Advances in Water Resources 34 (2011) 114–127

Contents lists available at ScienceDirect

Advances in Water Resources

journal homepage: www.elsevier.com/locate/advwatres

# Characterizing parameter sensitivity and uncertainty for a snow model across hydroclimatic regimes

Minxue He [a,f,g], Terri S. Hogue [a,*], Kristie J. Franz [b], Steven A. Margulis [a], Jasper A. Vrugt [c,d,e]

[a] Department of Civil and Environmental Engineering, University of California, Los Angeles, United States
[b] Department of Geologic and Atmospheric Sciences, Iowa State University, United States
[c] Center for Nonlinear Studies, Los Alamos National Laboratory, Los Alamos, United States
[d] Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, The Netherlands
[e] Department of Civil and Environmental Engineering, University of California, Irvine, United States
[f] Office of Hydrologic Development, NOAA/National Weather Service, Silver Spring, United States
[g] Riverside Technology, Inc., Fort Collins, United States

ARTICLE INFO

*Article history:*
Received 21 September 2009
Received in revised form 27 September 2010
Accepted 1 October 2010
Available online 15 October 2010

*Keywords:*
National Weather Service
SNOW17
Snowmelt
Uncertainty
Generalized Sensitivity Analysis
Differential Evolution Adaptive Metropolis

ABSTRACT

The National Weather Service (NWS) uses the SNOW17 model to forecast snow accumulation and ablation processes in snow-dominated watersheds nationwide. Successful application of the SNOW17 relies heavily on site-specific estimation of model parameters. The current study undertakes a comprehensive sensitivity and uncertainty analysis of SNOW17 model parameters using forcing and snow water equivalent (SWE) data from 12 sites with differing meteorological and geographic characteristics. The Generalized Sensitivity Analysis and the recently developed Differential Evolution Adaptive Metropolis (DREAM) algorithm are utilized to explore the parameter space and assess model parametric and predictive uncertainty. Results indicate that SNOW17 parameter sensitivity and uncertainty generally varies between sites. Of the six hydroclimatic characteristics studied, only air temperature shows strong correlation with the sensitivity and uncertainty ranges of two parameters, while precipitation is highly correlated with the uncertainty of one parameter. Posterior marginal distributions of two parameters are also shown to be site-dependent in terms of distribution type. The SNOW17 prediction ensembles generated by the DREAM-derived posterior parameter sets contain most of the observed SWE. The proposed uncertainty analysis provides posterior parameter information on parameter uncertainty and distribution types that can serve as a foundation for a data assimilation framework for hydrologic models.



## 1. Introduction

A range of studies demonstrate that atmospheric warming over the past 50 years has led to a steady decline in snowpack depth as well as altered melt patterns across large regions of the western US [34,35,44,54]. Changes in snowpack volume and spring melt timing substantially alter the volume and pattern of streamflow in snow-dominated watersheds, increasing the probability of extreme flooding events and droughts. Accurate prediction of snowmelt is vital to support optimal water resources planning and management practices [49]. The increasing western population and corresponding water demand, as well as potential climatic extremes [67], make accurate snowmelt and runoff predictions in the western regions of the US especially critical.

The National Weather Service (NWS), the US agency responsible for short- and long-term streamflow predictions across the nation, primarily applies the SNOW17 model [3] for operational forecasting of snow accumulation and melt in snow-dominated areas. The SNOW17 is a process-based model requiring only precipitation and air temperature as inputs [39,45]. Despite its conceptual nature and low input data requirements, the model demonstrates similar, and often better, performance than more detailed physically-based snow models in terms of snow water equivalent and subsequent streamflow predictions [15,16]. Although the SNOW17 model has been extensively utilized in the operational environment for several decades, comprehensive sensitivity and uncertainty analyses of its model parameters are rare in the literature. Tang et al. [56] conducted a sensitivity analysis of the SNOW17 model in conjunction with the NWS operational Sacramento Soil Moisture Accounting (SAC–SMA) rainfall–runoff model. However, only about half of the SNOW17 parameters were analyzed. The primary focus of the Tang et al. [56] study was to evaluate different sensitivity analysis methods rather than to explicitly quantify parameter behavior or uncertainty in the SNOW17 model. The NWS is currently investigating

* Corresponding author. Address: Department of Civil and Environmental Engineering, 5731F Boelter Hall, University of California, Los Angeles, CA 90095-1593, United States. Tel.: +1 310 794 4239.
*E-mail address:* thogue@seas.ucla.edu (T.S. Hogue).

0309-1708/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.advwatres.2010.10.002

distributed versions of the SNOW17 and SAC–SMA models for operational forecasts within the Hydrology Laboratory Research Distribution Hydrologic Model (HL-RDHM) [26,40,51] framework. Preliminary application of the distributed SNOW17 model to operational watersheds has been conducted [46,47]. More in-depth exploration of the full potential of distributed snow modeling requires a rigorous investigation of the sensitivity and uncertainty of SNOW17 parameters and would assist in parameter identification for subwatershed regions with varying climatic and geographic characteristics. Previous studies illustrate that hydrologic model parameter sensitivity varies significantly within watersheds with diverse climatic and geographic characteristics [48,58] and calibration (automatic or otherwise) of distributed hydrologic models has been shown to be challenging [30]. A sensitivity and uncertainty analysis can help identify potential relationships among watershed characteristics and related parameter behavior, as well as facilitate hydrologic forecasting in ungauged regions.

In the current study, we combine the Generalized Sensitivity Analysis (GSA) [23,53] and the DiffeRential Evolution Adaptive Metropolis (DREAM) algorithm to assess parameter sensitivity and uncertainty of the SNOW17 parameters. To reach broad conclusions (regionalization information), we use forcing and snow water equivalent data from 12 contrasting study sites. The GSA method serves as a first step in determining which SNOW17 parameters are sensitive across a range of climate conditions and should be included in DREAM. DREAM is used to determine posterior parameter distributions and related uncertainty. We are especially concerned with identifying potential regionalization relationships, correlating parameter sensitivity and uncertainty with site characteristics and analyzing posterior probability density functions of sensitive parameters. The derived parameter distributions serve as a basic building block for the development of an ensemble-based sequential data assimilation framework that requires parameter sets to be sampled from their distributions. The framework is currently being developed for the SNOW17 model and the SAC–SMA model.

## 2. Methodology

The current study focuses at the point-scale given that ground-based snow observations are most readily available at this resolution. We advocate that rigorous investigation at the point-scale (with ground-based measurements) provides a thorough assessment and understanding of model behavior before applying relevant models to larger scales using alternative data sources [15,22]. We utilize an extensive set of observations from SNOw TELmetry (SNOTEL) sites available from the Natural Resources Conservation Service (NRCS). Although SNOTEL observations are associated with some uncertainty, they are still the longest and most widely used record of snow depth and snow water equivalent (SWE) across a range of hydroclimatic conditions in the US [43].

### 2.1. SNOW17 model

The SNOW17 is a lumped process-based model that simulates snow accumulation and ablation (Fig. 1) [3,39]. Inputs to the model are air temperature and precipitation. Model output include snow melt plus any rain on bare ground, and SWE. Snow is modeled as a single layer and the heat storage of the snowpack, snow melt, and liquid water retention and transmission are computed using empirically-based relationships. The model has ten primary parameters for point-scale simulation (Table 1) [2,13]. When applied to the areal domain, the model implements an additional parameter SI (the mean areal water equivalent above which 100% snow cover exists) and an areal depletion curve (ADC) that

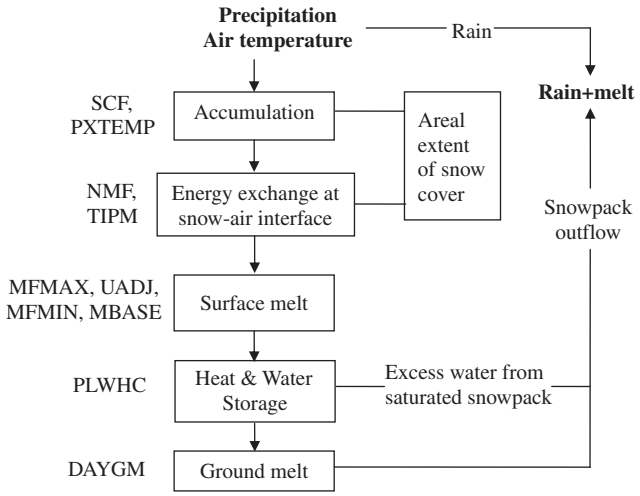


**Fig. 1.** Schematic of the SNOW17 model processes and corresponding parameters. The boxes designate model processes. Model inputs and output are highlighted in bold. *Source*: Adapted from Anderson [3].

accounts for the areal extent of snow cover. The SCF (snow correction factor) parameter is a multiplier of the precipitation input used to account for gage catch deficiencies. The PXTEMP parameter is the threshold temperature distinguishing snowfall from rainfall in the precipitation time series. The parameters SCF and PXTEMP determine the snow accumulation in the model.

Parameters NMF and TIPM are used to determine heat exchange during non-melt periods, which is typically less significant than during melt periods [3]. Heat content in the snowpack increases or decreases as a function of the gradient between the antecedent temperature (as determined by parameter TIPM) and the current air temperature [3]. Heat conduction through the snowpack is assumed to vary similarly to the non-rain melt factor and is scaled by a negative melt factor (NMF). Melt occurs when enough energy has been added to the snowpack to bring its heat content to zero.

During non-rain periods the depth of melt is determined by:

$$M = M_f(T_a - \mathrm{MBASE}) \tag{1}$$

where $M$ is the depth of melt (mm), $M_f$ is the seasonally varying melt factor (mm/°C), and MBASE is the temperature above which melt will occur. The melt factor is computed from a sinusoidal curve with limits defined by the maximum (MFMAX) and minimum (MFMIN) melt factor parameters as follows [3]:

$$M_f = \frac{\Delta t}{6}\left\{\frac{1}{2}\left(\sin\frac{2\pi T}{366} + 1\right)(\mathrm{MFMAX} - \mathrm{MFMIN}) + \mathrm{MFMIN}\right\} \tag{2}$$

where $\Delta t$ is the time step; $T$ is the day number since March 21. MFMAX is generally more critical than the MFMIN in determining actual melt rate when most of melt occurs after March 21, while MFMIN is more critical prior to March 21 [39]. Energy balance equations are used to compute melt during rain-on-snow events utilizing assumptions about meteorological conditions during rainy periods [3]. Melt during rain-on-snow events is controlled by UADJ, a parameter controlling the impact of wind advection. When there is sufficient rainfall and the air temperature is well above freezing, increasing UADJ results in more melt [39].

Excess water will occur in the pack when it is isothermal at 0 °C and the liquid water holding capacity of the pack (as determined by parameter PLWHC) is met. Excess water is lagged and attenuated to simulate flow through the pack based on a series of empirically derived equations for ripe snow. A constant daily rate of melt at the soil-snow interface (DAYGM) is parameterized in the model to account for the heat flux at the soil-snow interface [3].

**Table 1**
Parameters of the SNOW17 model with ranges estimated from Anderson [3] and Franz [13].

| Parameters | Description | Unit | Ranges |
|---|---|---|---|
| SCF | Snowfall correction factor | – | 0.7–1.4 |
| MFMAX | Maximum melt factor considered to occur on June 21 | mm/6 h/°C | 0.5–2.0 |
| MFMIN | Minimum melt factor considered to occur on December 21 | mm/6 h/°C | 0.05–0.49 |
| UADJ | The average wind function during rain-on-snow periods | mm/mb/°C | 0.03–0.19 |
| NMF | Maximum negative melt factor | mm/6 h/°C | 0.05–0.50 |
| MBASE | Base temperature for non-rain melt factor | °C | 0–1.0 |
| PXTEMP | Temperature that separates rain from snow | °C | −2.0 to 2.0 |
| PLWHC | Percent of liquid–water capacity | – | 0.02–0.3 |
| DAYGM | Daily melt at snow–soil interface | mm/day | 0–0.3 |
| TIPM | Antecedent snow temperature index | – | 0.1–1.0 |

### 2.2. Study sites and datasets

The SNOTEL network, managed by the Natural Resources Conservation Service (NRCS), consists of 730 sites throughout 11 western states (including Alaska), providing snowpack information for water supply forecasting and water resources management purposes [38]. A typical SNOTEL site has sensors which automatically measure SWE, snow depth, precipitation, and air temperature [8]. The data is provided at the daily timestep. SWE and precipitation data are generally available since the early 1980s. Air temperature is generally available since the late 1980s. The data quality is largely controlled by instrumentation sensitivities and environmental factors including snow drifting, wind scour, and falling debris [10,12].

The current study focuses on observations from 12 SNOTEL sites (Fig. 2; Table 2) with at least one site chosen from each of the eight regions defined by Serreze et al. [42]. These defined regions (Fig. 2) capture the major snow regimes of the Western US [42] and have been used in a regional analysis of snowfall events [43]. The 12 selected study sites are also within or close to current NWS operational basins. For some regions with distinctive characteristics, such as the Sierra Nevada region (higher precipitation and snowfall), two sites are selected. The selection is also based upon using a representative median elevation in each region (i.e. the elevation of the selected site(s) resembles the median site elevation of the corresponding region). Elevation of the study sites varies from 1207 m (site MC (WA)) to 3316 m (site VL (CO)); average annual precipitation ranges from 624 mm (site BL (CO)) to 2503 mm (site LL (CA)); average annual air temperature varies from −0.45 °C (site BL (CO)) to 7.94 °C (site WH (AZ)) (Table 2). The maximum monthly precipitation varies from 59 mm (April, site BL (CO)) to 396 mm (January, site LL (CA)). The range for maximum monthly SWE is from 81 mm (February, Site WH (AZ)) to 1519 mm (April, site LL (CA)). The maximum and minimum monthly air temperature range from 10 °C (July, Site BL (CO)) to 18.8 °C (July, Site WH (AZ)) and from −9.7 °C (December, Site BL (CO)) to 1.0 °C (January, Site HM (OR)), respectively (Fig. 3). These wide variations exemplify the diverse climatic and geographic characteristics of the selected study sites. A detailed discussion of these study sites is also presented by He [21].

Daily precipitation, temperature, and SWE data for the 12 selected SNOTEL sites were screened following the quality control

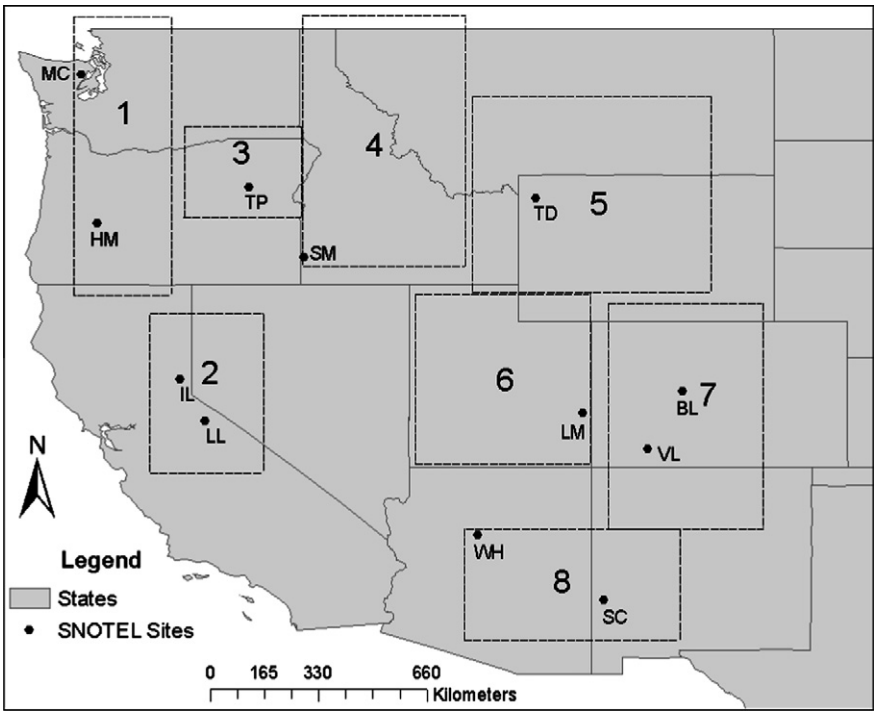


**Fig. 2.** Location of study sites as well as eight regions (boxes) encompassing the study sites (from Serreze et al. [42]). The regions and the median elevation of SNOTEL sites within the regions include: (1) Pacific Northwest (1422 m), (2) Sierra Nevada (2439 m), (3) Blue Mountains, Oregon (1646 m), (4) Idaho/western Montana (1905 m), (5) NW Wyoming (2479 m), (6) Utah (2774 m), (7) Colorado (3037 m), and (8) Arizona/New Mexico (2418 m). Detailed information on the regions and SNOTEL sites can be found in Serreze et al. [42].

**Table 2**
Summary of selected SNOTEL sites. Statistics are computed by water year (WY) October 1 to September 30. Max. SWE represents the maximum SWE value recorded at the site. Annual temperature and precipitation are mean annual values. Date of Max. SWE is the mean value of the dates of annual maximum SWE. Date of disappearance of snow and days of snowmelt are also mean values for the period of record. Days of snowmelt represent the period from the day SWE peaks till the day snow melts out.

| Site | Site name | State | Elevation (m) | Climate region | Data period (WY) | Max. SWE (mm) |
|---|---|---|---|---|---|---|
| MC | Mount Crag | WA | 1207 | 1 | 1991–2008 | 2073 |
| HM | Holland Meadows | OR | 1503 | 1 | 1984–2008 | 1255 |
| IL | Independence Lake | CA | 2546 | 2 | 1995–2008 | 2068 |
| LL | Leavitt Lake | CA | 2931 | 2 | 1990–2008 | 2807 |
| TP | Tipon | OR | 1570 | 3 | 1990–2008 | 439 |
| SM | South Mtn | ID | 1981 | 4 | 1985–2008 | 828 |
| TD | Thumb Divide | WY | 2432 | 5 | 1990–2008 | 836 |
| LM | Lasal Mtn | UT | 2914 | 6 | 1989–2008 | 582 |
| BL | Brumley | CO | 3231 | 7 | 1987–2008 | 485 |
| VL | Vallecito | CO | 3316 | 7 | 1987–2008 | 874 |
| SC | Silver Creek Divide | NM | 2743 | 8 | 1990–2008 | 630 |
| WH | White Horse Lake | AZ | 2188 | 8 | 1990–2008 | 343 |

| Site | Annual temperature (°C) | Annual precipitation (mm) | Snowfall fraction | Date of Max. SWE | Date of disappearance of snow | Days of snowmelt |
|---|---|---|---|---|---|---|
| MC | 4.35 | 1922 | 0.42 | 3-April | 6-June | 63 |
| HM | 6.36 | 1964 | 0.34 | 3-March | 18-May | 66 |
| IL | 4.38 | 2151[a] | 0.60 | 29-April | 2-July | 63 |
| LL | 2.74 | 2503[a] | 0.64 | 25-April | 4-July | 69 |
| TP | 5.37 | 589 | 0.52 | 18-March | 2-May | 44 |
| SM | 6.44 | 919[a] | 0.52 | 18-March | 6-May | 48 |
| TD | 1.02 | 758 | 0.58 | 11-April | 24-May | 42 |
| LM | 3.87 | 819 | 0.40 | 25-March | 11-May | 46 |
| BL | −0.45 | 624 | 0.48 | 16-April | 19-May | 32 |
| VL | 2.56 | 865 | 0.57 | 5-April | 20-May | 44 |
| SC | 5.64 | 671 | 0.44 | 4-March | 19-April | 45 |
| WH | 7.94 | 628 | 0.22 | 27-February | 2-April | 33 |

[a] Corrected values obtained by applying the method of Serreze et al. [42].

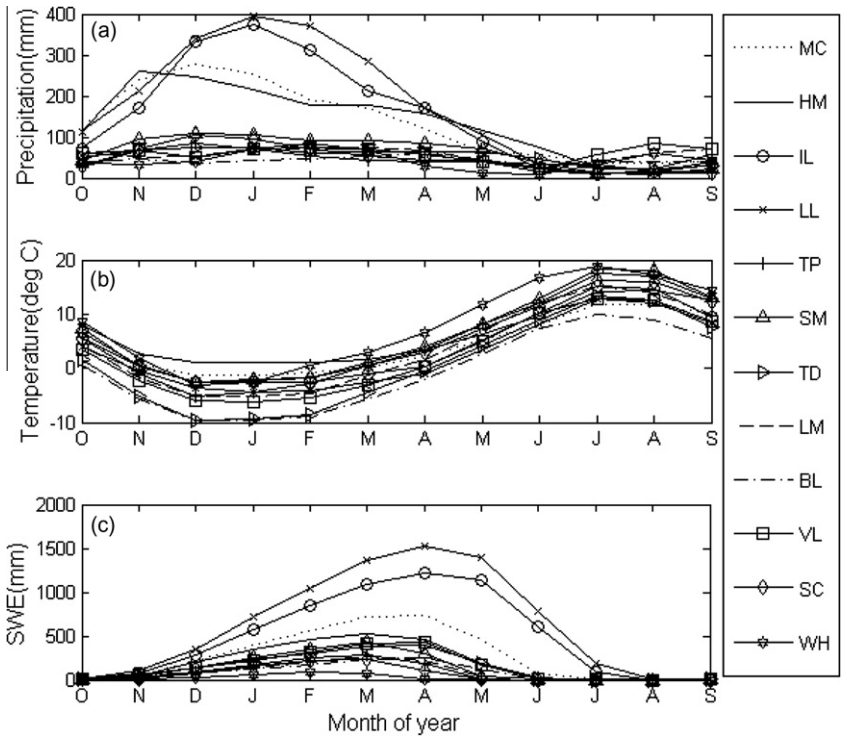


**Fig. 3.** Monthly (a) precipitation, (b) temperature, and (c) SWE for the 12 SNOTEL sites presented in Table 2.

procedures of Serreze et al. [42] to mask outliers and identify negative SWE and precipitation values. This methodology is applied in the studies of Serreze et al. [43] and Fassnacht et al. [12]. After screening, the longest historical record is 25 years (site HM (OR)) and the shortest is 14 years (site IL (CA)). In a few cases, temperature data are missing (about 1% of the record length). Under these conditions, data from the prior date and subsequent date is used to linearly interpolate missing data. For sites IL (CA), LL (CA), and SM (ID), the snowfall fraction (ratio of SWE to precipitation) exceeds 1.0 in several years. In such cases, the under-catch in precipitation is estimated using the method proposed by Serreze et al. [42]. The method uses the ratio between daily SWE increments exceeding

2.5 cm and the corresponding precipitation data in January and February to upwardly adjust precipitation measurements. As the shortest data period for the selected sites is 14 years (water year 1995–2008, site IL (CA)), this common 14-year period is used for analysis.

### 2.3. Sensitivity and uncertainty analysis

It is generally accepted that sensitivity and uncertainty analyses should be conducted as part of calibration efforts and model implementation. These analyses are especially necessary for data assimilation studies which require the structure of all uncertainty sources specified *a priori* [31,41,50]. Generally speaking, a sensitivity analysis is the practice of identifying the primary parameters which dominate model performance [20], while an uncertainty analysis is the practice of determining the reliability of model outputs by addressing potential sources of uncertainty, including uncertainty associated model structure, forcing, parameters and observations [4,25,36]. A recent study [7] reviewed the advantages and disadvantages of a range of sensitivity analysis methods. Their review noted that the Latin Hypercube Sampling (LHS), utilized within the Generalized Sensitivity Analysis (GSA) [23,53], was acceptable for sensitivity analysis.

A range of uncertainty analysis methods have also been reported in the hydrologic literature, including the generalized likelihood uncertainty estimation (GLUE) [5], the Bayesian recursive estimation (BARE) [57], the dynamic identifiability analysis (DYNIA) [66], and Markov Chain Monte Carlo (MCMC) methods including the Metropolis algorithm [28], the parameter identification method based on the localization of information (PIMLI) [59], the Shuffled Complex Evolution Metropolis (SCEM-UA) algorithm [60], and the Differential Evolution Adaptive Metropolis (DREAM) algorithm [62]. The DREAM algorithm is an adaptation of the SCEM-UA and has advantages over the SCEM-UA in the context of maintaining detailed balance and ergodicity [62,63]. The DREAM algorithm also exhibits the best search efficiency in comparison to its counterparts. Most of the above methods focused only on the uncertainty that is associated with the selection of parameters. The disregard of other uncertainty sources (observational data, model structure and states) in many of the current methods stems from the fact that there is no rigorous or widely agreed upon method to characterize them [29,33]. Some methods have recently emerged to investigate model forcing and/or structural uncertainty [1,24,50,63], however, they have not been commonly applied.

The current work sequentially applies the GSA and DREAM algorithm to assess parameter sensitivity and uncertainty. The GSA method serves as a simple screening tool for determining which SNOW17 parameters are sensitive across a range of climate conditions and should be included in the more complex DREAM algorithm which determines posterior parameter distributions and related uncertainty.

#### 2.3.1. The Generalized Sensitivity Analysis (GSA) method

The GSA method was originally proposed by Spear and Hornberger [53] and has received extensive application in hydrology [e.g., 17,48,52,56,64]. The GSA has three primary components: Monte Carlo sampling, "behavioral/non-behavioral" classification, and sensitivity assessment. Monte Carlo sampling generates parameter sets in the feasible parameter space from prior distributions. The current study applies Latin Hypercube Sampling (LHS) to generate random realizations of parameters because it has been shown to be effective in previous studies [6,7,32,36,48,55,69]. The LHS is a stratified Monte Carlo sampling method in which parameter ranges are equally divided into $N$ intervals ($N$ is the number of samples) and one realization is sampled from each interval. This particular approach is easy to implement and use, yet could become computationally cumbersome and inefficient in the presence of high-dimensionality. In the present study, we assume a uniform sampling distribution for each of the SNOW17 parameter with ranges that are listed in Table 1. We also utilize the Nash–Sutcliffe efficiency (*NSE*) measure [37] to quantify model performance and to distinguish behavioral from non-behavioral parameter sets. The *NSE* is defined as:

$$NSE = 1 - \frac{\sum_{i=1}^{n}\left(Q_i^{obs} - Q_i^{sim}\right)^2}{\sum_{i=1}^{n}\left(Q_i^{obs} - Q_{obs}^{mean}\right)^2} \tag{3}$$

where $Q^{obs}$ and $Q^{sim}$ are observed and simulated states at time $i$, respectively; and $Q_{obs}^{mean}$ is the mean value of the observed states. Based on previous studies [e.g. 18] we utilize a threshold *NSE* value of 0.3 to distinguish between behavioral ($NSE > 0.3$) and non-behavioral ($NSE < 0.3$) solutions. The behavioral parameter sets are then divided into 10 equally sized groups based on sorted *NSE* values [17,65]. For each parameter, cumulative distributions of the parameter within each of the 10 groups are plotted, producing 10 cumulative distribution curves (CDCs). Similarity (closeness) between the produced CDCs reveals that the parameter is insensitive, whereas spread between these CDCs suggests high parameter sensitivity [17]. The Kolmogorov–Smirnov (*KS*) test [27] is applied to quantify the difference between the CDCs, where the calculated *KS* value is the maximum vertical distance between the 10 CDCs. The *KS* value ranges from 0 to 1, with higher values indicating higher parameter sensitivities.

Based on literature recommendations [56], 10,000 parameter sets are generated within their feasible ranges and a corresponding 10,000 SWE estimates (model simulations) are generated for each of the 12 sites. Behavioral sets are selected based on the *NSE* value, corresponding CDCs are generated, and the *KS* statistic is estimated. On the basis of the *KS* results, insensitive parameters are identified. Finally, the correlation between parameter sensitivity and site characteristics is assessed.

#### 2.3.2. The Differential Evolution Adaptive Metropolis (DREAM) algorithm

The Differential Evolution Adaptive Metropolis (DREAM) algorithm was recently introduced by Vrugt et al. [62] to estimate optimal parameter values and their underlying posterior probability density function within a single run. This MCMC scheme is an adaptation of the Shuffled Complex Evolution Metropolis [60] global optimization algorithm and has the advantage of maintaining detailed balance and ergodicity while showing good efficiency on complex, highly nonlinear, and multimodal target distributions [61–63].

In the DREAM algorithm, a preset number ($N$) of Markov Chains (a chain refers to a vector containing one parameter realization) are simultaneously run in parallel. The chains are initialized by randomly sampling the parameter space using the specified prior distribution. These chains form a population, conveniently stored as a $N \times d$ matrix $\mathbf{X}$, with $d$ the dimension of the parameter space. For each chain, $i \in \{1,2,\ldots,N\}$, a candidate point $\mathbf{z}_i$ (vector) is generated by taking a fixed multiple of the difference between randomly chosen pairs of chains (without replacement) of $\mathbf{X}_{-i}$ ($\mathbf{X}$ without $\mathbf{x}_i$) [62]:

$$\mathbf{z}_i = \mathbf{x}_i + (\mathbf{1}_d + \mathbf{e})\gamma(\delta, d_{eff})\left[\sum_{j=1}^{\delta}\mathbf{x}_{r_1(j)} - \sum_{n=1}^{\delta}\mathbf{x}_{r_2(n)}\right] + \boldsymbol{\varepsilon} \tag{4}$$

where $\delta$ signifies the number of chain pairs, $\gamma$ is a jumprate, and $r_1(j), r_2(n) \in \{1,2,\ldots,N\}$ but $r_1(j) \neq r_2(n) \neq i$. The value of $\mathbf{e}$ is drawn from $U_d(-b, b)$ with $|b| < 1$, and $\boldsymbol{\varepsilon} \sim N_d(0, b^*)$ is a white noise term with $b^*$ small compared to the width of the target distribution.

The Metropolis ratio is used to decide whether to accept the candidate point or not. If accepted, the chain moves from $\mathbf{x}_i$ to $\mathbf{z}_i$, otherwise the location of the chain remains unchanged. From the guidelines of $\gamma$ in Random Walk Metropolis (RWM), a good choice of $\gamma = 2.38/\sqrt{2\delta d_{eff}}$, where $d_{eff}$ denotes the number of dimensions that will be updated. With this approach, a Markov chain is obtained, the stationary distribution of which is the posterior distribution. The proof of this is presented in Vrugt et al. [61]. After a so-called burn-in period, the convergence of a DREAM run can be monitored with the $\widehat{R}$-statistic of Gelman and Rubin [19], which compares the variance within and between the chains. A value of $\widehat{R}$ smaller than 1.2 for each parameter ($\widehat{R}_k < 1.2,\ k = 1, 2, \ldots, d$) diagnoses convergence to a limiting distribution. A detailed description of DREAM appears in Vrugt et al. [62,63], and so will not be repeated here. The robustness and usefulness of the DREAM algorithm has been demonstrated previously. Using a variety of different case studies, Vrugt et al. [61] showed that DREAM works really well compared to existing MCMC schemes. Moreover, convergence results demonstrate that DREAM can also outperform the widely used SCE-UA global optimization algorithm [11].

In the current study, DREAM is applied with standard algorithmic settings to determine the posterior distribution of the SNOW17 parameters. This distribution contains all desired information about the posterior mean, uncertainty intervals, and cross-correlation of each individual parameter. In keeping with previous studies, we use a traditional least squares likelihood function containing the deviation between measured and SNOW17 predicted SWE. For each DREAM trial, a maximum total of 35,000 SNOW17 model runs were used. Preliminary tests have shown that this number of model evaluations is sufficient to obtain convergence to the posterior distribution according to the $\widehat{R}$-statistic of [19].

When using DREAM, the uncertainty analysis is presented in terms of: (I) performance of the traditional best parameter combination (best posterior parameter set), (II) performance of all posterior parameter samples (95% prediction uncertainty associated with the posterior parameter sets is used), and (III) uncertainty and correlation of posterior parameters. The best parameter set refers to the one producing the lowest (best) *RMSE* value. In presenting the best parameter set (simulation) at each site, other metrics including the *NSE* and percent bias are applied. In assessing the prediction of the ensemble of posterior parameter sets (with $\widehat{R}$-statistic less than 1.2), the focus is on the coverage of SWE observations by SWE predictions estimated via these parameter sets. The coverage is hereinafter defined as the containing ratio (*CR*) which stands for the percentage of SWE observations contained within the bounds of the SWE simulated by the ensembles, evaluated over the simulation period. The *CR* ranges from 0 to 1 with higher value indicating better coverage [68]. For comparison purposes, the performance of the prior parameter sets is also assessed. Note that the prior parameter sets are those generated from the initial uniform distribution using the LHS sampling method. To facilitate comparison, we also evaluate the *CR* in terms of the spread of the prediction ensemble, estimating a normalized containing ratio (*NCR*), defined as:

$$NCR = \frac{CR}{M} = \frac{CR}{\frac{1}{n}\sum_{i=1}^{n}\left(SWE_i^U - SWE_i^L\right)} \tag{5}$$

where *CR* designates the unitless containing ratio; *M* represents the time-averaged width of prediction bounds, or the difference between the upper $\left(SWE_i^U\right)$ and lower bound $\left(SWE_i^L\right)$ of SWE estimates, and *n* is the number of simulation days. A higher *NCR* value indicates a high containing ratio associated with low forecast uncertainty (low *M*), whereas a lower *NCR* value (near 0) could indicate a low containing ratio or, alternatively, a high containing ratio associated with high forecast uncertainty (high *M*). In either case, a low *NCR* would represent an unwanted situation in model predictions. The prior and posterior SWE prediction ensemble is also compared to the prediction produced from the widely used SCE-UA global optimization algorithm using a standard least squares objective function [11].

The performance of both the best parameter set and the ensemble of parameter sets are first evaluated at the 12 study sites within the 14-year study period. Next, performance is examined at annual time scales at three selected sites. These three sites, LL (CA) from the Pacific Northwest, BL (CO) from Colorado, and WH (AZ) from the Arizona/New Mexico region, represent a range of average annual snowfall, including high, medium, and scarce snowfall, respectively. The marginal distribution of posterior parameters and Spearman's rank correlation are investigated at the three selected sites for brevity.

## 3. Results and discussion

### 3.1. Sensitivity analysis

The *KS* values for model parameters (Table 1) at the 12 study sites are computed (Fig. 4). Results show that parameters TIPM and NMF have generally lower *KS* values for all the sites with maximum *KS* values of 0.20 and 0.25, respectively. This reveals that parameters TIPM and NMF involved in calculating heat transfer during non-melt periods are insensitive at the study sites. PLWHC also has consistently low *KS* values across the study sites. SCF, PXTEMP, and MFMAX have significantly higher *KS* values, with average *KS* values of 0.66, 0.52, and 0.46, respectively. As discussed previously, SCF and PXTEMP directly alter the volume of simulated SWE in the SNOW17. The high sensitivity of MFMAX is due to the fact that MFMAX is the primary control on snowmelt rate during non-rain periods after March 21; non-rain melt dominates in the Western US and the primary melt season occurs after March 21 at most of the study sites. Parameters UADJ and MFMIN have less overall sensitivity. The sensitivity of MBASE and DAYGM is small, but these parameters are more sensitive at several sites than parameters NMF, TIPM, and PLWHC. Given these results, parameters NMF, TIPM, and PLWHC are excluded from further analyses.

For the more sensitive parameters, the degree of sensitivity varies from site to site, suggesting that the physical site characteristics (topographic and climatic factors) influence parameter sensitivity (Fig. 4). To more closely examine this relationship, the correlation between *KS* values and site information are evaluated (Fig. 5). The metric applied is the Spearman's rank correlation coefficient (Spearman's $\rho$), which ranges from $-1$ to 1 with a higher absolute value indicating higher correlations. We also estimate a corresponding *p*-value statistic to determine if the Spearman rank correlation is significant or not, with a value less than 0.05 indicating significant correlation. In general, parameter sensitivity shows low correlations (correlation coefficient less than 0.6) with most site characteristics (Fig. 5). An exception is that site mean annual air temperature exhibits negative correlations with the sensitivity of parameters SCF and DAYGM. The corresponding Spearman's $\rho$ are $-0.76$ and $-0.62$, respectively. The *p*-values are 0.01 and 0.03, respectively, indicating that the correlation is significant. As noted, SCF controls the mass of the snowfall input in the accumulation period, with a higher SCF value leading to higher snowfall input under a given set of air temperature time series; whereas air temperature inversely regulates snowfall input in the same period with higher temperature resulting in lower snowfall with a fixed SCF. In the accumulation season, when the temperature is higher, the impact of SCF is expected to be less significant on model simulations. The dependency of model estimates on SCF is therefore

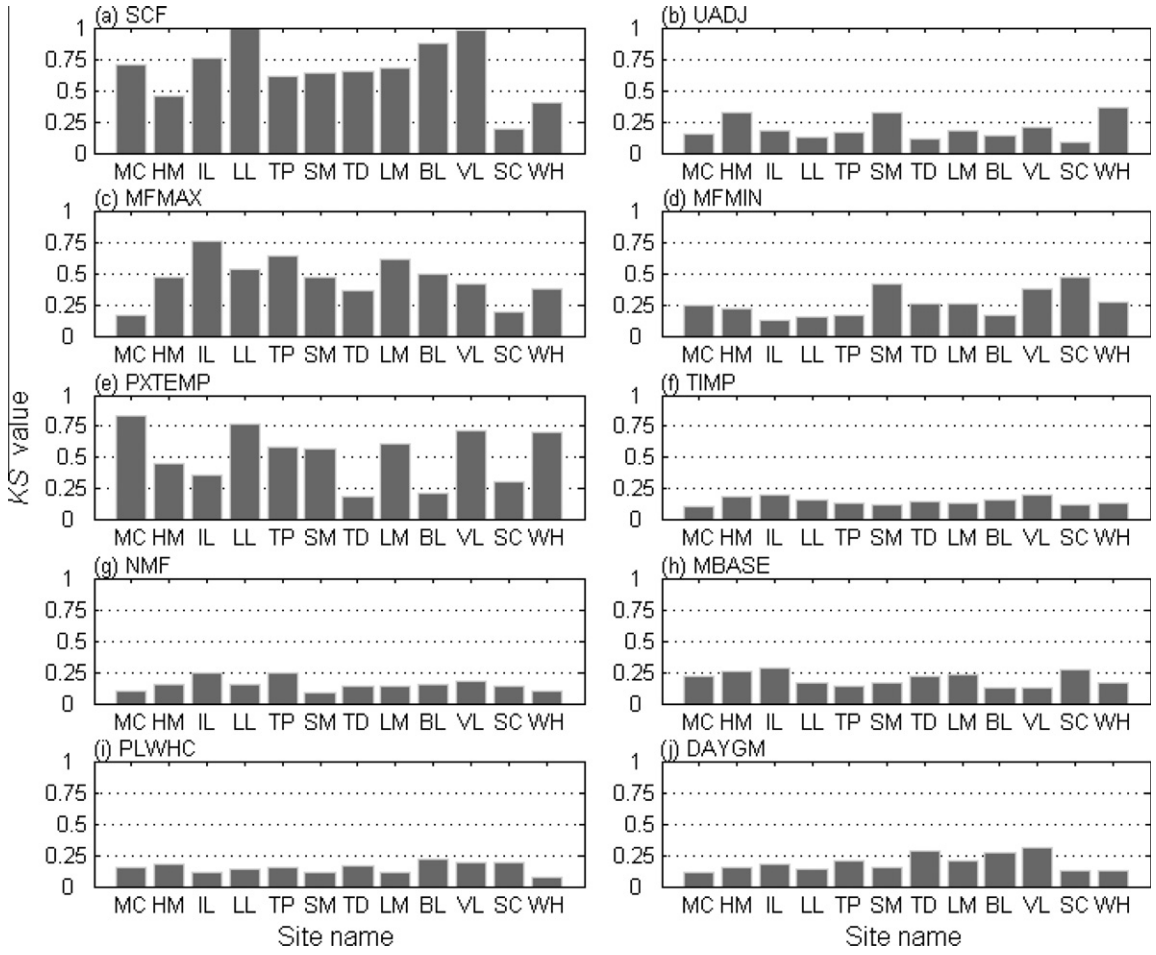


**Fig. 4.** *KS* values associated with parameters (a) SCF, (b) UADJ, (c) MFMAX, (d) MFMIN, (e) PXTEMP, (f) TIMP, (g) NMF, (h) MBASE, (i) PLWHC, and (j) DAYGM at 12 SNOTEL sites. *X*-axis includes abbreviations of study sites.

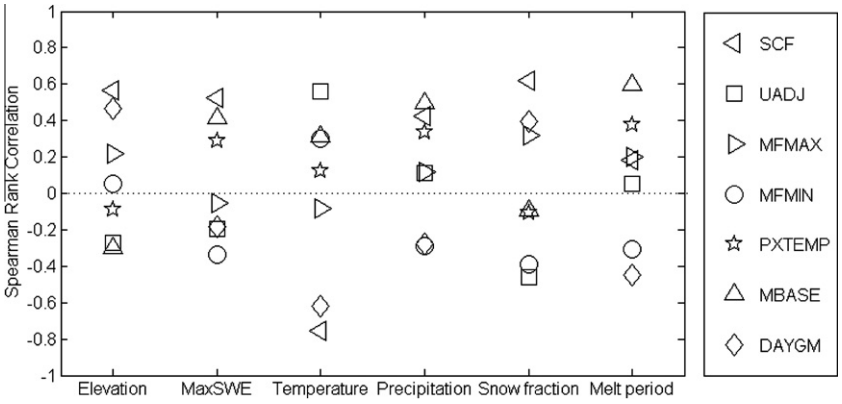


**Fig. 5.** Correlation between parameter sensitivity and topographic and climatic conditions associated with study sites for seven parameters: SCF, UADJ, MFMAX, MFMIN, PXTEMP, MBASE, and DAYGM.

less, which explains the negative correlation between air temperature and the sensitivity of parameter SCF. Similarly, under conditions where there is heat exchange between the soil surface and snowpack (soil surface is not isothermal at 0 °C), higher temperatures result in a shallower snowpack, which would be expected to have a colder base and lower heat exchange with the ground surface. Consequently, DAYGM would be lower for a shallower snowpack and less important in determining overall snowpack water balance. Site snowfall fraction is also correlated to SCF sensitivity, with a coefficient of 0.62 and *p*-value of 0.03. This is also expected since higher SCF values result in more snowfall input, thus increasing the snowfall fraction. It should be noted that the correlation between MBASE sensitivity and melting period is 0.59 which is close to, but less than, the threshold value (0.60).

Due to the insensitivity of parameters NMF, TIPM, and PLWHC at all study sites, these parameters are fixed at constant values in the DREAM analysis. At a few sites, the sensitivity of MBASE and DAYGM is evident; therefore these parameters are classified as

sensitive and are included in the uncertainty analysis. Parameter sensitivity is poorly correlated with elevation, maximum SWE, precipitation, and length of melting period. Parameter sensitivity might be affected by multiple site characteristics, in which case it would be difficult to distinguish the effect of one characteristic from another. Furthermore, other site physiographics, including aspect and vegetation cover, may have a significant impact on watershed-scale snow accumulation and ablation rates. Their influence on parameter sensitivity may not be evident in our point-scale analysis, but may be more apparent, and have more influence, at the basin-scale.

### 3.2. Uncertainty analysis

#### 3.2.1. Prediction of best posterior parameter set

The SWE estimates produced from the best posterior parameter set (maximum likelihood) are compared to the observed SWE. The *RMSE*, *NSE*, and bias metrics are calculated for the 14-year study period at each site (Fig. 6). First, the *NSE* value is consistently greater than 0.8 (Fig. 6a), indicating that the SNOW17 model configured with the best parameter set accurately reproduces the SWE at all sites. Second, the normalized *RMSE* (*RMSE* divided by long-term annual mean snowfall) shows significant variability (Fig. 6b), with the highest value at site LL (CA) and the lowest value at site WH (AZ), sites that receive the highest and lowest annual precipitation, respectively. In addition, sites MC (WA), HM (OR), IL (CA), and LL (CA), located in regions that receive higher precipitation (Pacific Northwest and Sierra Nevada regions), show significantly higher *RMSE* values. Last, the percent bias calculated at each site indicates a generally good match between simulated and observed SWE, with the highest percent bias at −5.5% (sites MC (WA) and TD (WY)). SWE is consistently under-estimated at all sites (Fig. 6c). Since the best parameter set is derived by minimizing the *RMSE*, it is most likely that the under-catch in wet years would dominate the *RMSE* calculation (by over-weighting the over-catch in dry years) and thus result in overall underestimation at the SNOTEL sites.

To further verify the robustness of the optimal DREAM parameter set, we investigate the model prediction at the annual scale for three sites: LL (Leavitt Lake, CA), BL (Brumley, CO), and WH (White Horse Lake, AZ). Sites LL (CA), BL (CO), and WH (AZ) receive 1815 mm, 307 mm, and 113 mm snowfall on an annual basis, respectively. The annual average *RMSE* accounts for 8%, 7%, and 12% of the annual average snowfall amount at three sites, respectively (Fig. 7). This indicates that the prediction at wet sites LL (CA) and BL (CO) is relatively better than at the drier site, WH (AZ). Furthermore, for each site, there are several years where the *RMSE* value is evidently higher. For instance, at site LL (CA), the *RMSE* calculated for WY2006 is the highest (Fig. 7a), accounting for 18% of the total snowfall of that year. Scrutinizing the hyetograph (not shown) of this year shows that most precipitation occurs in the early winter. Precipitation in December and January accounts for 45% of the annual total precipitation and the peak precipitation (216 mm) occurs in early December. On January 2, 2006 (following a four-day intense precipitation event (396 mm) starting from December 30, 2005), the model simulated SWE only accounts for 80% of the observation due to erroneous classification of snowfall from the precipitation by the model. The bias starts to accumulate from that day forward. Consequently, the model-predicted peak is 20% lower than the observed peak and modeled SWE melts out 15 days earlier. The discrepancy in peak and melting patterns leads to higher *RMSE* values. In summary, the performance of the best parameter set varies from year to year and, as expected, is more satisfactory in normal years than in extreme years. In addition, the performance is not solely affected by snowfall totals. The temporal distribution of snowfall (variance from the normal pattern) also significantly influences model predictions.

#### 3.2.2. Prediction of posterior parameter sets

The performance of DREAM-derived posterior parameter sets is first evaluated using the containing ratio (*CR*). Second, the performance is compared to the prior parameter sets. Last, the performance is evaluated against that of the widely used SCE-UA global optimization algorithm. The median value of the SWE prediction ensembles is used in the comparison with the calibration algorithm. The *CR* values range from 0.62 (at TD (WY) and VL (CO)) to 0.76 (at LL (CA)), with a mean value of 0.69 (Fig. 8a). The *CR* is also highly variable from year to year (Fig. 8b). The highest *CR*

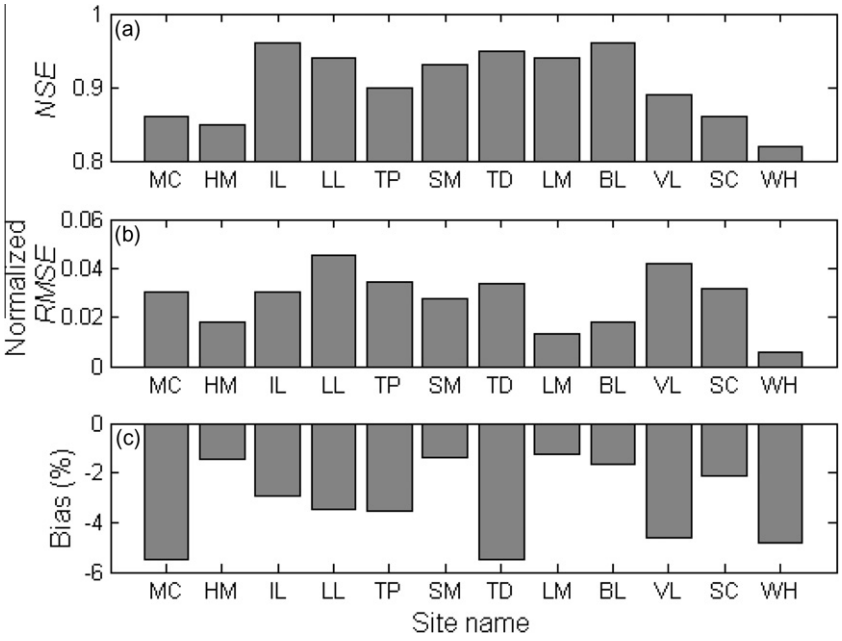


**Fig. 6.** Statistics of 14-year (WY1995-2008) SWE predictions simulated using the DREAM-derived optimal parameter set: (a) *NSE*, (b) normalized *RMSE* (*RMSE* divided by annual mean snowfall), and (c) percent bias.

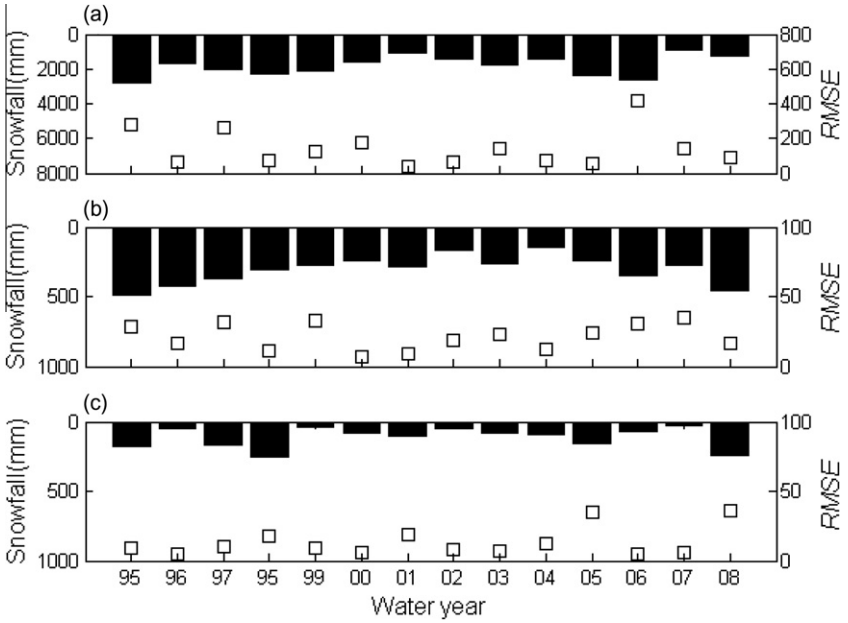


**Fig. 7.** Annual snowfall and *RMSE* using the optimal DREAM parameter set at the three selected sites: (a) Leavitt Lake (LL (CA)), (b) Brumley (BL (CO)), and (c) White Horse Lake (WH (AZ)). Bars designate snowfall and squares indicate *RMSE* (mm) values.

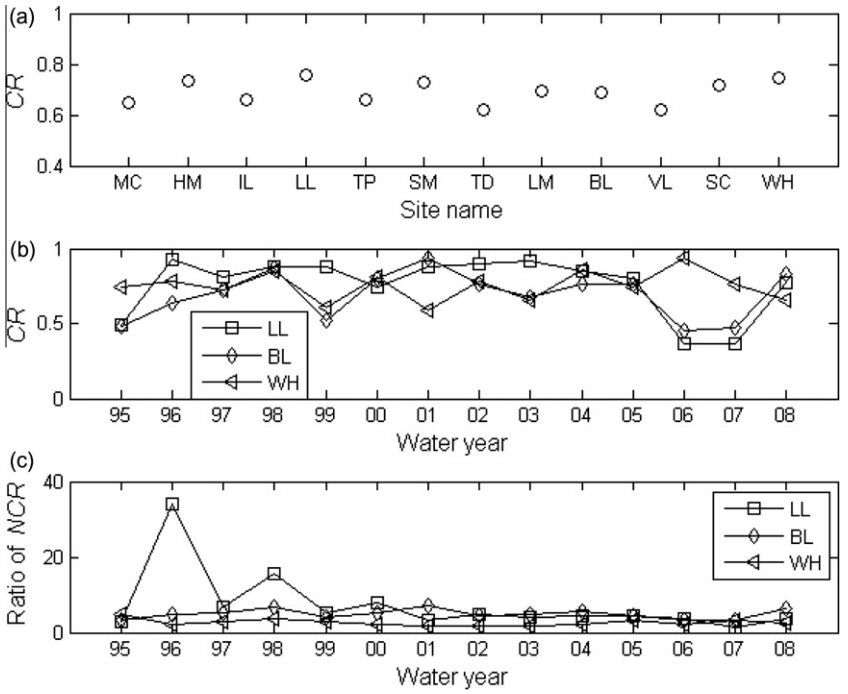


**Fig. 8.** The long-term (14-year average) containing ratio (*CR*) at (a) each of the 12 study sites and (b) *CR* at sites Leavitt Lake (LL (CA)), Brumley (BL (CO)), and White Horse Lake (WH (AZ)) for each of the study years, as well as (c) the ratio of the posterior *NCR* to the prior *NCR* ($NCR_{posterior}/NCR_{prior}$) at each of the three sites for each year.

values for sites LL (CA), BL (CO), and WH (AZ) are 0.93 (WY1996), 0.94 (WY2001), and 0.94 (WY2006), respectively. These years receive normal amount of precipitation, indicating excellent performance of the parameter sets in normal years. Correspondingly, the lowest *CR* values are 0.36 (WY2006), 0.45 (WY2006), and 0.59 (WY2001) for LL (CA), BL (CO), and WH (AZ), respectively. The low *CR* value for WY2006 at site LL (CA) is expected since, as described above, the model performs poorly in that year even with the optimal parameter set. Similarly, for site BL (CO), model performance in WY2007 is not satisfactory; under-estimation of SWE starting in early winter results in a low *CR* value. At site WH (AZ) there are two apparent accumulation–ablation periods (October to middle November and early January to late March) in WY2001 (not shown). The model fails to capture the first accumulation process (it misclassifies precipitation as rain since the mean air temperature of the process is greater than 2 °C) and subsequently the first melting process. The simulated SWE from the first period persists until the beginning of the next accumulation phase, leading to consistently higher SWE estimate in the second period. Overall, the SWE is over-estimated. The over-estimation likely re-

sults in low *CR* value at the site in WY2001. The annual mean *CR* values at the three sites are 0.76 (LL (CA)), 0.69 (BL (CO)), and 0.75 (WH (AZ)), respectively, suggesting acceptable long-term performance of the model-derived SWE ensembles at all three sites. However, the low *CR* value at site LL (CA) (0.36) illustrates that it is possible to produce SWE ensemble estimates without satisfactory coverage of observations. As discussed earlier, the skewed distribution of precipitation in WY2006 is most likely contributing to the low coverage of the observations.

To investigate the extent to which the posterior parameter sets outperform the prior, the coverage of observations associated with prior and posterior uncertainty intervals is calculated. The prior uncertainty interval (generally wider) would be expected to cover more observations. This is due to the fact that the prior parameters are generated from the entire parameter space while the posterior parameters are within a smaller region of the parameter space. To make the comparison meaningful, a metric reconciling both *CR* and the width of uncertainty interval (*NCR*, Eq. (5)), is applied. The ratio of the posterior *NCR* to the prior *NCR* for sites LL (CA), BL (CO), and WH (AZ) are evaluated (Fig. 8c). At the three selected sites, the posterior parameter ensemble outperforms the prior parameter ensemble (ratio greater than one) in all years. Specifically, the ratio at site LL (CA) in WY1996 reaches 34, a value indicating that the posterior parameter ensembles perform considerably better than the prior parameter ensembles.

Peak SWE estimates for three years from sites LL (CA), BL (CO), and WH (AZ) are also examined (Fig. 9). These years correspond to a relative wet year, normal year, and dry year with respect to snowfall received at each site. As expected, the uncertainty in the prior ensembles is consistently larger than the posterior ensembles. Comparing to the prior uncertainty bounds, the posterior bounds more adequately capture the SWE peaks. Furthermore, both the prior and posterior uncertainty bounds generally mimic the pattern of observed SWE. However, the posterior ensembles more closely and more appropriately match SWE variability. In addition, at the wet site (LL (CA)), some of the prior parameters provide non-zero SWE prediction in the summer due to the fact that the over-estimated peak SWE predictions take much longer to melt-out. Last, for all three years at site BL (CO), the melt-out dates of the prior predictions are significantly lagged in comparison to the SWE observations.

Overall, the median of the SWE prediction ensemble (produced from the DREAM-derived posterior parameters) more closely resembles the observed SWE than the prediction produced from SCE-UA derived optimal parameter set at three selected sites (Table 3). For sites BL (CO) and WH (AZ), the ensemble median outperforms the SCE-UA prediction in terms of the *NSE* and *RMSE* metrics investigated. For site LL (CA), in comparison with the ensemble median, the SCE-UA prediction has a higher *RMSE* (which indicates worse estimation on high SWE) and lower bias (which indicates smaller overall deviation from observations). This suggests that the SCE-UA predictions provide poorer estimation of lower SWE values as a result of the algorithm trying to minimize the overall least squared error between modeled and observed SWE (the objective function applied in the SCE-UA). Results further illustrate that the median of predictions derived from the prior parameters is generally poorer than the median posterior prediction and SCE-UA prediction (Table 3).

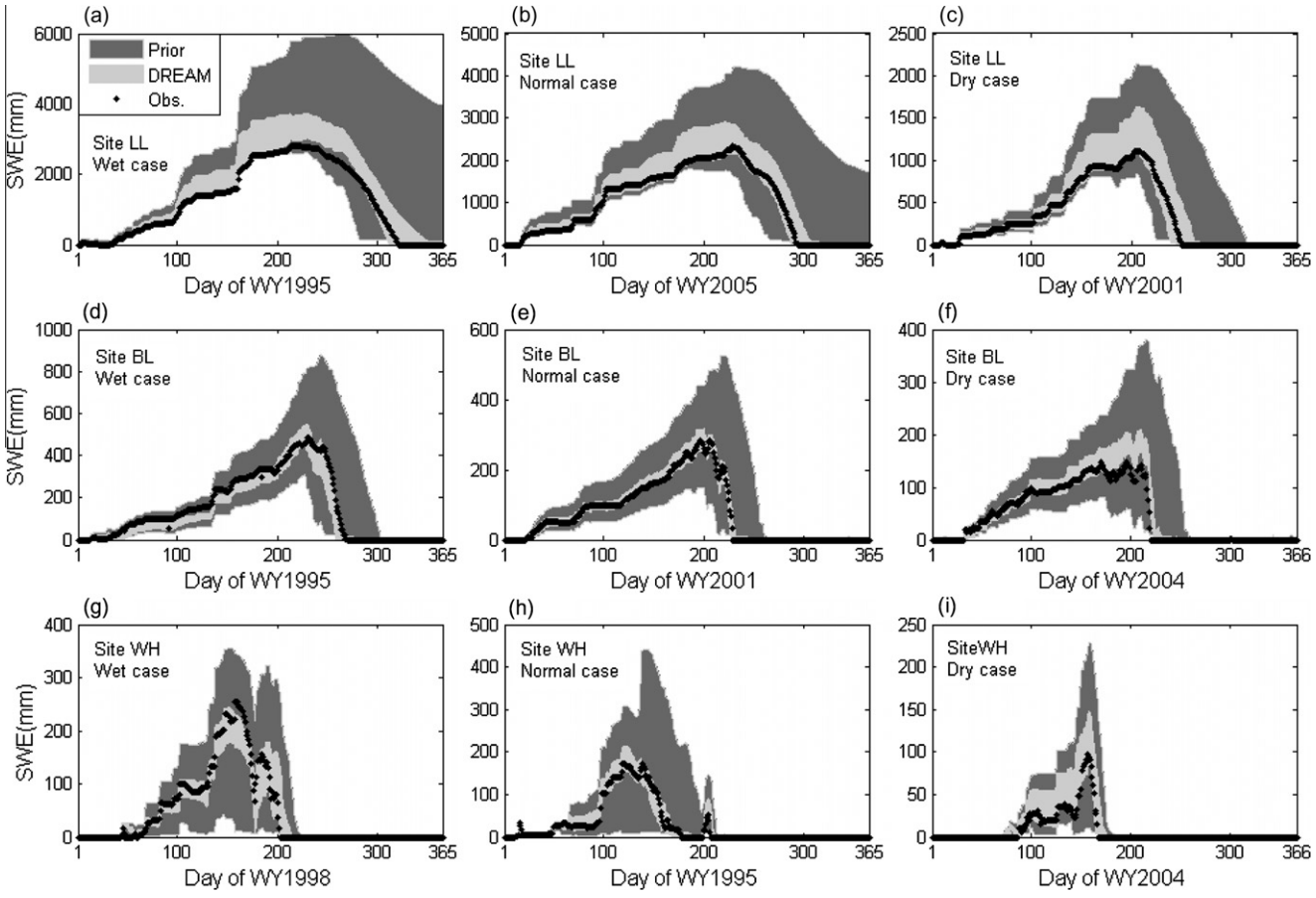


**Fig. 9.** Prior and posterior 95% SWE prediction uncertainty ranges (shaded regions) along with SWE observations (circles). The top, middle, and bottom rows show results for sites Leavitt Lake (LL (CA)), Brumley (BL (CO)), and White Horse Lake (WH (AZ)), respectively. The left, middle, and right columns illustrate examples of model performance for wet, normal, and dry year cases, respectively at each site. Note vertical scale differences at the three sites.

**Table 3**
Summary statistics of 14-year (WY1995–2008) SWE predictions for three selected SNOTEL sites using the median of DREAM-derived posterior parameter ensemble and the SCE-UA derived optimal parameter set. Cases 1 and 2 represent the median of prior and posterior SWE prediction, respectively. Case 3 represents the SCE-UA prediction.

| Site | Case | Statistics | | |
|---|---|---|---|---|
| | | *NSE* | *RMSE* (mm) | BIAS (%) |
| LL | Case 1 | −0.41 | 893.05 | 87.71 |
| | Case 2 | 0.94 | 177.65 | −3.61 |
| | Case 3 | 0.93 | 200.18 | −0.88 |
| BL | Case 1 | 0.76 | 55.16 | 24.48 |
| | Case 2 | 0.96 | 23.26 | −1.55 |
| | Case 3 | 0.94 | 26.64 | −6.61 |
| WH | Case 1 | 0.65 | 22.61 | 20.98 |
| | Case 2 | 0.81 | 16.51 | −5.50 |
| | Case 3 | 0.71 | 20.64 | −26.71 |

### 3.2.3. Uncertainty and correlation of posterior parameters

The correlation between posterior parameter ensemble width (uncertainty) for seven model parameters and various site characteristics is evaluated for all study sites (Fig. 10). The evaluation aims to identify potential regional relationships between parameter uncertainty and site hydroclimatic conditions. In general, there are only a few site/parameter combinations that show strong correlation. The uncertainty of PXTEMP is negatively correlated to site elevation (−0.91; *p*-value 0.00) and positively correlated to air temperature (0.75; *p*-value 0.01). Recall that PXTEMP (ranging from −2 °C to 2 °C) distinguishes snowfall from precipitation based on the criterion that precipitation is in the form of snowfall when air temperature is less than PXTEMP. More winter precipitation at higher elevation is more likely to be snowfall (air temperature less than −2 °C). In this case, PXTEMP has less impact on the determination of snow versus rain (i.e., most precipitation is in the form of snow), leading to smaller variations in its range. The uncertainty in parameter UADJ is found to be positively correlated to elevation (0.67; *p*-value 0.02) and negatively correlated to air temperature (−0.73; *p*-value 0.01). UADJ accounts for heat exchange due to turbulent transfer of heat during rain periods. It is likely that the larger uncertainty in UADJ with increasing elevation is due to increasing winds at higher elevation for which the constant UADJ cannot account. Last, a negative correlation is observed between the uncertainty in SCF and site precipitation (−0.61; *p*-value 0.04) and between SCF and length of the melting period (−0.61; *p*-value 0.03). As described previously, SCF is a mass correction factor. Recall that the posterior parameters are obtained by minimizing the *RMSE* of simulated and observed SWE. The more precipitation received at a site, the less variation is expected in SCF in order to produce SWE estimates deviating less from the observed SWE (smaller *RMSE*). Considering the fact that SCF functions in the accumulation period rather than the melting period, the

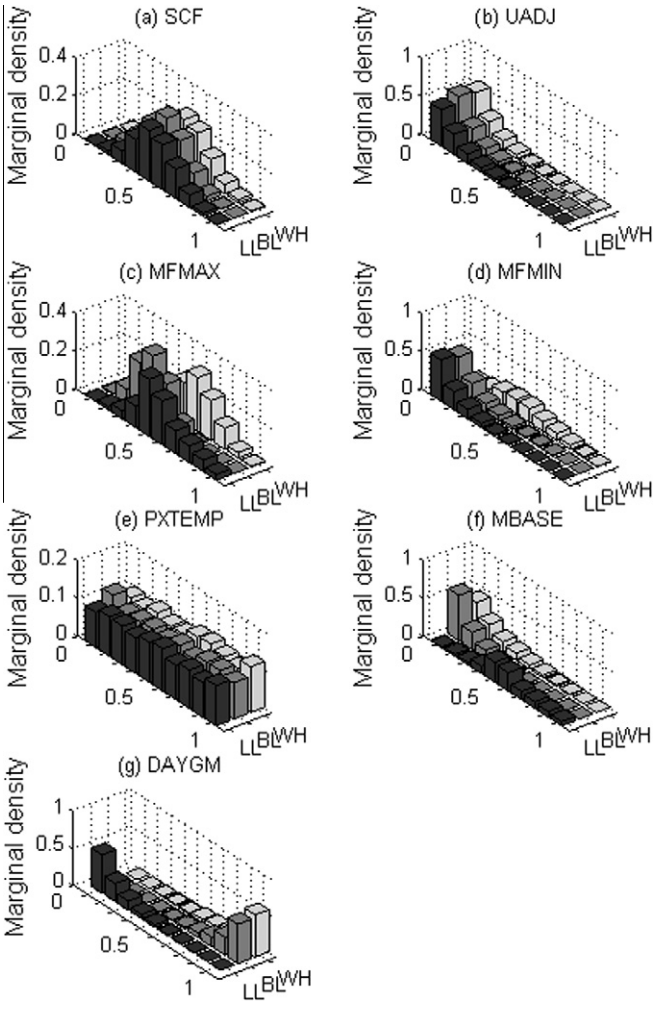


**Fig. 11.** Normalized posterior marginal parameter distributions at the three select sites: Leavitt Lake (LL (CA), dark gray), Brumley (BL (CO), medium gray), and White Horse Lake (WH (AZ), white).

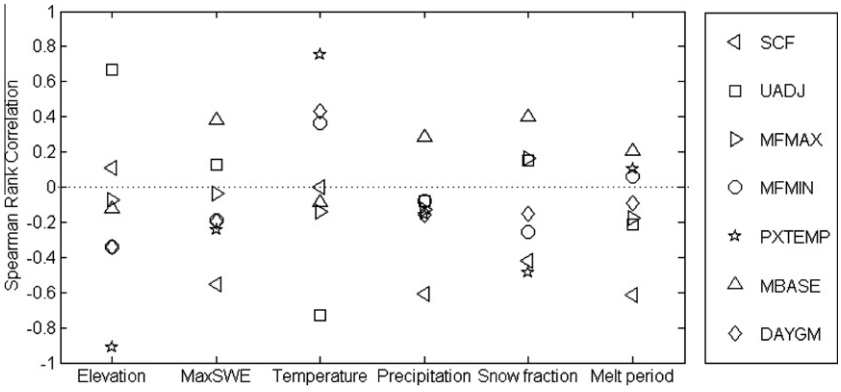


**Fig. 10.** Correlation between DREAM-derived posterior parameter ranges and select topographic and climatic characteristics associated with each study site for seven SNOW17 model parameters: SCF, UADJ, MFMAX, MFMIN, PXTEMP, MBASE, and DAYGM.

correlation between SCF and the length of melting period is most likely indirect, as the overall amount of precipitation is strongly related to the length of the melting period.

The normalized marginal distributions of the posterior parameters at sites LL (CA), BL (CO), and WH (AZ) are illustrated (Fig. 11). Parameter SCF (Fig. 11a) shows a normal distribution at all three sites. In addition, at each site, the peak of the distribution is in the middle part of the parameter range. Parameter MFMAX (Fig. 11c) is similar to that of SCF, with the exception that the distribution at site BL (CO) is skewed to the first half of parameter range. Parameters UADJ (Fig. 11b) and DAYGM (Fig. 11g) follow log-normal type distributions. Parameter PXTEMP (Fig. 11e) roughly follows a uniform distribution at all three sites. This finding implies that the prior and posterior marginal distribution of parameter PXTEMP follows a similar type, indicating that there is no significant high probability region of the parameter PXTEMP space. However, the posterior PXTEMP mean at site LL (CA) is constantly different from sites BL (CO) and WH (AZ) (Table 4). Parameters MFMIN (Fig. 11d) and MBASE (Fig. 11f) distributions appear to be site-dependent (log-normal or normal distributions). In general, most parameters (except for MFMIN and MBASE) have the same posterior distributions at different sites. In addition, for the same site, different parameters may follow different types of distributions, indicative of the complex form of the joint distribution of posterior parameters.

There are few significant correlations among most of the parameters at each of the three sites (Table 4). At site LL (CA), the correlation between parameters MFMAX and MBASE is 0.95. This is likely due to the fact that non-rain melt is the primary melt process at site LL (CA) and that the melt starts in the Spring at the site (on April 25, Table 2). Recall that parameters MFMAX, MFMIN, and MBASE control non-rain melt where MFMAX and MBASE dominate the calculation of melt amount after March 21. The high correlation between MFMAX and MBASE at the site is thus expected. At site WH (AZ), a correlation coefficient of 0.72 is found between parameters SCF and MFMIN. The $p$-values under both cases are much less than 0.05, indicating that the correlation is significant. Moderate correlations are seen between parameters SCF and MFMAX (−0.38) and SCF and MBASE (−0.55) at site LL (CA); between parameters SCF and MFMAX (0.44) and between MBASE and MFMAX (0.41) at site BL (CO); and between MFMIN and DAYGM (−0.33) at site WH (AZ).

## 4. Summary and conclusions

In the current study the sensitivity and uncertainty of SNOW17 model parameters are investigated for a set of SNOTEL sites with varying hydroclimatic characteristics using the commonly applied GSA and the advanced parameter uncertainty algorithm DREAM. The study demonstrates how different SNOW17 parametric sensitivity and uncertainty vary under different hydroclimatic (forcing) conditions. Results also show the regionalization information associated with the parametric sensitivity and uncertainty. In addition, the study has implications with regards to model behavior and merging observational data and the model via data assimilation techniques, which are addressed in future work. Findings from the current study are summarized as follows:

(1) Parameter sensitivity is highly variable from site to site. Parameters SCF, MFMAX, and PXTEMP are generally sensitive across the study sites, while the remaining parameters are only sensitive at a few of the study sites. Consistent sensitivity of SCF and PXTEMP is expected because they control snowfall input to the model. This observation highlights the importance of identifying and correcting precipitation bias (usually undercatch) as well as improving rain-snow partitioning in snow modeling, laying the foundation for improved modeling of accumulation processes and, subsequently, melting processes. Improved understanding of the accumulation phase also holds considerable potential for improving the areal extent of snow cover in distributed snow modeling, allowing better representation of the observed heterogeneity of watershed snowpack. MFMAX is sensitive because it is the key parameter determining non-rain melt which is the primary melt process in the Western US. Three parameters NMF, TIPM, and PLWHC are generally insensitive at all sites.

**Table 4**
Posterior mean and correlation structure between SNOW17 model parameters for three selected SNOTEL sites for the study period (WY1995–2008). Shading designates level of correlation: low (white), medium (light gray), high or significant (dark gray).

| Parameter | Posterior mean | SCF | UADJ | MFMAX | MFMIN | PXTEMP | MBASE | DAYGM |
|---|---|---|---|---|---|---|---|---|
| *Site Leavitt Lake (LL, CA)* | | | | | | | | |
| SCF | 0.72 | – | 0.00 | −0.38 | −0.03 | 0.00 | −0.55 | 0.07 |
| UADJ | 0.03 | – | – | 0.01 | 0.03 | −0.02 | 0.03 | 0.03 |
| MFMAX | 0.96 | – | – | – | 0.08 | −0.04 | 0.95 | 0.03 |
| MFMIN | 0.05 | – | – | – | – | 0.00 | 0.10 | 0.05 |
| PXTEMP | 1.95 | – | – | – | – | – | −0.04 | −0.05 |
| MBASE | 0.56 | – | – | – | – | – | – | 0.03 |
| DAYGM | 0.00 | – | – | – | – | – | – | – |
| *Site Brumley (BL, CO)* | | | | | | | | |
| SCF | 1.04 | – | −0.06 | 0.44 | 0.00 | 0.05 | −0.12 | 0.32 |
| UADJ | 0.03 | – | – | −0.11 | 0.02 | 0.00 | 0.02 | −0.02 |
| MFMAX | 1.18 | – | – | – | 0.05 | 0.02 | 0.41 | −0.01 |
| MFMIN | 0.05 | – | – | – | – | −0.02 | 0.09 | 0.00 |
| PXTEMP | 1.05 | – | – | – | – | – | −0.01 | 0.03 |
| MBASE | 0.01 | – | – | – | – | – | – | −0.05 |
| DAYGM | 0.30 | – | – | – | – | – | – | – |
| *Site White Horse Lake (WH, AZ)* | | | | | | | | |
| SCF | 1.06 | – | 0.05 | 0.19 | 0.72 | 0.03 | −0.06 | 0.10 |
| UADJ | 0.03 | – | – | 0.01 | 0.04 | 0.01 | −0.01 | −0.02 |
| MFMAX | 0.66 | – | – | – | −0.15 | 0.01 | 0.20 | 0.05 |
| MFMIN | 0.07 | – | – | – | – | 0.00 | 0.01 | −0.33 |
| PXTEMP | 1.05 | – | – | – | – | – | 0.00 | 0.04 |
| MBASE | 0.02 | – | – | – | – | – | – | 0.04 |
| DAYGM | 0.29 | – | – | – | – | – | – | – |

(2) The 95% SWE prediction uncertainty ranges corresponding to the posterior parameter sets contains, on average, 75% of the observations in respect to both peak and variation (pattern) at three selected sites with contrasting wetness. In addition, the median posterior simulation outperforms the prediction of the widely used deterministic calibration algorithm (SCE-UA), confirming the robustness of DREAM in estimating parameter and model predictive uncertainty.
(3) There is generally low correlation between SNOW17 parametric sensitivity/uncertainty and most site characteristics investigated. This shows that the regionalization information is very limited, and hence it will be difficult to identify useful empirical relationships for complex snow models to extrapolate parametric sensitivity and uncertainty to ungaged areas. Our results suggest that to accurately identify sensitivity and uncertainty patterns of model parameters for a given site and study period, the coupled sensitivity and uncertainty analysis should be conducted for that case specifically.
(4) For the sites studied, the mean posterior parameter values change from site to site. The inter-correlation among posterior parameters also varies with different study sites, although there are only a few pairs of parameters at each site which have significantly strong correlation. The marginal distributions of most model parameters have similar distribution types. The derived correlation structure and distributions provide critical information on formulating a realistic and accurate data assimilation framework for a range of operational forecasting locations across the western US. Science-based estimates of parameter distribution should also help improve probabilistic ensemble streamflow predictions, which the NWS relies on to produce long-term streamflow outlooks [9,14].

It is worth noting that our analyses do not explicitly address uncertainty in forcing (except for those represented by SCF and PXTEMP uncertainty), observations, and model structure. These different sources of uncertainty and their affect on regionalization will be addressed in due course. In addition, due to the limited availability of SNOTEL data on a time scale finer than a day, simulations are conducted using a 24-h time step only. Application of SNOW17 to analyze diurnal variations could potentially yield different results and hence requires further investigation.

## Acknowledgements

The authors thank Dr. Pedro Restrepo, Dr. Dong-Jun Seo, and Dr. Yuqiong Liu for their suggestions and advice on the development of this work. This work is partially supported by a grant from the National Oceanic Atmospheric Administration (NOAA) National Weather Service (NA07NWS4620013) and by a UCLA Graduate Division Fellowship. The work of the last author is supported by a J. Robert Oppenheimer Fellowship from the Los Alamos National Laboratory Postdoctoral Program.

## References

[1] Ajami NK, Duan Q, Sorooshian S. An integrated hydrologic Bayesian multimodel combination framework: confronting input, parameter, and model structural uncertainty in hydrologic prediction. Water Resour Res 2007;43:W01403.
[2] Anderson EA. Calibration of conceptual models for use in river forecasting. Hydrology Lab., Silver Spring, MD: National Weather Service; 2002.
[3] Anderson EA. National weather service river forecast system – snow accumulation and ablation model. Tech. memo NWS HYDRO-17. Silver Springs, Md.: National Oceanographic and Atmospheric Administration; 1973. p. 2491.
[4] Beck MB. Water quality modeling: a review of the analysis of uncertainty. Water Resour Res 1987;23:1393.
[5] Beven K, Binley A. Future of distributed models: model calibration and uncertainty prediction. Hydrol Process 1992;6:279.
[6] Blasone RS, Vrugt JA, Madsen H, Rosbjerg D, Robinson BA, Zyvoloski GA. Generalized likelihood uncertainty estimation (GLUE) using adaptive Markov Chain Monte Carlo sampling. Adv Water Resour 2008;31:630.
[7] Christiaens K, Feyen J. Use of sensitivity and uncertainty measures in distributed hydrological modeling with an application to the MIKE SHE model. Water Resour Res 2002;38:1169.
[8] Crook AG. SNOTEL: monitoring climatic factors to predict water supplies. J Soil Water Conserv 1977;32.
[9] Day GN. Extended streamflow forecasting using NWSRFS. J Water Res Pl-ASCE 1985;111:157.
[10] Doesken NJ, Schaefer GL. The contribution of SNOTEL precipitation measurements to climate analysis, monitoring and research. In: Western snow conference, Vancouver, B.C., Canada; 1987. p. 14.
[11] Duan Q, Sorooshian S, Gupta V. Effective and efficient global optimization for conceptual rainfall–runoff models. Water Resour Res 1992;28:1015.
[12] Fassnacht SR, Dressler KA, Bales RC. Snow water equivalent interpolation for the Colorado River basin from snow telemetry (SNOTEL) data. Water Resour Res 2003;39:1208.
[13] Franz KJ. Characterization of the comparative skill of conceptual and physically-based snow models for streamflow prediction. Irvine: University of California; 2006. p. 223.
[14] Franz KJ, Hartmann HC, Sorooshian S, Bales R. Verification of national weather service ensemble streamflow predictions for water supply forecasting in the Colorado River basin. J Hydrometeorol 2003;4:1105.
[15] Franz KJ, Hogue TS, Sorooshian S. Operational snow modeling: addressing the challenges of an energy balance model for National Weather Service forecasts. J Hydrol 2008;360:48.
[16] Franz KJ, Hogue TS, Sorooshian S. Snow model verification using ensemble prediction and operational benchmarks. J Hydrometeorol 2008;9:1402.
[17] Freer J, Beven K, Ambroise B. Bayesian estimation of uncertainty in runoff prediction and the value of data: an application of the GLUE approach. Water Resour Res 1996;32:2161.
[18] Garbrecht JD. Comparison of three alternative ANN designs for monthly rainfall–runoff simulation. J Hydraul Eng ASCE 2006;11:502.
[19] Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences. Stat Sci 1992;7:457.
[20] Hamby DM. A review of techniques for parameter sensitivity analysis of environmental models. Environ Monit Assess 1994;32:135.
[21] He M. Data assimilation in watershed models for improved hydrologic forecasting. Los Angeles: University of California; 2010. p. 172.
[22] Hogue TS, Bastidas LA, Gupta HV, Sorooshian S. Evaluating model performance and parameter behavior for varying levels of land surface model complexity. Water Resour Res 2006;42:W08430.
[23] Hornberger GM, Spear RC. An approach to the preliminary analysis of environmental systems. J Environ Manage 1981;12:7.
[24] Kavetski D, Kuczera G, Franks SW. Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory. Water Resour Res 2006;42:W03407.
[25] Kitanidis PK, Bras RL. Real-time forecasting with a conceptual hydrologic model I. Analysis of uncertainty. Water Resour Res 1980;16:1025.
[26] Koren V, Reed S, Smith M, Zhang Z, Seo DJ. Hydrology laboratory research modeling system (HL-RMS) of the US National Weather Service. J Hydrol 2004;291:297.
[27] Kottegoda NT, Rosso R. Statistics, probability, and reliability for civil and environmental engineers. New York: McGraw-Hill; 1997.
[28] Kuczera G, Parent E. Monte Carlo assessment of parameter uncertainty in conceptual catchment models: the Metropolis algorithm. J Hydrol 1998;211:69.
[29] Liu Y, Gupta HV. Uncertainty in hydrologic modeling: toward an integrated data assimilation framework. Water Resour Res 2007;43:W07401.
[30] Madsen H. Automatic calibration of a conceptual rainfall–runoff model using multiple objectives. J Hydrol 2000;235:276.
[31] Margulis SA, McLaughlin D, Entekhabi D, Dunne S. Land data assimilation and estimation of soil moisture using measurements from the Southern Great Plains 1997 field experiment. Water Resour Res 2002;42:W01407.
[32] McKay MD. Sensitivity and uncertainty analysis using a statistical sample of input values. Boca Raton, Fla: CRC Press; 1988. pp. 145–86.
[33] Montanari A. What do we mean by'uncertainty'? The need for a consistent wording about uncertainty assessment in hydrology. Hydrol Process 2007;21:841.
[34] Mote PW. Trends in snow water equivalent in the Pacific Northwest and their climatic causes. Geophys Res Lett 2003;30:1601.
[35] Mote PW, Hamlet AF, Clark MP, Lettenmaier DP. Declining mountain snowpack in Western North America. Bull Am Meteorol Soc 2005;86:39.
[36] Muleta MK, Nicklow JW. Sensitivity and uncertainty analysis coupled with automatic calibration for a distributed watershed model. J Hydrol 2005;306:127.
[37] Nash JE, Sutcliffe JV. River flow forecasting through conceptual models: Part I. A discussion of principles. J Hydrol 1970;10:282.
[38] NRCS. SNOTEL data collection system. Natural Resources Conservation Service; 1997.
[39] NWS. National Weather Service River forecast system – user's manual. Silver Spring, MD.: Hydrology Lab, National Weather Service; 2004.

[40] Reed S, Koren V, Smith M, Zhang Z, Moreda F, Seo D-J, et al. Overall distributed model intercomparison project results. J Hydrol 2004;298:27.
[41] Seo DJ, Koren V, Cajina N. Real-time variational assimilation of hydrologic and hydrometeorological data into operational hydrologic forecasting. J Hydrometeorol 2003;4:627.
[42] Serreze MC, Clark MP, Armstrong RL, McGinnis DA, Pulwarty RS. Characteristics of the western United States snowpack from snowpack telemetry (SNOTEL) data. Water Resour Res 1999;35:2145.
[43] Serreze MC, Clark MP, Frei A. Characteristics of large snowfall events in the montane western United States as examined using snowpack telemetry (SNOTEL) data. Water Resour Res 2001;37:675.
[44] Service RF. As the west goes dry. Science 2004;303:1124.
[45] Shamir E, Carpenter TM, Fickenscher P, Georgakakos KP, Asce M. Evaluation of the National Weather Service operational hydrologic model and forecasts for the American River basin. J Hydrol Eng 2006;11:392.
[46] Shamir E, Georgakakos KP. Distributed snow accumulation and ablation modeling in the American River basin. Adv Water Resour 2006;29:558.
[47] Shamir E, Georgakakos KP. Estimating snow depletion curves for American River basins using distributed snow modeling. J Hydrol 2007;334:162.
[48] Sieber A, Uhlenbrook S. Sensitivity analyses of a distributed catchment model to verify the model structure. J Hydrol 2005;310:216.
[49] Singh P, Singh VP. Snow and glacier hydrology. The Netherlands: Kluwer Academic Publishers; 2001.
[50] Slater AG, Clark MP. Snow data assimilation via an ensemble Kalman filter. J Hydrometeorol 2006;7:478.
[51] Smith MB, Seo D-J, Koren VI, Reed SM, Zhang Z, Duan Q, et al. The distributed model intercomparison project (DMIP): motivation and experiment design. J Hydrol 2004;298:4.
[52] Spear RC, Grieb TM, Shang N. Parameter uncertainty and interaction in complex environmental models. Water Resour Res 1994;30:3159.
[53] Spear RC, Hornberger GM. Eutrophication in peel inlet – II. Identification of critical uncertainties via generalized sensitivity analysis. Water Res 1980;14:43.
[54] Stewart IT, Cayan DR, Dettinger MD. Changes in snowmelt runoff timing in Western North America under a 'business as usual' climate change scenario. Climatic Change 2004;62:217.
[55] Tang Y, Reed P, van Werkhoven K, Wagener T. Advancing the identification and evaluation of distributed rainfall–runoff models using global sensitivity analysis. Water Resour Res 2007;43:W06415.
[56] Tang Y, Reed P, Wagener T, Van Werkhoven K. Comparing sensitivity analysis methods to advance lumped watershed model identification and evaluation. Hydrol Earth Syst Sci 2007;11:793.
[57] Thiemann M, Trosset M, Gupta H, Sorooshian S. Bayesian recursive parameter estimation for hydrologic models. Water Resour Res 2001;37:2521.
[58] van Werkhoven K, Wagener T, Reed P, Tang Y. Characterization of watershed model behavior across a hydroclimatic gradient. Water Resour Res 2008;44:W01429.
[59] Vrugt JA, Bouten W, Gupta HV, Sorooshian S. Toward improved identifiability of hydrologic model parameters: the information content of experimental data. Water Resour Res 2002;38:1312.
[60] Vrugt JA, Gupta HV, Bouten W, Sorooshian S. A shuffled complex evolution metropolis algorithm for optimization and uncertainty assessment of hydrological model parameters. Water Resour Res 2003;39:1201.
[61] Vrugt JA, ter Braak C, Diks CGH, Roberts DA, Hyman JM, Higdon D. Accelerating Markov Chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling. Int J Nonlinear Sci Numer Simulat 2009;10:273.
[62] Vrugt JA, ter Braak C, Gupta H, Robertson D. Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling? Stoch Environ Res Risk Assess 2008. doi:10.1007/s00477-008-0274-y.
[63] Vrugt JA, ter Braak CJF, Clark MP, Hyman JM, Robinson BA. Treatment of input uncertainty in hydrologic modeling: doing hydrology backwards with Markov chain Monte Carlo simulation. Water Resour Res 2008;44:W00B09.
[64] Wagener T, Boyle DP, Lees MJ, Wheater HS, Gupta HV, Sorooshian S. A framework for development and application of hydrological models. Hydrol Earth Syst Sci 2001;5:13.
[65] Wagener T, Kollat J. Numerical and visual evaluation of hydrological and environmental models using the Monte Carlo analysis toolbox. Environ Modell Softw 2007;22:1021.
[66] Wagener T, McIntyre N, Lees MJ, Wheater HS, Gupta HV. Towards reduced uncertainty in conceptual rainfall–runoff modelling: dynamic identifiability analysis. Hydrol Process 2003;17:455.
[67] Washington WM, Weatherly JW, Meehl GA, Semtner Jr AJ, Bettge TW, Craig AP, et al. Parallel climate model (PCM) control and transient simulations. Clim Dynam 2000;16:755.
[68] Xiong L, O'Connor KM. An empirical method to improve the prediction limits of the GLUE methodology in rainfall–runoff modeling. J Hydrol 2008;349:115.
[69] Zhang Y, Pinder G. Latin hypercube lattice sample selection strategy for correlated random hydraulic conductivity fields. Water Resour Res 2003;39:1226.