. P., and Nijssen, B. (2002). "A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States." *J. Clim.*, 15(22), 3237–3251.

Mitchell, K. E., et al. (2004). "The multi-institution North American land data assimilation system (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system." *J. Geophys. Res.: Atmos.*, 109, D07S90.

Moriasi, D. N., Arnold, J. G., Liew, M. W. V., Bingner, R. L., Harmel, R. D., and Veith, T. L. (2007). "Model evaluation guidelines for systematic quantification of accuracy in watershed simulations." *ASABE*, 50(3), 885–900.

Nash, J. E., and Sutcliffe, J. V. (1970). "River flow forecasting through conceptual models. I: A discussion of principles." *J. Hydrol.*, 10(3), 282–290.

Nijssen, B., O'Donnell, G. M., Lettenmaier, D. P., Lohmann, D., and Wood, E. F. (2001). "Predicting the discharge of global rivers." *J. Clim.*, 14(15), 3307–3323.

Park, C., Lee, J., and Koo, M.-H. (2013). "Development of a fully-distributed daily hydrologic feedback model addressing vegetation, land cover, and soil water dynamics (VELAS)." *J. Hydrol.*, 493, 43–56.

Ricard, S., Bourdillon, R., Roussel, D., and Turcotte, R. (2013). "Global calibration of distributed hydrological models for large-scale applications." *J. Hydrol. Eng.*, 10.1061/(ASCE)HE.1943-5584.0000665, 719–721.

Rodell, M., et al. (2004). "The global land data assimilation system." *Bull. Am. Meteorol. Soc.*, 85(3), 381–394.

Sharif, H., Sparks, L., Hassan, A., Zeitler, J., and Xie, H. (2010). "Application of a distributed hydrologic model to the November 17, 2004, flood of Bull Creek watershed, Austin, Texas." *J. Hydrol. Eng.*, 10.1061/(ASCE)HE.1943-5584.0000228, 651–657.

Smith, M. B., et al. (2012). "The distributed model intercomparison project–Phase 2: Motivation and design of the Oklahoma experiments." *J. Hydrol.*, 418–419, 3–16.

Sorooshian, S., and Gupta, V. K. (1983). "Automatic calibration of conceptual rainfall-runoff models: The question of parameter observability and uniqueness." *Water Resour. Res.*, 19(1), 260–268.

Todini, E. (1996). "The ARNO rainfall-runoff model." *J. Hydrol.*, 175(1–4), 339–382.

Wood, E. F., Lettenmaier, D. P., and Zartarian, V. G. (1992). "A land-surface hydrology parameterization with subgrid variability for general circulation models." *J. Geophys. Res.: Atmos.*, 97(D3), 2717–2728.

Wu, H., Kimball, J. S., Mantua, N., and Stanford, J. (2011). "Automated upscaling of river networks for macroscale hydrological modeling." *Water Resour. Res.*, 47(3), W03517.

Xia, Y., et al. (2012). "Continental-scale water and energy flux analysis and validation for North American land data assimilation system project phase 2 (NLDAS-2). II. Validation of model-simulated streamflow." *J. Geophys. Res.: Atmos.*, 117(D3), D03110.

Xue, X., et al. (2013). "Statistical and hydrological evaluation of TRMM-based multi-satellite precipitation analysis over the Wangchu Basin of Bhutan: Are the latest satellite precipitation products 3B42V7 ready for use in ungauged basins?" *J. Hydrol.*, 499, 91–99.

Xue, X. W. (2010). "Study of karstic watershed hydrological model and parameters regionalisation in Southwest China." College of Hydrology and Water Resources, Hohai Univ., Nanjing, 122.

Yao, C., Li, Z. J., Bao, H. J., and Yu, Z. B. (2009). "Application of a developed grid-Xinanjiang model to Chinese watersheds for flood forecasting purpose." *J. Hydrol. Eng.*, 10.1061/(ASCE)HE.1943-5584.0000067, 923–934.

Yao, C., Li, Z. J., Yu, Z. B., and Zhang, K. (2012). "A priori parameter estimates for a distributed, grid-based Xinanjiang model using geographically based information." *J. Hydrol.*, 468, 47–62.

Yong, B., et al. (2010). "Hydrologic evaluation of multisatellite precipitation analysis standard precipitation products in basins beyond its inclined latitude band: A case study in Laohahe Basin, China." *Water Resour. Res.*, 46(7), W07542.

Yong, B., et al. (2014). "Intercomparison of the version-6 and Version-7 TMPA precipitation products over high and low latitudes basins with independent gauge networks: Is the newer version better in both real-time and post-real-time analysis for water resources and hydrologic extremes?" *J. Hydrol.*, 508, 77–87.

Zak, S. K., and Beven, K. J. (1999). "Equifinality, sensitivity and predictive uncertainty in the estimation of critical loads." *Sci. Total Environ.*, 236(1–3), 191–214.

Zhang, X. J., Tang, Q. H., Pan, M., and Tang, Y. (2014). "A long-term land surface hydrologic fluxes and states dataset for China." *J. Hydrometeorol.*, 15(5), 2067–2084.

Zhang, Y., Zhang, Z., Reed, S., and Koren, V. (2011). "An enhanced and automated approach for deriving a priori SAC-SMA parameters from the soil survey geographic database." *Comput. Geosci.*, 37(2), 219–231.

Zhang, Z., et al. (2012). "SAC-SMA a priori parameter differences and their impact on distributed hydrologic model simulations." *J. Hydrol.*, 420–421, 216–227.

Zhijia, L., Penglei, X., and Jiahui, T. (2013). "Study of the Xinanjiang model parameter calibration." *J. Hydrol. Eng.*, 10.1061/(ASCE)HE.1943-5584.0000527, 1513–1521.