Characterization of watershed model behavior across a hydroclimatic gradient

Kathryn van Werkhoven,[1] Thorsten Wagener,[1] Patrick Reed,[1] and Yong Tang[1]

Received 15 June 2007; revised 20 September 2007; accepted 8 October 2007; published 24 January 2008.

[1] A fundamental tradeoff exists in watershed modeling between a model's flexibility for representing watersheds with different characteristics versus its potential for overparameterization. This study uses global sensitivity analysis to investigate how a commonly used intermediate-complexity model, the Sacramento Soil Moisture Accounting Model (SAC-SMA), represents a wide range of watersheds with diverse physical and hydroclimatic characteristics. The analysis aims to establish a detailed understanding of model behavior across watersheds and time periods with the ultimate objective to guide model calibration and evaluation studies. Sobol's sensitivity analysis is used to evaluate the SAC-SMA in 12 Model Parameter Estimation Experiment (MOPEX) watersheds in the US. The watersheds span a wide hydroclimatic gradient from arid to humid systems. Four evaluation metrics reflecting base flows, midrange flows, peak flows, and long-term water balance were used to comprehensively characterize trends in sensitivity and model behavior. Results show significant variation in parameter sensitivities that are correlated with the hydroclimatic characteristics of the watersheds and time periods analyzed. The sensitivity patterns are consistent with the expected dominant processes and demonstrate the need for moderate model complexity to represent different hydroclimatic regimes. The analysis reveals that the primary model controls for some aspects of the simulated hydrograph are different from those typically assumed for the SAC-SMA. Results also show that between 6 and 10 parameters are regularly identifiable from daily hydrologic data, which is about twice the range that is often assumed (i.e., 3 to 5). Synthesized results provide comprehensive SAC-SMA calibration guidance, demonstrate the flexibility of the model for representing multiple hydroclimatic regimes, and highlight the great difficulty in generalizing model behavior across watersheds.

Citation: van Werkhoven, K., T. Wagener, P. Reed, and Y. Tang (2008), Characterization of watershed model behavior across a hydroclimatic gradient, *Water Resour. Res.*, *44*, W01429, doi:10.1029/2007WR006271.

1. Introduction

[2] Watershed-scale hydrologic models are essential for flood and drought prediction, water resources planning and allocation, erosion and sedimentation studies, nonpoint source pollution and remediation, climate and land use change assessments, hydropower operations, etc. [*Singh and Frevert*, 2006]. The extensive array of models that have been developed to date include both simple [e.g., *Jakeman et al.*, 1990] and highly complex structures [e.g., *Abbott et al.*, 1986; *Reggiani et al.*, 2000]. For some applications, such as operational forecasting, a single model structure may be used to represent a wide range of watersheds with varying physical and hydroclimatic characteristics. Such cases may require a model with sufficient flexibility (and therefore complexity) to represent the different watersheds. However, as a consequence of increasing model flexibility (and/or complexity) there is an associated increase in the number of model parameters that must be estimated. The potential for equifinality and overparameterization thus also increases for complex models, resulting in parameter values that are not always easily identifiable in the calibration process [*Beven*, 1989]. Studies have shown [e.g., *Jakeman and Hornberger*, 1993; *Wagener et al.*, 2003; *Wagener and Wheater*, 2006] that as model complexity increases, the number of unidentifiable parameters also increases, preventing (for those unidentifiable parameters) the possibility of locating one parameter value that is any better than another in the calibration process.

[3] Sensitivity analysis has become a popular tool in watershed modeling to explore high-dimensional parameter spaces, assess parameter identifiability, and understand sources of uncertainty. [*Hornberger and Spear*, 1981; *Freer et al.*, 1996; *Saltelli et al.*, 1999; *Wagener et al.*, 2001, 2003; *Wagener*, 2003; *Hall et al.*, 2005; *Muleta and Nicklow*, 2005; *Sieber and Uhlenbrook*, 2005; *Bastidas et al.*, 2006; *Pappenberger et al.*, 2006, 2008; *Van Griensven et al.*, 2006; *Demaria et al.*, 2007; *Tang et al.*, 2007b, 2007c]. In this context, sensitivity analysis is commonly used to determine which parameters have a significant impact on the model response and should be the focus of estimation efforts, and conversely, which have an insignificant impact (e.g., owing to overparameterization) and could be fixed to some a

[1]Department of Civil and Environmental Engineering, Pennsylvania State University, University Park, Pennsylvania, USA.

0043-1397/08/2007WR006271

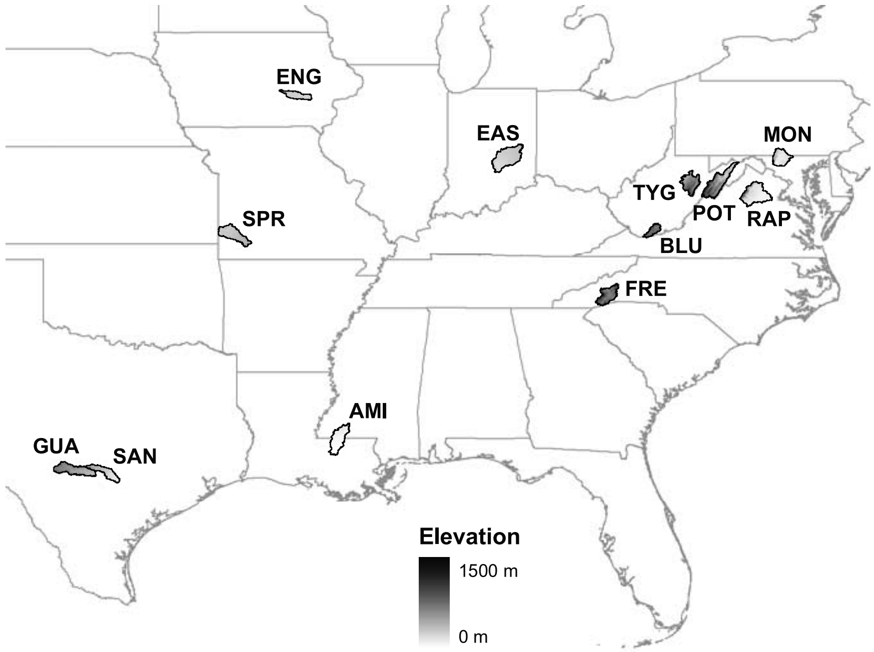

Figure 1. Location and elevation of 12 Model Parameter Estimation Experiment (MOPEX) watersheds.

priori or regional estimates. In several previous studies, model output sensitivity to parameter values (herein referred to as ''parameter sensitivity'') has been shown to vary significantly across watersheds, time periods and timescales, and evaluation metrics [*Wagener et al.*, 2001; *Sieber and Uhlenbrook*, 2005; *Demaria et al.*, 2007; *Tang et al.*, 2007c, 2007b]. However, no studies have characterized this variation for a single model across a well-defined hydroclimatic gradient using multiple metrics and time periods/timescales. Therefore an understanding of model behavior, and its dependency on hydroclimatic regime, remains limited.

[4] For the case in which a model is used to simulate watersheds with widely varying characteristics, an assessment of why and how parameter sensitivities vary across watersheds for a suite of flow condition metrics can help to determine if the model structure is fully exploited and/or if it is overparameterized in all cases. Some studies have suggested that relatively few (e.g., 3 to 5) parameters are identifiable from observations of streamflow for hydrologic models [*Jakeman and Hornberger*, 1993]. However, the existence of significant sensitivity variation across watersheds, time periods/timescales, and evaluation metrics would suggest that the number of identifiable parameters found in their study is more a function of the experimental design and cannot be generalized. A comprehensive analysis of varying model behavior would test this hypothesis as well as provide valuable understanding and guidance for calibration. Such information is particularly useful for watershed models that are used extensively in operational environments. These models often require calibration across a wide range of watersheds and are used in multiple applications. As an example, the Sacramento Soil Moisture Accounting Model (SAC-SMA) is the primary model used by the National Weather Service (NWS) River Forecast Centers (RFC) throughout the United States. Several studies have presented automatic or semiautomatic methods that could facilitate calibrating this model for the hundreds of watersheds across the United States, as is required for a countrywide forecasting system [*Brazil*, 1988; *Sorooshian et al.*, 1993; *Duan et al.*, 1994; *Boyle et al.*, 2000, 2001; *Vrugt et al.*, 2003a; *Hogue et al.*, 2006; *Tang et al.*, 2006; *Vrugt et al.*, 2006]. Other studies have focused on the development and assessment of a priori parameter estimates for the SAC-SMA [*Koren*, 2000; *Duan et al.*, 2001; *Koren et al.*, 2003; *Anderson et al.*, 2006; *Gan and Burges*, 2006]. However, to date, few sensitivity analyses of the SAC-SMA exist in the literature despite its common use in operations and hydrologic research. Furthermore, our previous work [*Tang et al.*, 2007b, 2007c] has suggested that some common assumptions about parametric sensitivities of the SAC-SMA model structure are not valid. These findings have implications for methodologies based on a priori assumptions about parameter sensitivity, such as ''stepwise'' or ''stepped'' calibration approaches [e.g., *Hogue et al.*, 2000, 2006; *Fenicia et al.*, 2007] and warrant further investigation. Finally, as efforts continue to reformulate this model into a distributed configuration [*Koren et al.*, 2004], a complete understanding of its parameter sensitivities becomes increasingly vital.

[5] The main objective of this study is to use Sobol's variance-based global sensitivity analysis to build a comprehensive picture of parametric sensitivity for the SAC-SMA and understand its variation across hydroclimatic regimes, flow types, time periods, and timescales. Further objectives are to determine how the variation informs us about model behavior and to what extent the variation is related to the hydroclimatic characteristics of the watersheds and/or simulated time periods. The analysis is intended to demonstrate (1) if moderate model complexity is warranted (or proven excessive) when modeling a range of watersheds, (2) comprehensive trends in SAC-SMA model be-

Table 1. MOPEX Watershed Characteristics

ID	River	Outlet Location	Area, km^2	Mean Annual, mm	Mean Annual ROC (Q/P)	Mean Annual PE, mm
GUA	Guadalupe	Spring Branch, TX	3406	765	0.15	1528
SAN	San Marcos	Luling, TX	2170	827	0.22	1449
ENG	English	Kalona, IA	1484	893	0.30	994
SPR	Spring	Waco, MO	3015	1076	0.28	1094
RAP	Rappahannock	Fredericksburg, VA	4134	1030	0.37	920
MON	Monocacy	Frederick, MD	2116	1041	0.40	896
EAS	East Fork White	Columbus, IN	4421	1015	0.37	855
POT	S. Branch Potomac	Springfield, WV	3810	1042	0.33	761
BLU	Bluestone	Pipestem, WV	1021	1018	0.41	741
AMI	Amite	Denham Springs, LA	3315	1564	0.39	1073
TYG	Tygart Valley	Phillipi, WV	2372	1166	0.63	711
FRE	French Broad	Ashville, NC	2448	1383	0.58	819

havior for calibration guidance, and (3) the validity of typical parameter sensitivity and identifiability assumptions.

2. MOPEX Basins and Data

[6] The hydrometeorological data sets used in this study were developed as part of the Model Parameter Estimation Experiment (MOPEX) and include data for 12 watersheds in the United States that span different hydroclimatic regimes and geographic locations [*Duan et al.*, 2006]. Previous studies [*Duan et al.*, 2006; *Gan and Burges*, 2006] summarized the performance of the SAC-SMA in these watersheds.

[7] From the MOPEX data set, daily precipitation and daily streamflow for 39 years (1960–1998) of data were used, along with mean monthly estimates of potential evaporation and vegetation adjustments. The relative locations of the 12 watersheds are shown in Figure 1 and their characteristics are listed in Table 1. In this table and in subsequent figures, the watersheds are ordered from dry to wet based on the wetness index, which is defined as the ratio of mean annual precipitation (P) to mean annual potential evaporation (PE). Throughout this paper, the watersheds will be identified using the three-letter IDs listed in Table 1.

[8] As illustrated in Figure 1, the watersheds' drainage areas range from the smallest (BLU) case encompassing 1021 km^2 to the largest (EAS) watershed draining 4421 km^2. They are located in the general southeastern region of the United States and include a variety of topographic and land cover characteristics. The wide ranges of mean annual P (765–1564 mm/a), mean annual runoff coefficient (ROC) (0.15–0.63), and mean annual PE (711–1528 mm/a) exemplify the diverse hydroclimatic regimes represented in the data set. A further summary of watershed characteristics is presented in Figures 2a–2e. Additional physical characteristics of these watersheds are presented by *Duan et al.* [2006] and *Gan and Burges* [2006].

3. Methods

3.1. Sacramento Soil Moisture Accounting Model (SAC-SMA)

[9] The SAC-SMA is a conceptual rainfall-runoff model that represents the soil column by an upper and lower zone of multiple storages [*Burnash*, 1995]. It has been used extensively in both research and operational applications where it is the primary rainfall-runoff model used for river forecasting by the National Weather Service (NWS) River Forecast Centers (RFCs) across the United States. Figure 3 shows the structure of the SAC-SMA and the main function of its 16 model parameters (shown in bold). Beyond these main functions, several parameters have secondary functions as part of the percolation component, which connects upper and lower zones. The representation of the percolation process is somewhat different in the SAC-SMA than in some other commonly used watershed models (e.g., PRMS, VIC, TOPMODEL). In the SAC-SMA, percolation is a function of both the upper zone moisture availability and the lower zone moisture deficit (versus only moisture availability as in many other models). Therefore parameters that control the moisture content of both the upper and lower zones also impact the amount of percolation.

[10] In the parameter estimation process, two of the 16 SAC-SMA model parameters are typically set to standard values (SIDE and RSERV) for all watersheds. The remaining 14 parameters must be estimated by some means (calibration or otherwise) for each watershed. These 14 parameters were the focus of this study and are described in Table 2 along with the allowable ranges used in the sensitivity analyses [*Anderson*, 2002]. Our objective was to investigate the parameter sensitivities within the ranges defined as reasonable by the NWS for standard SAC-SMA model calibration over the variety of watershed types found in the United States.

3.2. Sobol Sensitivity Analysis

[11] *Sobol*'s [1993] sensitivity analysis method is a variance-based approach in which the model output variance is decomposed into relative contributions from individual parameters and parameter interactions. This method was selected based on previous work that demonstrated it to be more robust than other sensitivity analysis methods for the evaluation of hydrologic models [*Tang et al.*, 2007b]. In addition, Sobol's method explicitly includes the effects of parameter interactions and quantifies sensitivity with easily compared indices, a necessity for our analysis. The method's primary drawback is its relatively large computational requirements.

[12] In Sobol's method, sensitivity to each parameter or parameter interaction is assessed based on its percent contribution to the total output variance. The variance in model output is typically measured as the variance in a model evaluation metric such as the root mean square error

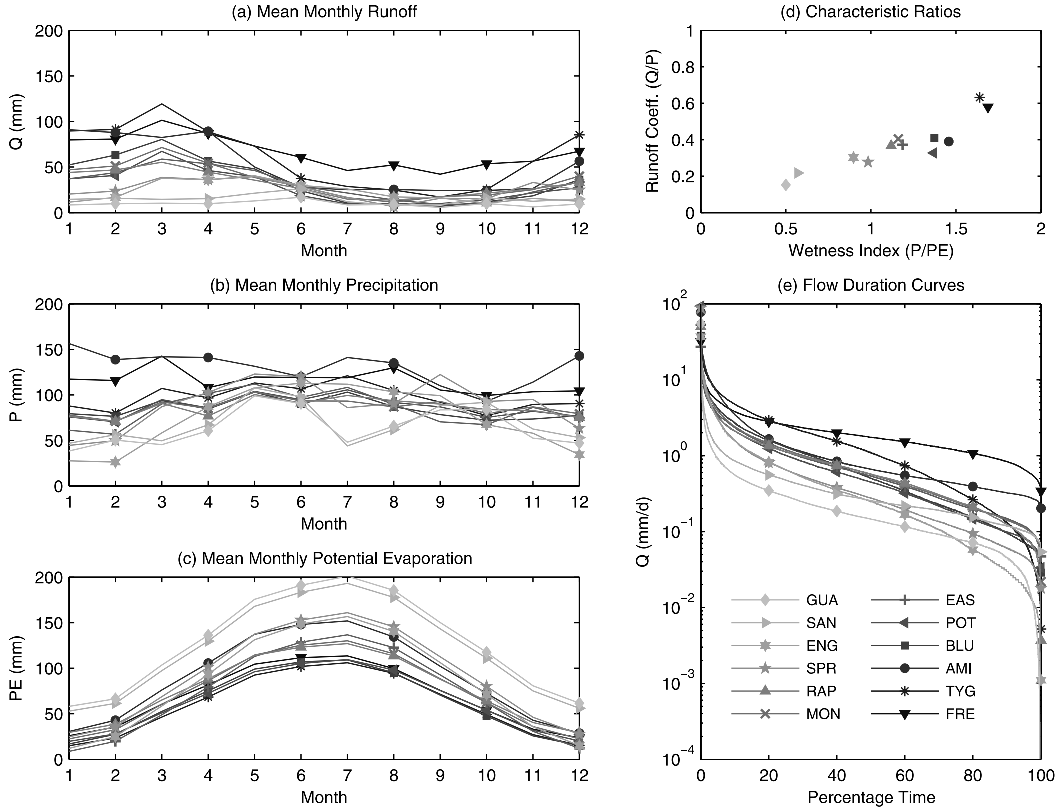

Figure 2. MOPEX watershed characteristics: (a) mean monthly runoff volume in millimeters, (b) mean monthly precipitation, (c) mean monthly potential evaporation, (d) hydrologic ratios, and (e) flow duration curves.

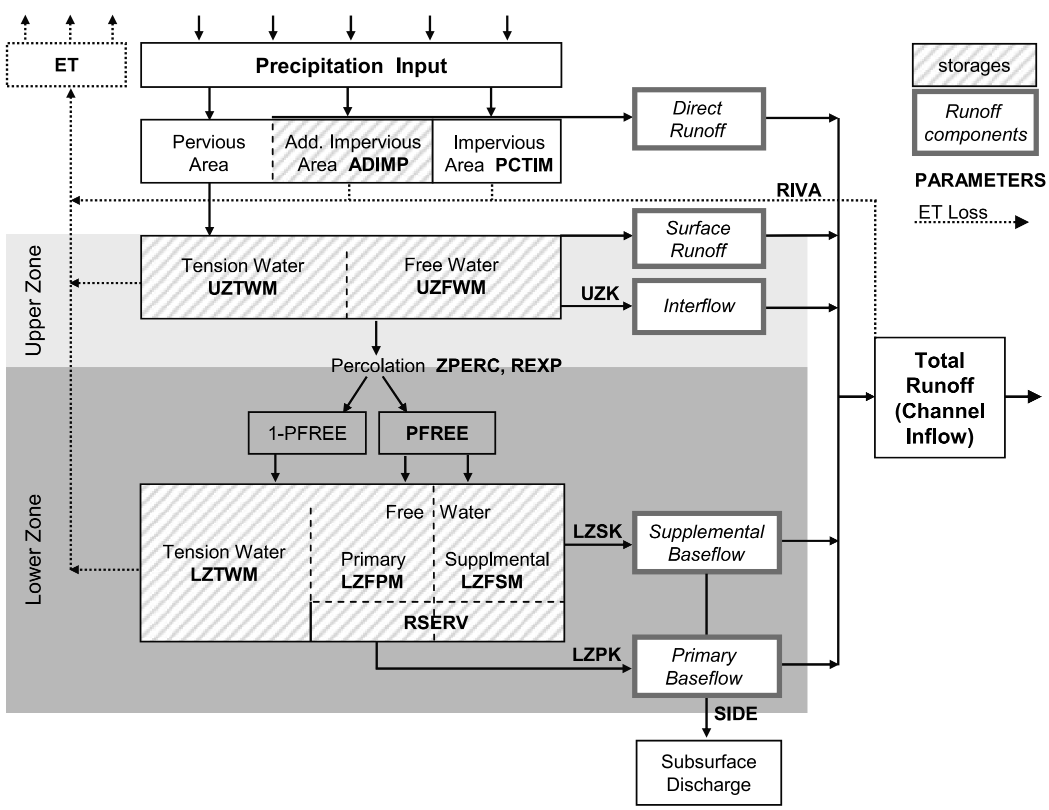

Figure 3. Conceptualization of the SAC-SMA.

Table 2. Description of SAC-SMA Parameters and Ranges Analyzed in This Study

Parameter	Units	Description	Allowable Range
UZTWM	mm	Upper zone tension water maximum storage	25–125
UZFWM	mm	Upper zone free water maximum storage	10–75
UZK	d^{-1}	Upper zone free water withdrawal rate	0.2–0.5
PCTIM	%/100	Percent permanent impervious area	0.0–0.05
ADIMP	%/100	Percent area contributing as impervious when saturated	0.0–0.2
RIVA	%/100	Percent area affected by riparian vegetation	0.0–0.2
ZPERC	dimensionless	Maximum percolation rate under dry conditions	20–300
REXP	dimensionless	Percolation equation exponent	1.4–3.5
PFREE	%/100	Percent of percolation going directly to lower zone free water	0–0.5
LZTWM	mm	Lower zone tension water maximum storage	75–300
LZFPM	mm	Lower zone free water primary maximum storage	40–600
LZFSM	mm	Lower zone free water supplementary maximum storage	15–300
LZPK	d^{-1}	Lower zone primary withdrawal rate	0.001–0.015
LZSK	d^{-1}	Lower zone supplementary withdrawal rate	0.03–0.2

(RMSE). Throughout this section, references to the variance in model output should be interpreted as the variance in an evaluation metric. The four metrics used in this study will be described in section 3.3. Sobol's variance decomposition can be represented as:

$$D(f) = \sum_i D_i + \sum_{i<j} D_{ij} + \sum_{i<j<k} D_{ijk} + D_{12\ldots p} \tag{1}$$

where $\boldsymbol{f}$ is the distribution of model output, $D(\boldsymbol{f})$ is the total output variance; D_i is the output variance due to the ith component of the input parameter vector Θ; D_{ij} is the output variance due to the interaction of parameter θ_i and θ_j; D_{ijk} represents third-order interactions; $D_{12\ldots p}$ represents all interactions greater than third-order; and p defines the total number of parameters. In this study, we were primarily interested in each parameter's total contribution to output variance, as well as how much of that contribution was due to individual effects versus interactions with other parameters (i.e., the difference between the total and individual effects). The first-order and total Sobol sensitivity indices were calculated to measure these contributions and are defined as:

$$\text{firstorder index}: \quad S_i = \frac{D_i}{D} \tag{2}$$

$$\text{total index}: \quad S_{Ti} = 1 - \frac{D_{\sim i}}{D} \tag{3}$$

where the first-order index, S_i, measures the model sensitivity to the individual effect of parameter θ_i, and the total index, S_{Ti}, measures the sensitivity due to the combined effect of parameter θ_i plus its interactions with all other parameters in the analysis. In equation (3), the term $D_{\sim i}$ refers to the variance resulting from all of the parameters *except* θ_i. In other words, if parameter θ_i were removed from the analysis, the resulting reduction in output variance is equivalent to the total impact of parameter θ_i. Since the indices are ratios of a portion to the total output variance, their values range from 0 to 1 and can be directly compared. If a particular parameter has a small first-order index but a large total sensitivity index, then that parameter impacts the model primarily through parameter interactions.

[13] The variance terms (i.e., D terms) in equations (1)–(3) can be approximated by numerical integration in a Monte Carlo framework. Distributions of model parameters are sampled and evaluated to generate distributions of model output. The total output variance, D, is simply the statistical variance of the output distribution, as follows:

$$\hat{f}_o = \frac{1}{n}\sum_{s=1}^{n} f(\Theta_s) \tag{4}$$

$$\hat{D} = \frac{1}{n}\sum_{s=1}^{n} f^2(\Theta_s) - \hat{f}_o^2 \tag{5}$$

where f is the model output, f_o is the mean model output, n is the sample size, and Θ_s is the sampled parameter vector. Calculation of the variance contributions is somewhat more complicated. An important aspect of Sobol's method is the use of two different samples, generated by the same scheme and with the same number of elements. The model is evaluated using the first sample to calculate the overall output mean and variance (i.e., the combined effects of all parameters). The second sample is then used to resample each parameter, rather than setting each to a fixed value, for the calculation of total and individual variance contributions. For the latter calculations, parameter vectors are constructed systematically, with values selected from the two samples in specific combinations defined by which parameter's contribution is being calculated. The resulting distributions of the parameter vectors are evaluated to obtain the corresponding distributions of model output that are used in the approximations for D_i and $D_{\sim i}$. The expressions for D_i and $D_{\sim i}$ as defined by *Sobol* [1993, 2001], *Hall et al.* [2005], and *Saltelli* [2002] are

$$\hat{D}_i = \frac{1}{n}\sum_{s=1}^{n} f\left(\Theta_s^{(a)}\right) f\left(\Theta_{(\sim i)s}^{(b)}, \Theta_{is}^{(a)}\right) - \hat{f}_o^2 \tag{6}$$

$$\hat{D}_{\sim i} = \frac{1}{n}\sum_{s=1}^{n} f\left(\Theta_s^{(a)}\right) f\left(\Theta_{(\sim i)s}^{(a)}, \Theta_{is}^{(b)}\right) - \hat{f}_o^2 \tag{7}$$

where (a) and (b) are two different samples (both of size n). The Θ symbols, defined in Table 3, indicate from which samples the parameters values are taken. In this study Sobol's quasi-random sequence was used to sample points

Table 3. Definition of the Θ Symbols in Equations (4)–(7) of Sobol's Method

Symbol	Definition
Θ_s	Sampled parameter vector
$\Theta_{is}^{(a)}$	Parameter θ_i taken from sample (a)
$\Theta_{is}^{(b)}$	Parameter θ_i taken from sample (b)
$\Theta_{(\sim i)s}^{(a)}$	All parameters except θ_i taken from sample (a)
$\Theta_{(\sim i)s}^{(b)}$	All parameters except θ_i taken from sample (b)

more uniformly in the parameter space than uncorrelated random sampling. Details of this sampling scheme can be found in the work of *Sobol* [1967, 1993], *Brantley and Fox* [1988], and *William et al.* [1999].

3.3. Metrics for Model Evaluation

[14] Applications of watershed models are inherently multiobjective [*Gupta et al.*, 1998; *Madsen*, 2000; *Buras*, 2001; *Vrugt et al.*, 2003b; *Bekele and Nicklow*, 2005; *Tang et al.*, 2007a]. In this study, we used four different model evaluation metrics to assess parameter sensitivity, two of which are common statistical metrics and two that are aggregate measures of overall hydrologic response. Each metric replaces the function f in the equations for Sobol's method defined above. Figure 4 illustrates that the metrics capture four important components of the hydrograph, including high flows, low flows, variability in midrange flows (streamflow regime), and the long-term water balance. The high flow metric is the commonly used root mean square error (RMSE), defined as:

$$RMSE = \sqrt{\frac{1}{m}\sum_{t=1}^{m}\left(Q_{s,t} - Q_{o,t}\right)^2} \tag{8}$$

where m is the number of time steps, $Q_{s,t}$ is the simulated flow for time step t, and $Q_{o,t}$ is the observed flow in time step t. For the low flow metric, the simulated and observed flow time series are first transformed by a Box-Cox transformation (equation (9)) with a λ value of 0.3, which has a similar effect as a log transformation. The RMSE of the transformed flows is then calculated to obtain a metric that emphasizes low flow, referred to here as the transformed root mean square error (TRMSE) (equation (10)).

$$Z = \frac{(1+Q)^{\lambda} - 1}{\lambda} \tag{9}$$

$$TRMSE = \sqrt{\frac{1}{m}\sum_{t=1}^{m}\left(Z_{s,t} - Z_{o,t}\right)^2} \tag{10}$$

where m is again the number of time steps, $Z_{s,t}$ is the transformed simulated flow for time step t, and $Z_{o,t}$ is the transformed observed flow in time step t. The next metric, referred to as slope of the flow duration curve error (SFDCE), measures how well the model captures the distribution of midlevel flows. The slope of a watershed's flow duration curve indicates the variability, or flashiness, of its flow magnitudes. The SFDCE metric is thus simply the absolute error in the slope of the flow duration curve between the 30 and 70 percentile flows as follows:

$$SFDCE = abs\left(\frac{Q_{s,70} - Q_{s,30}}{40} - \frac{Q_{o,70} - Q_{o,30}}{40}\right) \tag{11}$$

where $Q_{s,30}$ and $Q_{s,70}$ are the 30 and 70 percentile flows of simulated flow duration curve and $Q_{o,30}$ and $Q_{o,70}$ are the 30 and 70 percentile flows of observed flow duration curve. Since this metric first combines the flows into one value (in this case slope) before calculating the error, it is an aggregate measure of overall model response and less biased by individual events. Similarly, the final metric, the runoff coefficient error (ROCE) captures the overall accuracy of the water balance by first combining the flows into one characteristic hydrologic descriptor, the mean annual runoff coefficient. The absolute error in the runoff coefficient is then calculated and thus the ROCE is defined as

$$ROCE = abs\left(\frac{\overline{Q}_s}{\overline{P}} - \frac{\overline{Q}_o}{\overline{P}}\right) \tag{12}$$

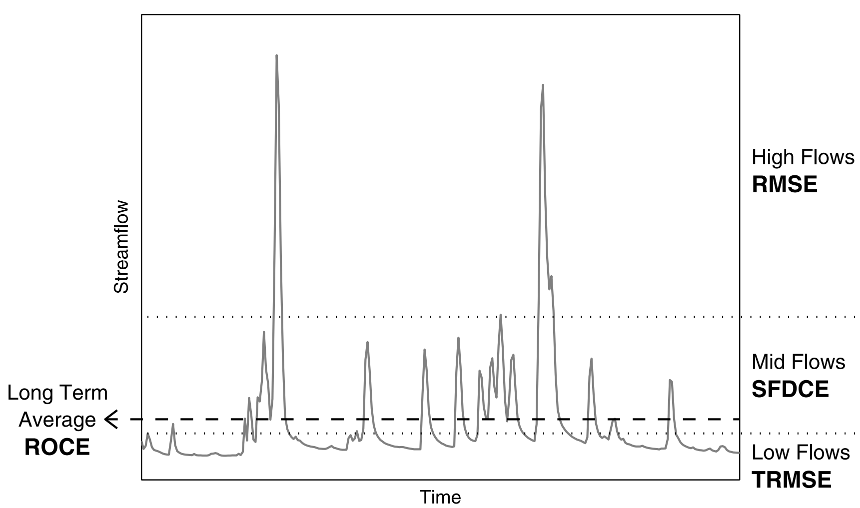

Figure 4. Hydrograph components captured by the four selected evaluation metrics.

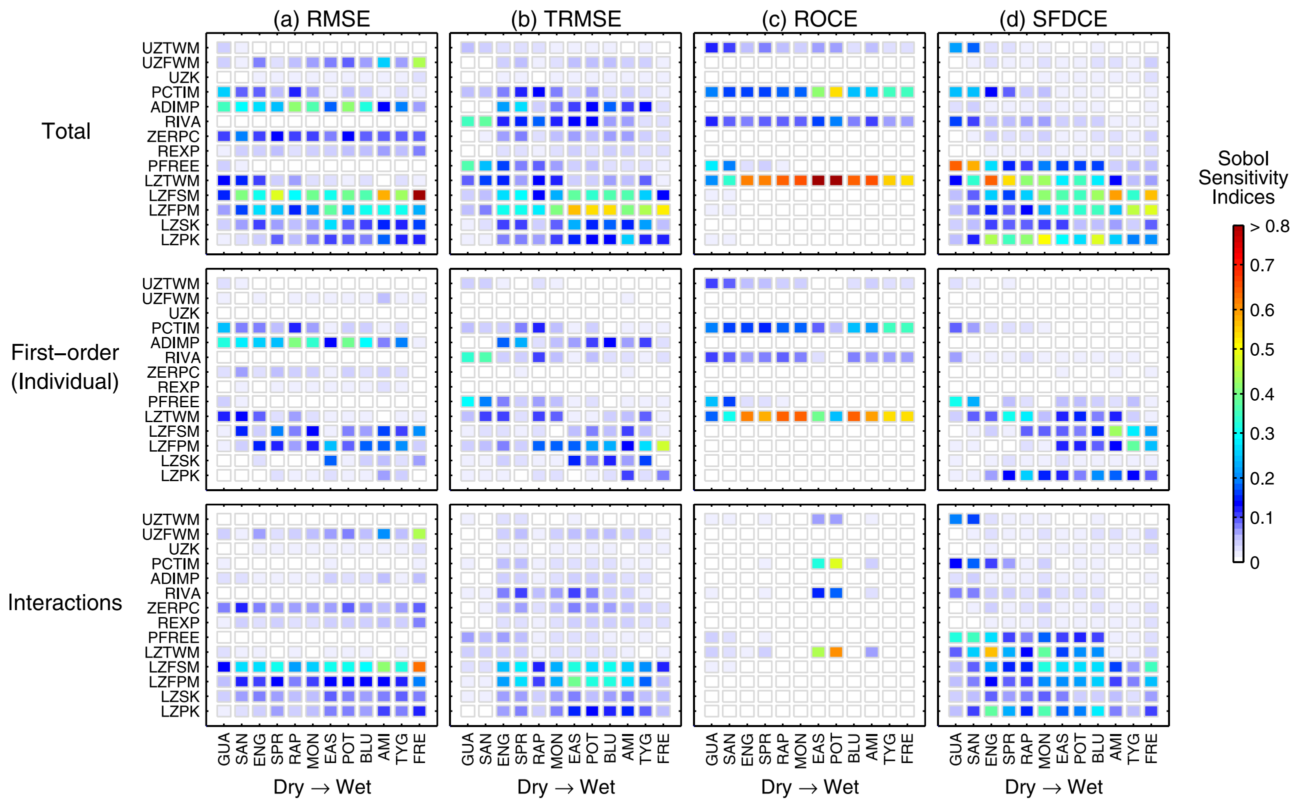

Figure 5. Long-term Sobol sensitivity indices for 12 watersheds and 14 parameters based on a 38-year period: Total (individual plus interactions) sensitivities, first-order (individual contributions) and interaction contributions for (a) RMSE, (b) TRMSE, (c) ROCE, and (d) SFDCE.

where $\overline{Q}_s$ and $\overline{Q}_o$ and are the simulated and observed mean annual runoff volume, and $\overline{P}$ is the mean annual precipitation.

3.4. Spearman Rank Correlation

[15] The Spearman rank correlation coefficient was used in this study to assess the relationships between parameter sensitivities and hydroclimatic characteristics. It provided a means to quantify the strength of a monotonic relationship between two variables, with no assumptions of frequency distribution or linearity [*Lehmann and D'Abrera*, 1998]. Since some of the relationships were highly nonlinear in this study (as will be shown in section 5), the Spearman rank coefficient was preferable to a traditional linear correlation coefficient. To calculate it, values of each of the variables (in this case watershed characteristics and sensitivity indices) are ranked and the correlation is calculated based on the difference in rankings as follows:

$$R = 1 - \frac{6\sum d^2}{v(v^2 - 1)} \tag{13}$$

where d is the difference in rank between the variables for a given value and v is the number of values.

4. Approach

[16] The methods described above were used to perform a comprehensive sensitivity analysis of the SAC-SMA for 12 watersheds using four model evaluation metrics for both a long-term 39 year period as well as yearly periods. A Monte Carlo sampling scheme [*Saltelli*, 2002; *Tang et al.*, 2007b] was used with 8096 samples and a warm-up period of 1 year (i.e., the first year was not included in the sensitivity calculations to allow the model states to warm up and remove any impact of uncertain initial conditions). The method was repeated for the four evaluation metrics described in section 3.3, resulting in 48 separate sets of sensitivity results (a "set" refers to a group of 14 individual and 14 total indices that result for the 14 model parameters in each run). The total-order Sobol indices were compared across watersheds and across the objectives to identify any visible patterns of variation in sensitivity. To quantify the variation, relationships between parameter sensitivity and several hydroclimatic characteristics were developed as scatter plots and correlation was calculated by the Spearman rank correlation method. Results of the long-term sensitivity and correlation analysis are presented in section 5.1.

[17] Beyond the long-term analysis, an interannual analysis was performed to investigate the year-to-year variation in sensitivity within each watershed. Sobol's method was applied using the same sampling scheme described above. For each sample, model simulations were again generated for a 39-year period, however, evaluation metrics and Sobol's indices were calculated separately for each individual calendar year. Therefore the results of each year are based on the same parameter samples and can be directly compared. As in the long-term analysis, the first year was used as a warm-up period and not included (leaving 38 separate years for analysis). Annual sensitivity indices were generated using this method for the 12 watersheds, 38 individual years, and four evaluation metrics. The results were plotted to identify patterns of temporal variation within

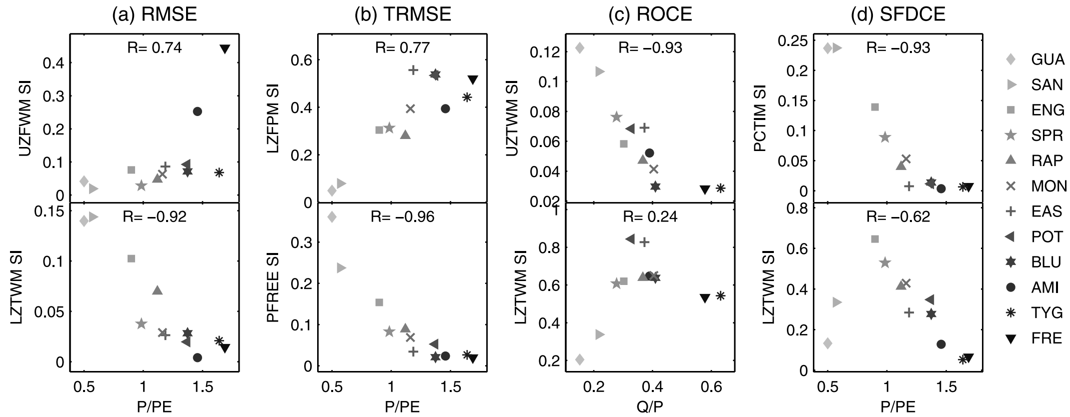

Figure 6. Scatterplots demonstrating Spearman rank correlation (R) between selected parameters' long-term Sobol sensitivity indices (SI) and long-term hydroclimatic variables across watersheds for (a) RMSE, (b) TRMSE, (c) ROCE, and (d) SFDCE. Watersheds are indicated by the marker symbols and the selected parameter is labeled on the y-axis.

watersheds, as well as to identify trends across watersheds that were masked in the long-term analysis. Relationships between interannual hydroclimatic characteristics and the interannual sensitivity indices were plotted and quantified by Spearman rank correlation. Results of the interannual sensitivity and correlation analysis are presented in section 5.2. As a final step, the results from both the long-term and interannual sensitivity analyses are synthesized to provide some general overall guidance for SAC-SMA model identification in section 5.3.

5. Results and Discussion

5.1. Long-Term Sensitivities

[18] In the following sections, the results are organized by the different aspects of the flow that have been assessed through the four evaluation metrics. In each section, dominant patterns of sensitivity are highlighted first, followed by discernible trends across the watersheds. Grids of long-term Sobol sensitivity indices across the 12 watersheds and 14 model parameters are presented in Figures 5a–5d. Note that in each grid, watersheds (x-axis) are ordered from dry to wet (left to right) based on the wetness index as in previous figures and tables. The model parameters (y-axis) are generally structured with upper zone parameters at the top, percolation parameters in the middle, and lower zone parameters at the bottom. Total indices (individual plus interactions), first-order indices (individual), and indices representing all parameter interactions (total index minus first-order index) are displayed separately to demonstrate the varying impact of parameter interactions.

5.1.1. High Flows: RMSE

[19] Some dominant patterns are observable in the total indices for the high flow metric (RMSE) (top grid of Figure 5a), where the amount of variable contributing area (ADIMP), the percolation multiplier (ZPERC), and the sizes of the lower zone free storages (LZFPM, LZFSM) are consistently sensitive across the watersheds. The sensitivity of variable contributing area (ADIMP) reflects that parameter's impact on high peaks (mainly at the high end of its allowable range). The strong sensitivity to lower zone parameters for this metric, however, is initially surprising and counter to typical a priori assumptions that mainly upper zone parameters dominate high flow simulations. However, in the SAC-SMA, the lower zone free water storages and recession rates are directly involved in the calculation of percolation. Since percolation controls the partitioning of water between the upper and lower zones, it also impacts how much of a given event is generated by faster, higher-peaking runoff components from the upper zone (i.e., interflow, surface runoff, or direct runoff) versus slower, lower-peaking components from the lower zone (base flow). In this capacity, parameters that control percolation (including ZPERC and lower zone storages and recessions) impact high-flow simulations as reflected by the sensitivity indices. The importance of interactions among these parameters is illustrated in Figure 5a (middle and lower grids). In some cases (e.g., lower zone recessions, LZPK and LZSK), a parameter's total sensitivity comes almost entirely from the effects of interactions. Overall the contributions by parameter interactions are a significant part of the total sensitivity picture for RMSE.

[20] Comparing parameter sensitivities across the watersheds for RMSE (i.e., across rows of Figure 5a), several trends are visible that provide insight into model behavior across hydroclimatic regimes. For example, the strong trend of increasing sensitivity from dry to wet watersheds for the upper zone free storage (UZFWM) and the opposite (though not as strong) trend for percent impervious area (PCTIM) demonstrates a shift in mechanisms for generating peaks. These trends suggest that in wet watersheds, simulated peaks are more often generated by saturation of the upper zone free water storage, while in dry watersheds peaks are largely controlled by direct runoff from impervious areas. This observation is intuitive as dry watersheds rarely receive enough rainfall to saturate both the tension storages and upper zone free water (which is required to generate surface runoff). Thus in many events impervious area will be the dominant or even the only mechanism of producing runoff in the model. Conversely, wet watersheds regularly saturate and produce surface runoff (which "overshadows" impervious runoff in the RMSE measure). Figure 6a (top) illustrates the trend of increasing upper zone free storage sensitivity with watershed wetness index. The Spearman

rank correlation coefficient for this relationship is 0.74. Another informative trend is the higher sensitivities of the tension storages in dry watersheds. Figure 6a (bottom) shows the strong negative rank correlation (R = −0.92) between lower zone tension storage (LZTWM) sensitivity and watershed wetness coefficient. This indicates the greater importance or "activation" of thresholds in dry watersheds for simulating peaks. In these watersheds, longer and more frequent dry periods (and less overall volume of precipitation) lead to tension storages drying out more often. Therefore the sizes of the tension storages become important in determining if and when thresholds are crossed and when runoff is generated. On the other hand, in wet watersheds the tension storages may regularly be full (for nearly any size) and thus not have a significant impact on peak simulations.

5.1.2. Low Flows: TRMSE

[21] Moving to Figure 5b and the low-flow (TRMSE) evaluation metric, the pattern of sensitivity is similar to that of the high-flow (RMSE) metric (e.g., prominent lower zone sensitivity) though some differences exist. For example, the switch in the parameter with highest sensitivities from the secondary to primary base flow storage (i.e., from LZFSM to LZFPM) reflects the greater importance of slower-receding (primary) base flow for low-flow versus high-flow periods. In addition, the reduced sensitivity of the percolation curve multiplier (ZPERC) suggests that the percolation in dry conditions (which is when ZPERC controls percolation) is less important than it was for high flows. This result is expected since much of percolation in dry conditions goes to lower-zone tension storage and does not recharge base flow. The lower-zone sensitivities in TRMSE are likely due to both the main parameter functions (i.e., control of the potential volume of base flow and slope of recession) as well as its role in the percolation during saturated soil conditions. Another difference between the high-flow (RMSE) and low-flow (TRMSE) results is the reduction in upper zone free storage (UZFWM) sensitivity. This difference makes sense since this parameter primarily impacts interflow and surface runoff generation (high-flow components) rather than base flow. The emergence of some sensitivity in parameters that control evapotranspiration (ET) losses (RIVA, LZTWM, PFREE) represents additional overall shifts in model control, as ET losses have a larger impact on low flows than on high flows. Finally, in Figure 5b, the contribution of interactions to total parameter sensitivity is again apparent (as for RMSE), further supporting the importance of accounting for parametric interactions.

[22] Comparing sensitivities across watersheds for TRMSE (i.e., across rows of Figure 5b), we find that the overall variability is somewhat larger than it was for RMSE. There are fewer parameters in this case that are sensitive across all watersheds. Most noticeably, the two driest watersheds (SAN, GUA) have distinctly different patterns of sensitivity than the other (particularly the wettest) watersheds. The lower sensitivity for the lower-zone free primary storage (LZFPM) in the dry watersheds (Figure 6b, top) reflects the limited importance of base flow for low flow simulations in these watersheds (where base flow may be intermittent). In contrast, the greater impact of ET on low flows in dry watersheds stands out strongly in the higher lower zone tension storage (LZTWM) sensitivities, riparian vegetation area (RIVA) sensitivities, and lower-zone partitioning (PFREE) sensitivities. The latter trend is shown and quantified in Figure 6b (bottom) with Spearman rank correlations of −0.96. These parameters' impacts on ET loss are discussed further in the next section.

5.1.3. Water Balance: ROCE

[23] The overall parameter sensitivity pattern for the long-term water balance metric (ROCE) is distinctly different than that of the other metrics (Figure 5c). Rather than being dominated by the lower-zone parameters, the pattern for ROCE is controlled across all watersheds by parameters that affect the volume of ET losses (UZTWM, PCTIM, RIVA, LZTWM, PFREE). This result reflects the fact that these parameters largely control the volume (rather than the shape) of the hydrograph, which impacts the long-term water balance. In the SAC-SMA, ET losses occur primarily from the upper and lower zone tension storages and from riparian areas. The amount of loss from each store depends on the demand (potential ET for that time of year) and on the supply (water content of the storage). The parameters that are sensitive to the long-term water balance are those that affect not only the size of these storages (i.e., the potential volume of losses) but also the amount of water that goes into these storages. For example, the percent impervious area controls the volume of runoff that enters the channel directly and is therefore unavailable to ET (i.e., it is the volume that does not enter the upper zone tension storage). Similarly, in the lower zone, percolated water is partitioned between tension and free storages by parameter PFREE. The volume that goes directly into free storage (rather than tension storage) is effectively unavailable for ET loss. These two parameters thus may be sensitive in addition to the parameters controlling the size of the ET source zones (UZTWM, LZTWM, RIVA).

[24] Evaluating the effects of interactions in Figure 5c (lower row), it is clear that interactions are not significant for ROCE, in contrast to RMSE and TRMSE. The individual sensitivity pattern (middle) is nearly identical to the total sensitivity pattern (top) and the contributions from interactions (lower) are largely zero, with a few exceptions. This observation reflects that these parameters are sensitive due to their main, independent functions in the model (rather than due to any interacting process like percolation). The reasons for interaction sensitivities for a few watersheds (e.g., POT and EAS) could not be determined though these interactions may be a result of hydroclimatic characteristics not included in this analysis (e.g., precipitation distribution) or of errors in the data.

[25] Comparing results across watersheds for ROCE, large sensitivity differences are apparent between the two driest watersheds (SAN, GUA) and the rest of the watersheds. Here the upper zone tension storage (UZTWM) and the lower zone partitioning (PFREE) become sensitive. These trends likely relate to the frequency or infrequency of saturation of the upper and lower zone tension storages. If the lower zone is usually dry, the percolation partitioning (PFREE) is more important due to its control over the volume of percolation going into the tension storage (and eventually lost to ET as discussed above) during unsaturated lower zone conditions. If the lower zone is saturated, no percolation goes into tension storage, and the partitioning parameter has no effect. Similarly, if the upper zone tension

water storage is frequently full, the volume of ET loss is less variable than if this storage is often dry. ET losses always occur at the potential rate under saturated conditions, thus if a watershed's upper zone is often saturated, water balance sensitivity to upper zone tension storage will be low (and vice versa). The strong negative rank correlation associated with this trend is presented in Figure 6c (top). The second trend illustrated for ROCE involves the lower zone tension storage sensitivity (Figure 6c, bottom). This trend is different than others in that it shows a nonmonotonic relationship with watershed characteristics since the highest sensitivities occur for the midwetness watersheds. Lower sensitivities occur for both the wettest and driest watersheds, producing the inverse V shape seen in the scatterplot of this parameter in Figure 6c. The reduction in sensitivity for wet watersheds is again likely due to more frequent saturation of the lower-zone tension storage and thus less impact of its size on the volume of losses to ET. Similar to UZTWM, in some years the storage may fill for all parameter values and the volume of ET loss is unaffected by that parameter for those periods. Conversely, in very dry watersheds there are potentially long periods in which the upper-zone tension storage never fills and thus no water percolates to the lower zone (and LZTWM has no effect).

5.1.4. Medium Flow Regime: SFDCE

[26] The final metric, SFDCE, evaluates the error in the slope of the flow duration curve between the 30 and 70 percentile flow magnitudes. It thus captures the distribution (i.e., the variability of flow magnitudes) within the range of midlevel flows. The hydrograph components that fall into the 30–70% range vary by watershed, but will generally include small peaks and high base flows (e.g., just after large storms). The distribution of these flow magnitudes (i.e., relative frequencies or "flashiness") determines the steepness (or mildness) of the FDC slope. Figure 5d shows that for this metric, lower-zone parameters again dominate the sensitivity pattern as they did for RMSE and TRMSE. In contrast to those metrics' results, however, the lower-zone tension storage parameter (LZTWM) and lower-zone partitioning parameter (PFREE) are also sensitive for most of the watersheds. These sensitivities, along with lower-zone storages and recessions, reflect the importance of both percolation and lower-zone partitioning (between tension and free storages) for reproducing the flow regimes of the watersheds. The percolation function, as mentioned, determines how much water infiltrates to slow-responding base flow (less variable flow magnitudes) versus how much moves through the upper zone to become faster-responding (more variable) interflow or surface runoff. Lower-zone partitioning then impacts the amount of percolation that recharges base flow after events (versus enters tension storage and is lost to ET). Watersheds with higher percolation and sustained base flow generally show less variability in flow magnitude (less "flashy" regime) than those with less percolation and therefore more surface runoff and interflow. Figure 5d (bottom) shows that the effects of interactions for the SFDCE metric are again significant as they were for RMSE and TRMSE (where percolation was also important). If only individual effects were considered (middle plot) the parameter sensitivity results would be incomplete due to the large contribution by parameter interactions (bottom plot). The importance of interactions for RMSE, TRMSE and SFDCE, though not for ROCE, supports a hypothesis that parameter interactions in the SAC-SMA are largely a result of the percolation function.

[27] Comparing sensitivities across the watersheds for this metric, it is seen that for the drier watersheds the ET-controlling parameters of the upper zone (UZTWM, PCTIM, and RIVA) again become sensitive for SFDCE as they did for ROCE. In this case, however, the reason for their sensitivity is their impact on the variability in flow magnitudes (rather than their impact on long-term runoff volume as for ROCE). This impact is greater for dry watersheds than wet watersheds for two reasons. First, in dry watersheds the parameters will be more frequently "activated" over the 30–70 percentile range of flows (whereas in wet watersheds the upper zone will be more often full over this range of flows and UZTWM and RIVA will have less impact). Second, since relatively high flows occur less frequently in dry watersheds than in wet watersheds, the 30–70 percentile range will shift downward to include lower flows (relative the that watershed's range of flows). Therefore the small peaks generated by impervious area and the recessions that are affected by riparian and upper zone ET will more likely fall into the 30–70 range. Figure 6d (top) demonstrates the decreasing dry to wet trend ($R = -0.93$) between watershed wetness and percent impervious area sensitivity (PCTIM). Figure 6d (bottom) also shows a bimodal trend for lower zone tension storage (LZTWM), similar to the trend for the ROCE evaluation metric (and likely for similar reasons of saturation frequency/infrequency as discussed in section 5.1.3.).

5.2. Interannual Sensitivities

[28] The interannual sensitivity analysis provides additional information about the variability of parameter sensitivities and model behavior across the watersheds. Temporal patterns (Figures 7–8) were plotted to observe how sensitivities change from year to year and how much interannual variability (or consistency) is present within each watershed. For all metrics and watersheds in Figures 7–8, variability in sensitivity patterns is evident. On the basis of the results in section 5.1, it is reasonable to infer that differences in flow and forcing characteristics from year-to-year could also result in different sensitivity patterns (as seen in Figures 7–8) from year-to-year. The interannual correlation analysis supports this premise as trends were found for all watersheds between the hydroclimatic characteristics and the parametric sensitivity results for each year. Some trends found in the interannual analysis were similar to trends in the long-term analysis (i.e., patterns across wet and dry years within a watershed are similar to long-term patterns across wet and dry watersheds). Other interannual trends, however, did not exist in the long-term analysis since they seem to represent more specific combinations of long- and short-term climatic characteristics at a certain location.

[29] In one bounding case the wet conditions in the FRE watershed show an increasing trend in high flow sensitivity for the upper zone free water storage (UZFWM) with annual maximum flow in watershed FRE (Figure 9a, top). In the long-term analysis, this watershed (FRE) had the highest sensitivity to the upper zone free storage (UZFWM) parameter (see Figure 5a). The high sensitivity is attributed to more frequent saturation of the upper zones. In the

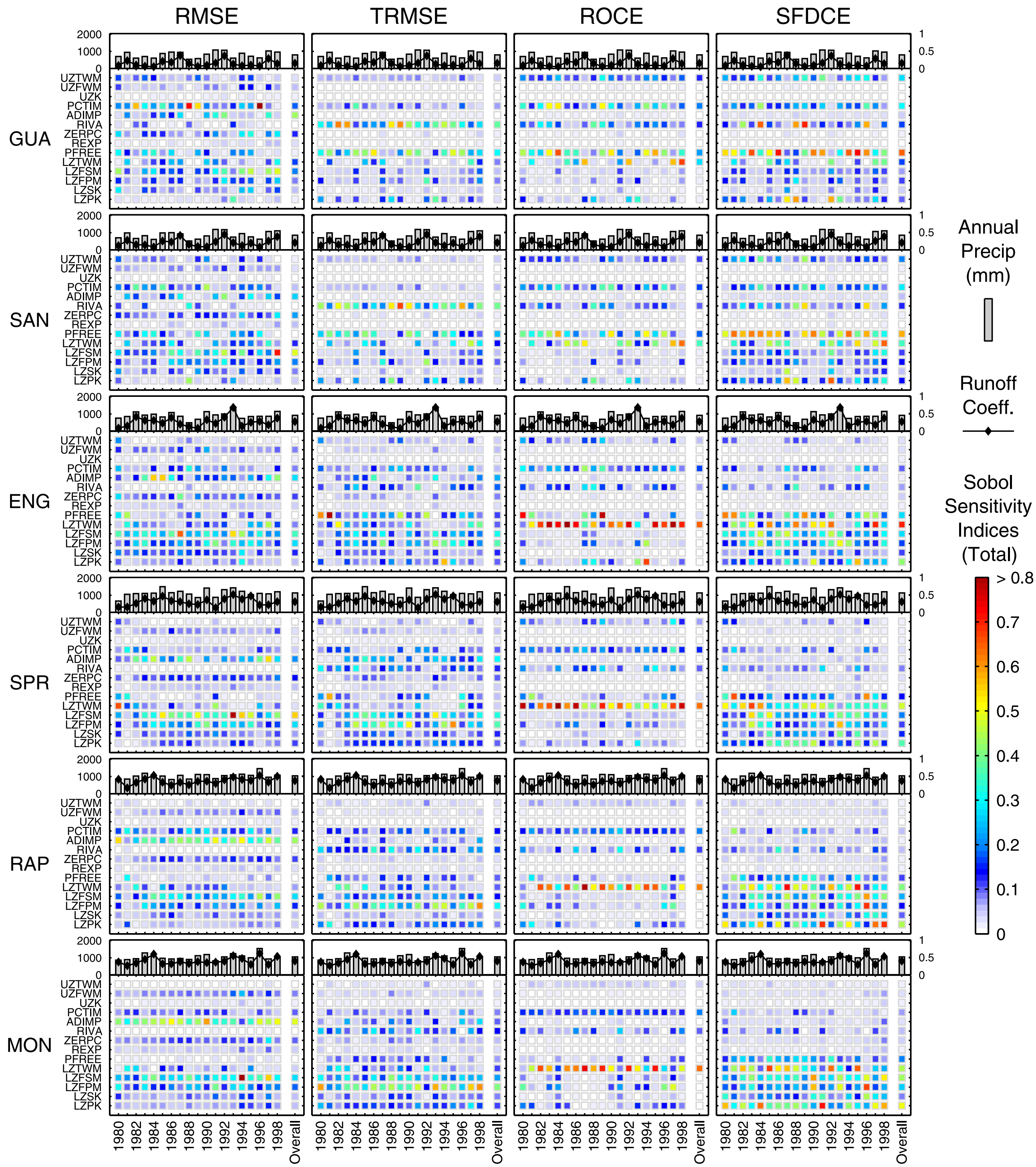

Figure 7. Interannual Sobol sensitivity indices for the six driest watersheds over the period 1980–1998. The overall sensitivity (last column in each grid) is the corresponding result from the long-term analysis shown in Figure 5. Total annual precipitation (left axis) and the runoff coefficient (right axis) for each year is shown above each sensitivity grid.

interannual analysis, years with high annual maximum flow indicate when large event(s) occurred and thus when the upper zone storages of the SAC-SMA are most likely to be saturated. In such cases (as discussed in section 5.1), runoff is produced in the model mainly by surface runoff (i.e., saturation excess). Thus the size of the upper zone free storage has a large impact on simulations particularly in those years, as demonstrated in Figure 9a. Another similar trend between the long-term and interannual analysis is the bimodal behavior of lower-zone tension store sensitivity. In dry watersheds the interannual trend is positive (Figure 9d, top) and in wet watersheds the trend is negative (Figure 9d, bottom). Similar reasons apply as were discussed in section 5.1 (relative frequencies of lower zone saturation).

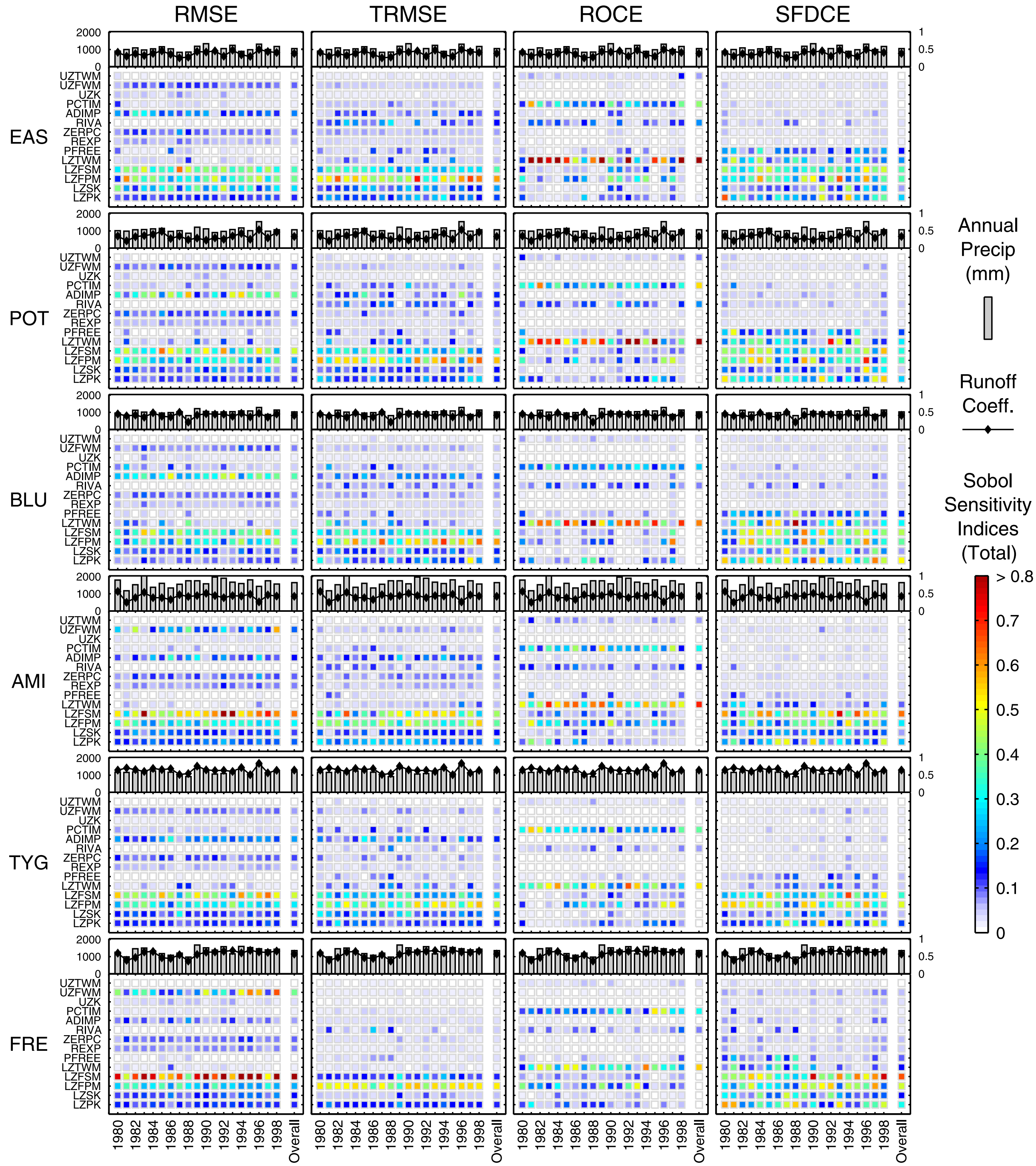

Figure 8. Interannual Sobol sensitivity indices for the six wettest watersheds over the period 1980–1998. The overall sensitivity (last column in each grid) is the corresponding result from the long-term analysis shown in Figure 5. Total annual precipitation (left axis) and the runoff coefficient (right axis) for each year is shown above each sensitivity grid.

[30] An example of a trend not apparent in the long-term analysis is the water balance (ROCE) sensitivity trend seen in watershed GUA for the percent impervious area (Figure 9c, top). The long-term sensitivity for this case (watershed GUA and parameter PCTIM in Figure 5c) was actually lower than most other watersheds. In the interannual analysis, however, percent impervious area becomes more sensitive in the dry years of watershed GUA than it does in nearly all years of other watersheds (Figure 7). This suggests a unique model behavior develops in the driest years of the driest watersheds that is not present for other conditions. In such cases total runoff is likely so low and

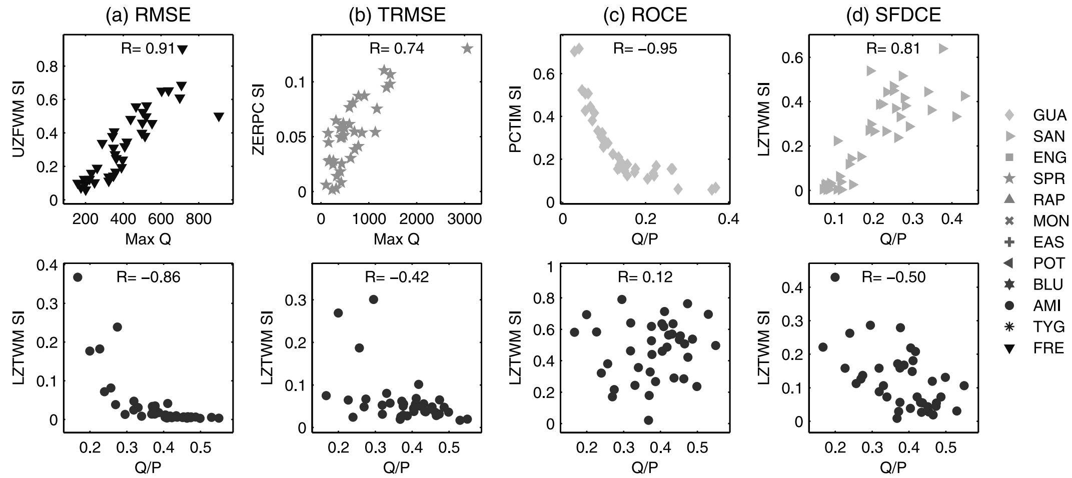

Figure 9. Scatterplots demonstrating Spearman rank correlation (R) between parameters' annual Sobol sensitivity indices (SI) and annual hydroclimatic variables. Each plot displays the result for a particular watershed and parameter combination. The watershed is indicated by the marker and the parameter is labeled on the y-axis.

storages so regularly dry that impervious runoff from the few infrequent events is the only runoff-generating mechanism and thus has a large impact on the water balance for that year. Another difference with long-term results is exemplified by the lower zone partitioning parameter (PFREE) in watersheds SAN and GUA. This parameter is sensitive for high flows in several of the individual years (Figure 7). However, the sensitivity does not appear in the long-term results (last column). As discussed in section 5.1, PFREE controls lower zone partitioning between tension and free water zones and generally affects low to middle flows and the water balance (in dry watersheds). In the dry years for these watersheds, peaks may be relatively low and PFREE will also impact the high flow metric. However, the total variance contribution to long-term results in those years is likely small (relative to years with high peaks) and the influence is therefore not discernible in the long-term result. This observation points out the potential of certain events to dominate long-term results and the ability to extract more information from the data when analyses are performed at shorter timescales. In addition, similarities and differences between the interannual and long-term analyses highlight the influence of timescale on model behavior; a parameter's influences will be most discernible on timescales for which the processes it controls are dominant.

[31] A final point in the interannual results is the tendencies of some watersheds to "look" like other (wet or dry) watersheds in certain years. For example, watershed RAP (see Figure 7) has a low-flow (TRMSE) sensitivity pattern in 1998 that is generally similar (e.g., with a limited sensitivity to PFREE) to the dominant pattern in the wet watersheds. However, its pattern in 1982 resembles the overall pattern of the dry watersheds. This fluctuating sensitivity pattern could be related to the fact that this watershed has a long-term P/PE value near 1 (thus it may fluctuate between energy-limited and water-limited years). As another example, the sensitivity patterns in years when major floods occur cause patterns of drier watershed to resemble patterns of wet watershed. For example, the Mississippi Flood of 1993 is evident in watershed ENG and the Winter Flood of 1996 in the Mid-Atlantic shows up in watersheds MON and POT. This also demonstrates that two watersheds with different locations and hydroclimatic characteristics can potentially have the same sensitivity pattern in a given year (such as ENG and AMI in 1993) under extreme changes in forcing. Overall, Figures 7 and 8 emphasize the risk of assuming a model's sensitivity based on results from a different watershed or analysis time period.

5.3. Synthesis of Results

[32] Sensitivity analysis is often performed in an effort to determine which parameters are most identifiable (i.e., most sensitive) and should be the focus of calibration efforts. The results in the previous sections have demonstrated that parameter sensitivities vary depending on the hydroclimatic characteristics of the watershed and time period of analysis, as well as on the metrics used for the analysis. On the basis of these findings, it becomes clear that to determine parameter sensitivity for a particular watershed and period of record, it is best to perform a complete sensitivity analysis for that specific case. However, computational and time costs may often make this infeasible. Therefore to provide some general guidance for SAC-SMA model identification, the results of the interannual and the long-term analyses are combined to create a summary of expected parameter sensitivity by watershed type and time period (Figure 10). Watersheds are grouped into three categories (dry, middle, and wet watersheds) based on indicated ranges of the wetness index. For each watershed group their behavior for dry, middle, and wet years is classified as highly sensitive, sensitive, or not sensitive based on sensitivity indices of the corresponding watersheds and time periods. Highly sensitive subcategories have a majority sensitivity index greater than 0.1 (e.g., the majority of the indices in the driest years of the dry watersheds are greater than 0.1). Sensitive subcategories have a majority index between 0.01 and 0.1 and nonsensitive subcategories have a majority index less than 0.01. Given a particular watershed and period of record (i.e., how wet/dry is the period),

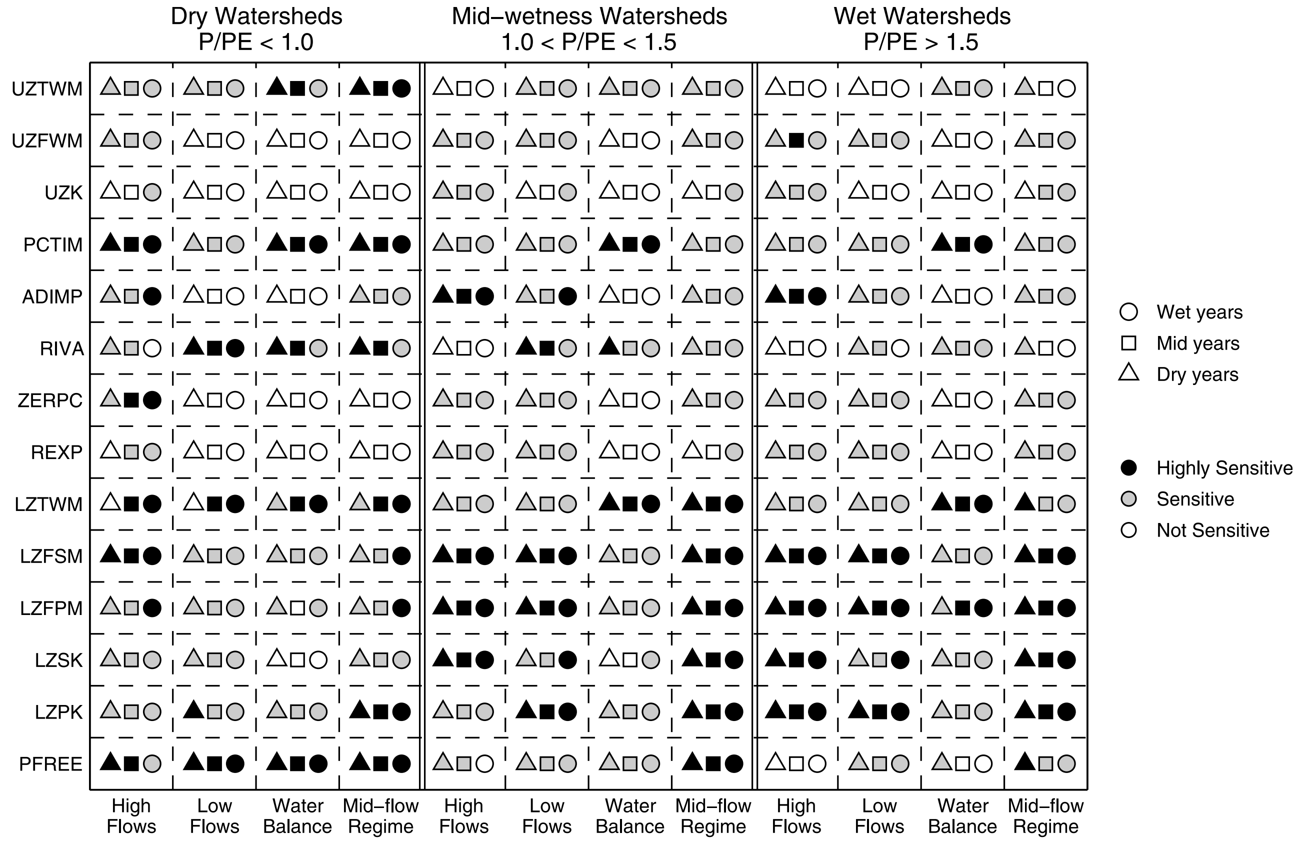

Figure 10. Summary of SAC-SMA parameter sensitivity results and guidance for parameter estimation. Highly sensitive (black) parameters are those where the majority of years have an SI > 0.1, sensitive (gray) are those where the majority are in the range 0.01 < SI < 0.1; and not sensitive (white) are those where the majority of years have an SI < 0.01.

Figure 10 provides some general indication as to which parameters are likely to be the most identifiable, which are less (but still somewhat) identifiable, and which (if any) are largely unidentifiable.

[33] The most evident indication in Figure 10 is the need to focus on the lower zone parameters (LZFSM, LZFPM, LZSK, LZPK) in model identification efforts, particularly for midwetness and wet watersheds. Section 5.1 discussed the involvement of these parameters in percolation and therefore their impact on low, middle, and high flow types. Significant information is also present for these watersheds to identify percent impervious area and lower zone tension storage using the water balance and midflow regime metrics. While high flows contain consistent information for identification of additional impervious area (ADIMP). Information for most of the remaining parameters is present in midwetness and wet watersheds but is weaker and less consistent (i.e., not present in all years). In dry watersheds, model identification focus should shift to parameters impacting ET losses (UZTWM, PCTIM, RIVA, LZTWM, and PFREE) as discussed throughout section 5.1. The remaining parameters (including percolation-related parameters) are less consistently sensitive in dry watersheds and would require targeting specific time periods for identification.

[34] As a final point, it should be noted that similar variation in parameter sensitivity was found at shorter time scales (e.g., intra-annual) than those presented in this paper. We performed the sensitivity analysis on some test cases at a monthly timescale. Sensitivity patterns were still largely driven by hydroclimatic variation as in the longer time scales. These results were not included in this paper as long-term and interannual variation were deemed most relevant for current calibration approaches (i.e., static parameter values are determined based on aggregates measures of model predictions for years to decades).

6. Summary and Conclusions

[35] This study demonstrates that intermediate-complexity watershed models, like the SAC-SMA, include necessary flexibility for representing a wide range of watersheds located in different hydroclimatic regions. An in-depth analysis is presented of the SAC-SMA's parametric sensitivity variation across watersheds composing a hydroclimatic gradient for multiple time periods/timescales and a suite of flow types. The sensitivity patterns demonstrate how different model components become dominant due to changes in forcing (hydroclimatic) conditions. This flexibility, combined with the lack of any consistently insensitive parameter (see Figure 10), substantiate that SAC-SMA's level of complexity is warranted (i.e., the model is not consistently overparameterized) for simulating watersheds across a range of hydroclimatic conditions. Results also provide detailed guidance for SAC-SMA calibration and refute some commonly employed a priori assumptions about the model's parametric sensitivity. The analysis has

broader implications with respect to hydrologic model behavior and identification in general, as discussed below.

[36] Across watersheds, we found that model behavior could be explained largely based on saturation of the upper and/or lower zones. In the upper zone the free water storage is important in watersheds or time periods in which that zone is frequently saturated and therefore producing surface runoff. In contrast, the impact of the upper zone tension storage is greater under conditions of infrequent saturation and therefore variable ET loss. In each case the respective conditions for sensitivity reflect conditions when the storage is actively impacting model predictions and reasonably represents the expected dominant processes. The structure of the upper zone in the SAC-SMA is similar to many hydrologic models in its partitioning between tension (ET-drained) and free water (gravity-drained) storages. Thus this moisture- and ET-driven model behavior of the upper zone would be relevant across many models.

[37] The lower zone structure of the SAC-SMA, however, is somewhat unique (among models of similar complexity) in its coupling with the upper zone through the demand-based percolation function. As a result of the percolation structure, lower-zone model behavior was found to extend beyond its more typical and assumed influences (low flows) to impact middle and high flows across most watersheds. This behavior was most evident in watersheds with appreciable levels of lower-zone saturation (midwetness and wet watersheds) due to the variable percolation demand in such watersheds.

[38] The patterns in model behavior for upper zone, lower zone, and percolation components of the SAC-SMA were interpreted across dry to wet watersheds based on the above-discussed, moisture-related mechanisms. Wet watersheds and dry watersheds resulted in distinctly different patterns of parametric control. The differences were reasonable and intuitive based on differential model forcing, evapotranspiration, and storage across watersheds. Similarly within watersheds, reasonable patterns of sensitivity were produced by wet years and dry years of the analysis record. It follows that similar model behavior variations would be expected to result from nonuniform forcing and storage across a spatially distributed model domain. Wet cells would be expected to have patterns of parametric control similar to wet year/watersheds of this analysis. And likewise dry cells would be expected to follow patterns of dry years/watersheds. Thus implications of this study are likely to apply (and be compounded) for a distributed configuration of the SAC-SMA.

[39] The differences in model behavior demonstrate that a moderate level of complexity (as in the SAC-SMA) is warranted to appropriately represent the hydrology of watersheds across a hydroclimatic gradient. Although some model components were found to be inactive for a single given watershed and/or flow type, a comprehensive evaluation across a range of watersheds and conditions revealed that nearly every model component is important in certain cases. Therefore generalizing model behavior and reducing the number of parameters that require calibration for this model would be difficult if all watersheds and all aspects of the hydrograph should be well represented. These results also demonstrate counter-evidence for the premise that no more than 3–5 parameters can regularly be identified from hydrologic data. We show that when results are combined across four metrics, substantial information exists for 6–10 parameters (highly sensitive parameters), and "some" information exists for most of the remaining parameters. Additionally, the fact that the dominant parameters are often similar across multiple metrics (particularly the commonly used statistical metrics) limits the feasibility of dividing parameters into nonintersecting groups for calibration, which is the basic premise of "step-wise" or "stepped" calibration procedures.

[40] The need to actively couple sensitivity analysis with calibration procedures is clearly indicated by the variation in sensitivity patterns found in this study. Assumptions of parametric controls based on extrapolation of sensitivity analysis results from different watersheds or time periods are likely to be invalid or inapplicable. Furthermore, some common assumptions with respect to the SAC-SMA (e.g., that lower zone parameters do not impact high flows) are incorrect, demonstrating the difficulty of discerning parametric controls a priori without rigorous computational analysis. To most effectively identify important model parameters for calibration of a given watershed and analysis period, sensitivity analysis should be performed for that specific case. An extension of this study is currently in progress to investigate how parameters of varying sensitivity impact the overall model performance. Results of that study will address the feasibility for different watershed types of removing (fixing to constants) parameters from the calibration process.

[41] **Acknowledgments.** Support for the first author was provided by the Henry Luce Foundation in the form of a Clare Booth Luce Fellowship, a Penn State College of Engineering Fellowship, and a GE Faculty for the Future Fellowship. Partial support for the second author was provided by SAHRA under NSF-STC grant EAR-664 9876800, and the National Weather Service Office of Hydrology under grants 665 NOAA/NA04NWS4620012 and NOAA/DG 133W-03-666 SE-0916. The third and fourth authors were partially supported by the National Science Foundation under grants EAR-0310122, EAR-0609791, and EAR-0609741. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the writers and do not necessarily reflect the views of any of the funding institutions.

References

Abbott, M. B., J. C. Bathurst, J. A. Cunge, P. E. O'Connell, and J. Rasmussen (1986), An introduction to the European hydrological system - Systeme Hydrologique Europeen, "SHE", 1. History and philosophy of a physically-based, distributed modeling system, *J. Hydrol.*, *87*(1–2), 45–59.

Anderson, E. A. (2002), Calibration of conceptual hydrologic models for use in river forecasting, *NOAA Tech. Rep. NWS 45*, Hydrol. Lab., Silver Spring, Md.

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

Bastidas, L. A., T. S. Hogue, S. Sorooshian, H. V. Gupta, and W. J. Shuttleworth (2006), Parameter sensitivity analysis for different complexity land surface models using multicriteria methods, *J. Geophys. Res.*, *111*, D20101, doi:10.1029/2005JD006377.

Bekele, E. G., and J. W. Nicklow (2005), Multiobjective management of ecosystem services by integrative watershed modeling and evolutionary algorithms, *Water Resour. Res.*, *41*, W10406, doi:10.1029/2005WR004090.

Beven, K. J. (1989), Changing ideas in hydrology - The case of physically-based models, *J. Hydrol.*, *105*(1–2), 157–172.

Boyle, D. P., H. V. Gupta, and S. Sorooshian (2000), Toward improved calibration of hydrologic models: Combining the strengths of manual and automatic methods, *Water Resour. Res.*, *36*(12), 3663–3674.

Boyle, D. P., H. V. Gupta, S. Sorooshian, V. Koren, Z. Y. Zhang, and M. Smith (2001), Toward improved streamflow forecasts: Value of semi-distributed modeling, *Water Resour. Res.*, *37*(11), 2749–2759.

Brantley, P., and B. L. Fox (1988), Algorithm 659: Implementing Sobol's quasi-random sequence generator, *ACM Trans. Math. Software*, *14*, 88–100.

Brazil, L. E. (1988), Multilevel calibration strategy for complex hydrologic simulation models, Ph.D. thesis, 217 pp., Colo. State University, Fort Collins.

Buras, N. (2001), Water resources - Unresolved issues, *J. Water Resour. Plann. Manage.*, *127*(6), 353.

Burnash, R. J. C. (1995), The NWS River Forecast System - Catchment model, in *Computer Models of Watershed Hydrology*, edited by V. P. Singh, pp. 311–366, Water Resour. Publ., Highlands Ranch, Colo.

Demaria, E. M., B. Njissen, and T. Wagener (2007), Monte Carlo sensitivity analysis of land surface parameters using the Variable Infiltration Capacity model, *J. Geophys. Res.*, *112*, D11113, doi:10.1029/2006JD007534.

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

Duan, Q., J. Schaake, and V. Koren (2001), A Priori estimation of land surface model parameters, in *Land Surface Hydrology, Meteorology, and Climate, Observations and Modeling*, *Water Sci. and Appl. Ser.*, vol. 3, edited by V. Lakshmi, J. Albertson and J. Schaake, pp. 77–94, AGU, Washington, D. C.

Duan, Q., et al. (2006), Model Parameter Estimation Experiment (MOPEX): An overview of science strategy and major results from the second and third workshops, *J. Hydrol.*, *320*(1–2), 3–17.

Fenicia, F., H. H.G. Savenije, P. Matgen, and L. Pfister (2007), A comparison of alternative multiobjective calibration strategies for hydrological modeling, *Water Resour. Res.*, *43*, W03434, doi:10.1029/2006WR005098.

Freer, J., K. J. Beven, and B. Ambroise (1996), Bayesian estimation of uncertainty in runoff prediction and the value of data: An application of the GLUE approach, *Water Resour. Res.*, *32*(7), 2161–2173.

Gan, T. Y., and S. J. Burges (2006), Assessment of soil-based and calibrated parameters of the Sacramento model and parameter transferability, *J. Hydrol.*, *320*(1–2), 117–131.

Gupta, H. V., S. Sorooshian, and P. O. Yapo (1998), Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information, *Water Resour. Res.*, *34*(4), 751–763.

Hall, J. W., S. Tarantola, P. D. Bates, and M. S. Horritt (2005), Distributed sensitivity analysis of flood inundation model calibration, *J. Hydrol. Eng.*, *131*(2), 117–126.

Hogue, T. S., S. Sorooshian, H. V. Gupta, A. Holtz, and D. Braatz (2000), A multistep automatic calibration scheme for river forecasting models, *J. Hydrometeorol.*, *1*, 524–542.

Hogue, T. S., H. V. Gupta, and S. Sorooshian (2006), A 'user-friendly' approach to parameter estimation in hydrologic models, *J. Hydrol.*, *320*(1–2), 202–217.

Hornberger, G. M., and R. C. Spear (1981), An Approach to the Preliminary-Analysis of Environmental Systems, *J. Environ. Manage.*, *12*(1), 7–18.

Jakeman, A. J., and G. M. Hornberger (1993), How much complexity is warranted in a rainfall-runoff model, *Water Resour. Res.*, *29*(8), 2637–2649.

Jakeman, A. J., I. G. Littlewood, and P. G. Whitehead (1990), Computation of the instantaneous unit-hydrograph and identifiable component flows with application to 2 small upland catchments, *J. Hydrol.*, *117*(1–4), 275–300.

Koren, V. (2000), Use of soil property data in the derivation of conceptual rainfall-runoff model parameters, paper presented at 15th Conference on Hydrology, Am. Meteorol. Soc., Long Beach, Calif.

Koren, V., M. Smith, and Q. Duan (2003), Use of a priori parameter estimates in the derivation of spatially consistent parameter sets on rainfall-runoff models, in *Calibration of Watershed Models*, *Water Sci. and Appl. Ser.*, vol. 6, edited by Q. Duan et al., pp. 239–254, AGU, Washington, D. C.

Koren, V., S. Reed, M. Smith, Z. Zhang, and D. J. Seo (2004), Hydrology Laboratory Research Modeling System (HL-RMS) of the U. S. National Weather Service, *J. Hydrol.*, *291*(3–4), 297–318.

Lehmann, E. L., and H. J. M. D'Abrera (1998), *Nonparametrics: Statistical Methods Based on Ranks*, Prentice-Hall, Englewood Cliffs, N. J.

Madsen, H. (2000), Automatic calibration of a conceptual rainfall-runoff model using multiple objectives, *J. Hydrol.*, *235*(3–4), 276–288.

Muleta, M. K., and J. W. Nicklow (2005), Sensitivity and uncertainty analysis coupled with automatic calibration for a distributed watershed model, *J. Hydrol.*, *306*(1–4), 127–145.

Pappenberger, F., I. Iorgulescu, and K. J. Beven (2006), Sensitivity analysis based on regional splits and regression trees (SARS-RT), *Environ. Modell. Software*, *21*(7), 976–990.

Pappenberger, F., K. J. Beven, M. Ratto, and P. Matgen (2008), Multi-method global sensitivity analysis of flood inundation models, *Adv. Water Resour.*, *31*(1), doi:10.1016/j.advwatres.2007.04.09.

Reggiani, P., M. Sivapalan, and S. M. Hassanizadeh (2000), Conservation equations governing hillslope responses: Exploring the physical basis of water balance, *Water Resour. Res.*, *36*(7), 1845–1863.

Saltelli, A. (2002), Making best use of model evaluations to compute sensitivity indices, *Comput. Phys. Commun.*, *145*(2), 280–297.

Saltelli, A., S. Tarantola, and K. Chan (1999), A quantitative model-independent method for global sensitivity analysis of model output, *Technometrics*, *41*, 39–56.

Sieber, A., and S. Uhlenbrook (2005), Sensitivity analyses of a distributed catchment model to verify the model structure, *J. Hydrol.*, *310*(1–4), 216–235.

Singh, V. P., and D. K. Frevert (Ed.) (2006), *Watershed Models*, CRC Press, Boca Raton, Fla.

Sobol, I. M. (1967), On the distribution of points in a cube and the approximate evaluation of integrals, *USSR Comput. Math. Math. Phys.*, *7*, 86–112.

Sobol, I. M. (1993), Sensitivity estimates for nonlinear mathematical models, *Math. Model. Comput. Exp.*, *1*, 407–417.

Sobol, I. M. (2001), Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates, *Math. Comput. Simul.*, *55*, 271–280.

Sorooshian, S., Q. Y. Duan, and V. K. Gupta (1993), Calibration of rainfall-runoff models - Application of global optimization to the Sacramento soil-moisture accounting model, *Water Resour. Res.*, *29*(4), 1185–1194.

Tang, Y., P. M. Reed, and T. Wagener (2006), How effective and efficient are multiobjective evolutionary algorithms at hydrologic model calibration?, *Hydrol. Earth Syst. Sci.*, *10*(2), 289–307.

Tang, Y., P. M. Reed, and J. B. Kollat (2007a), Parallelization strategies for rapid and robust evolutionary multiobjective optimization in water resources applications, *Adv. Water Resour.*, *30*(3), 335–353.

Tang, Y., P. M. Reed, T. Wagener, and K. van Werkhoven (2007b), Comparing sensitivity analysis methods to advance lumped watershed model identification and evaluation, *Hydrol. Earth Syst. Sci.*, *11*, 793–817.

Tang, Y., P. M. Reed, T. Wagener, and K. van Werkhoven (2007c), Advancing the identification and evaluation of distributed rainfall-runoff models using Sobol's global sensitivity analysis, *Water Resour. Res.*, *43*, W06415, doi:10.1029/2006WR005813.

Van Griensven, A., T. Meixner, S. Grunwald, T. Bishop, A. Diluzio, and R. Srinivasan (2006), A global sensitivity analysis tool for the parameters of multi-variable catchment models, *J. Hydrol.*, *324*(1–4), 10–23.

Vrugt, J. A., H. V. Gupta, W. Bouten, and S. Sorooshian (2003a), A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters, *Water Resour. Res.*, *39*(8), 1201, doi:10.1029/2002WR001642.

Vrugt, J. A., H. V. Gupta, L. A. Bastidas, W. Bouten, and S. Sorooshian (2003b), Effective and efficient algorithm for multiobjective optimization of hydrologic models, *Water Resour. Res.*, *39*(8), 1214, doi:10.1029/2002WR001746.

Vrugt, J. A., H. V. Gupta, S. C. Dekker, S. Sorooshian, T. Wagener, and W. Bouten (2006), Application of stochastic parameter optimization to the Sacramento Soil Moisture Accounting Model, *J. Hydrol.*, *325*(1–4), 288–307.

Wagener, T. (2003), Evaluation of catchment models, *Hydrol. Process.*, *17*(16), 3375–3378.

Wagener, T., and H. S. Wheater (2006), Parameter estimation and regionalization for continuous rainfall-runoff models including uncertainty, *J. Hydrol.*, *320*(1–2), 132–154.

Wagener, T., D. P. Boyle, M. J. Lees, H. S. Wheater, H. V. Gupta, and S. Sorooshian (2001), A framework for development and application of hydrological models, *Hydrol. Earth Syst. Sci.*, *5*(1), 13–26.

Wagener, T., N. McIntyre, M. J. Lees, H. S. Wheater, and H. V. Gupta (2003), Towards reduced uncertainty in conceptual rainfall-runoff modelling: Dynamic identifiability analysis, *Hydrol. Process.*, *17*(2), 455–476.

William, H., S. A. Teukolsky, T. V. William, and P. F. Brian (1999), *Numerical Recipes in C*, 2nd ed., Cambridge Univ. Press, New York.

P. M. Reed, Y. Tang, K. van Werkhoven, and T. Wagener, Department of Civil and Environmental Engineering, Pennsylvania State University, 212 Sackett Building, University Park, PA 16802, USA. (klg292@psu.edu)