



Exploring snow model parameter sensitivity using Sobol' variance decomposition



Elizabeth S. Houle^a, Ben Livneh^{a, b}, Joseph R. Kasprzyk^{a, *}

^a Civil, Environmental, and Architectural Engineering, University of Colorado Boulder, Boulder, CO, USA

^b Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, CO, USA

ARTICLE INFO

Article history:

Received 7 June 2016

Received in revised form

20 November 2016

Accepted 23 November 2016

Available online 30 December 2016

Keywords:

Snow hydrology

Parameter sensitivity

Snow modeling

Snow water equivalent

Model performance

Sobol' sensitivity analysis

ABSTRACT

This study advances model diagnostics for snowmelt-based hydrological systems using Sobol' sensitivity analysis, illuminating parameter sensitivities and contrasting model structural differences. We consider several distinct snow-dominated locations in the western United States, running both SNOW-17, a conceptual degree-day model, and the Variable Infiltration Capacity (VIC) snow model, a physically-based model. Model performance is rigorously evaluated through global sensitivity analysis and a temperature warming analysis is conducted to explore how model parameterizations affect portrayals of climate change. Both VIC and SNOW-17 produce comparable results with SNOW-17 performing slightly better for shallower snowpacks and VIC performing better for deeper snowpacks. However, the lack of sensitivity of SNOW-17 to climate warming suggests that it may not be as reliable as a more sensitive model like VIC. Inter-model differences presented here offer insights into physical features with greatest uncertainty and may inform future model development and planning activities.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Between 50% and 80% of annual runoff in snowmelt-dominated rivers in the western United States (U.S.) is derived from snowmelt (Stewart et al., 2004); these systems are especially sensitive to climatic variations that affect the spring snowpack (Barnett et al., 2005; Mote et al., 2005). In this context, numerical snow models provide insight on how snow water equivalent (SWE) estimation is impacted by these changes. However, it is critical to understand a model's structure and parametric sensitivity (i.e., model diagnostics) in order to properly interpret model depiction of hydrologic processes (Gan et al., 2014; Sarrazin et al., 2016) and how the processes change under climatic variations.

There are two broad categories of forward-running, continuous snow models: conceptual and physically-based. Conceptual models represent the physical processes in a simplified and sometimes indirect form (e.g. a degree-day snow model), whereas physically-based models represent physical processes by solving energy and mass balance equations. Several snow model comparison studies have demonstrated that model complexity is not directly equitable

with the fidelity of simulation results (Essery et al., 2013; Etchevers et al., 2004). Thus, it is informative to study both conceptual and physically-based models in a model diagnostic exercise. The main lens through which to study the model performance uses quantified objective functions that compare model simulations with the observed data, assuming that we can learn something about proper values of the model parameters by gauging how well the parameterization reproduces observed data. Our approach to model evaluation supplements this analysis by using multiple objective functions that capture different aspects of the model response, supplementing this with qualitative observation of visualizations of performance (Wagener, 2003). In this paper, we explore parameter sensitivities of two widely-used snow models: the conceptual model, SNOW-17 (Anderson, 1973), and the snow model within the physically-based Variable Infiltration Capacity (VIC) model (Liang et al., 1994).

The parameters of conceptual models do not always have a direct relationship to physical properties of hydrologic systems. Thus, conceptual models such as SNOW-17 are designed to be calibrated before they are used for predictions. Previous studies have used a variety of methods to calibrate SNOW-17 parameters, both with automatic calibration on objective functions associated with the SWE response (Leisenring and Moradkhani, 2011), and manual adjustment of some of the model parameters (Franz et al.,

* Corresponding author. Civil Environmental and Architectural Engineering Department, University of Colorado Boulder, UCB 607, Boulder, CO, 80309, USA.

E-mail address: joseph.kasprzyk@colorado.edu (J.R. Kasprzyk).

2008) Furthermore, there is a need to address the parameter sensitivity of these conceptual models to gain further insight into how the parameter sets affect the robustness of model predictions. For example, He et al. (2011a) used global sensitivity and uncertainty analysis methods to explore the parameter space and sensitivity of SNOW-17 model parameters, across 12 point locations with varying climate regimes. They determined the same set of model parameters remains sensitive among all site locations, but sensitivities vary in magnitude among sites.

Studies that perform diagnostics on physically-based Land Surface Models (LSMs), such as VIC, have been more limited. Demaria et al. (2007) used a Monte Carlo framework to evaluate the sensitivity and identifiability of 10 parameters controlling surface and subsurface runoff generation in VIC. They concluded that modeled streamflow is sensitive to three parameters, allowing for a more parsimonious reduced set of model parameters. Raje and Krishnan (2012) extended this analysis with Markov Chain Monte Carlo methods to model uncertainty and eventually project future discharge with respect to climate change. They concluded that uncertainty resulting from a general circulation model (GCM) is larger than that from parameter uncertainties in VIC. Of particular interest to this work is a comparative analysis of snow processes within LSMs (Feng et al., 2008), which conducted simple parameter sensitivity analyses of snow processes within LSMs including the Simplified Simple Biosphere Model (Xue et al., 1991), VIC (Liang et al., 1994), and others. In their investigation of the impact of snow model complexity on simulation performance, they conducted an albedo sensitivity analysis, but only changed the fresh snow albedo parameter within VIC and ignored albedo decay parameters. There has been limited additional research to date with respect to the sensitivity of physically-based snow model parameters, specifically parameters controlling uncertain quantities that are rarely observed, namely liquid water holding capacity (Denoth, 2003), new snow density, snow roughness, new snow albedo, and albedo decay rates.

This study employs the Sobol' sensitivity analysis method (Sobol'), a global, variance-based sensitivity assessment method that varies all of the model's parameters in predefined regions of the parameter space within upper and lower bounds, to quantify the amount that each parameter contributes to the variance of the output (Sobol', 1993, 1967). Global sensitivity analysis methods are sophisticated diagnostic tools that are especially useful for exploring parameter sensitivities (Borgonovo, 2006; Neumann, 2012; Norton, 2015; Pianosi et al., 2016; Saltelli et al., 2004, 1999; Tian, 2013) within complex models. Tang et al. (2007) compared several local and global sensitivity analysis methods applied to the Sacramento Soil Moisture Accounting Model coupled with SNOW-17 and determined the Sobol' sensitivity method provides robust and detailed sensitivity rankings. Additional studies employed Sobol' to determine model parameter sensitivities (Butler et al., 2014; Confalonieri et al., 2010; Cosenza et al., 2013; Fieberg and Jenkins, 2005; Ganji et al., 2016; Hall et al., 2005; Kontoravdi et al., 2005; Nossent et al., 2011; Rosolem et al., 2012; Van Werkhoven et al., 2008; Vezzaro and Mikkelsen, 2012).

This paper seeks to contribute the first combined Sobol' sensitivity analysis comparing physically-based and conceptual snow models. This complements prior work in snow model investigations, specifically showing sensitivity indices that suggest the relative contribution of each parameter to the model output variance. Our investigation explores VIC snow model parameters which are typically fixed in modeling studies. The paper is a complement to prior published work in that we look at model behavior at multiple sites as well as across multiple model structures (physically-based vs. conceptual). Finally, we present an analysis of how the models' estimations of SWE changes are modulated by the

parameter sensitivities under a temperature change scenario. Overall, our analysis here seeks to contribute a holistic model diagnostic exercise that can aid users of snow models. Specifically, the analysis has the potential to inform future model calibration exercises¹ by potentially fixing some insensitive parameters and guiding users on the most important model parameters.

2. Methodology

2.1. Data

Five sites from the U.S. Natural Resources Conservation Service SNOW TElemetry (SNOTEL) system, located in the headwaters of major U.S. watersheds and representing a range of snow-driven climates in the western U.S., are selected for this research. Table 1 describes the five SNOTEL locations, which include CSS Lab, Squaw Valley G.C., Harts Pass, Copper Mountain, and Galena Summit, their climatologies, and the water years used in the analysis. Fig. 1 depicts a map of the SNOTEL locations, including 64-year average annual precipitation, from Livneh et al. (2015), across the western U.S. Different water years are selected for Harts Pass due to long periods of missing data during the water years used for the other sites.

Model simulations are forced by daily precipitation, air temperature, and wind speed, and are validated against SWE data. As previously mentioned, precipitation, air temperature, and SWE data are from SNOTEL. Daily wind data are from the North American Research Re-analysis, using the grid cell closest to the study sites from a wind field re-sampled to 1/16° data (Livneh et al., 2013).

Two quality-related issues with the data are present. First, inspection reveals approximately 2% of the temperature data are missing or faulty. In an approach from Livneh et al. (2014), two data points, one before and one after the missing or faulty data, are used in a linear interpolation to replace the missing or faulty data. The second data-related issue stems from the fact that obtaining reliable precipitation estimates at high elevations can be difficult due to wind redistribution processes in snow-covered catchments, which affect snow accumulation. These redistribution processes can lead to a phenomenon termed under-catch of frozen precipitation. Precipitation undercatch results from wind-induced updrafts at the gauge orifice and wetting losses on the internal walls of the gauge (Groisman and Legates, 1994) and can significantly affect the accuracy of precipitation observations (Rasmussen et al., 2012). One potential result of undercatch is that a dataset could contain less precipitation than SWE, which is physically impossible. As a result, a straightforward procedure outlined by Livneh et al. (2014) is applied to adjust precipitation data to ensure it is always equal-to or greater-than observed SWE. In summary, an algorithm was applied to estimate true precipitation and prevent erroneous estimates during the cold season whereby the precipitation gauge or snow pillow may be prone to errors. In essence it uses a seven-day moving window in which accumulated gauge-measured precipitation is compared with increases in SWE from the snow pillow.

2.2. SNOW-17 model

SNOW-17 is a conceptual snow accumulation and ablation model first developed by Anderson (1973) as part of the U.S. National Weather Service (NWS) River Forecast System for use in river

¹ For more information regarding model calibration, see Houle (2015), which describes some multi-objective calibration of the models in this paper.

Table 1

Characteristics of the geography and climate of SNOTEL sites selected for this study.

SNOTEL location (site number)	State	Basin	Lat.	Long.	Elevation (m)	Mean annual precipitation (mm)	Water years	Climate	Climate description
Copper Mountain (415)	CO	Upper Colorado River	39.49	−106.17	3216	734	1998–2002	Continental	- Shallow, less dense snowpack - Less frequent storms with periods of drought - Colder temperatures
CSS Lab (428)	CA	Truckee River	39.33	−120.37	2089	1704	1998–2002	Maritime	- Deep, high density snowpack - Mild regime
Galena Summit (490)	ID	Salmon Basin	43.87	−114.71	2767	579	1998–2002	Inter-Mountain	- Intermediate snowpack depth - Mild temperatures - Fewer storms and rain than maritime climates
Harts Pass (515)	WA	Upper Columbia	48.72	−120.66	1978	1132	2000–2003	Maritime	- Deep, high density snowpack - Frequent storms - Mild regime
Squaw Valley G.C. (784)	CA	Truckee River	30.19	−120.26	2447	1656	1998–2002	Maritime	Same as CSS

forecasting. The model description in this section is adapted from the most recent reference for the model: [Anderson \(2006\)](#). The NWS has used SNOW-17 for decades to model snow accumulation and ablation in snow-dominated regions to create short- and long-lead stream flow forecasts. SNOW-17 is an index model that uses average daily air temperature as the only index to simulate heat storage of the snowpack, snowmelt, liquid water retention, and melt water transmission to determine the energy exchange across the snow-air interface with empirically-based relationships ([He et al., 2011a](#)). It also requires temperature and precipitation as input variables, and it outputs SWE and runoff timeseries.

This analysis focuses on eight of ten parameters that govern SNOW-17 performance, discussed further in [Table 2](#)²: snow correction factor (SCF), maximum negative melt factor (NMF), antecedent temperature index (TIPM), temperature above which melt typically occurs (MBASE), minimum melt factor on December 21st (MFMIN), maximum melt factor on June 21st (MFMAX), average wind function during rain-on-snow events (UADJ), and percent liquid water holding capacity of a well-aged snow cover (PLWHC). Following prior studies ([Anderson, 2006](#); [He et al., 2011a](#)), we exclude two parameters: PXTEMP, which controls the rain/snow temperature partitioning and DAYGM, ground heat flux. The existence of snow measurement error as an artifact of wind speed, gauge type, and site location, is well documented ([Rasmussen et al., 2012](#)). As a result, SNOW-17 includes the variable SCF as a multiplicative factor that adjusts the amount of newly fallen snow to account for gauge under-catch errors before the new snow is added to the existing snow cover using the following equation

$$P_n = P \cdot f_s \cdot SCF \quad (1)$$

where P_n is the water equivalent of new snowfall (mm), P is the total precipitation input to the model (mm), and f_s is the fraction of precipitation in the form of snow. SCF is an average value over all of the accumulation periods and is typically used to calibrate the model. The value is usually chosen to give a best estimate of the amount of water in snow cover at the beginning of the melt season ([Anderson, 2006](#)).

Most important physical processes that take place within the snowpack are explicitly included within SNOW-17, but in simplified form, using temperature as an index. For example, Equation (2)

describes melt during non-rain periods in the conceptual temperature-index SNOW-17 model. Melt occurs when the air temperature (T_a), is above a base value (MBASE); otherwise melt (M) does not occur. The SNOW-17 melt equation is given by

$$M = M_f(T_a - MBASE) \quad (2)$$

where M is the depth of melt (mm), M_f is the seasonally varying melt factor ($\text{mm}/^\circ\text{C}$), T_a is the air temperature ($^\circ\text{C}$), and MBASE is the temperature threshold above which melt will not occur ($^\circ\text{C}$).

2.3. Variable Infiltration Capacity (VIC) snow model

The VIC model is a physically-based, distributed, hydrologic model that requires minimum and maximum daily temperature, precipitation, and wind speed as inputs while providing a wide range of outputs, including SWE, streamflow, evapotranspiration, and water and energy balance terms. In this section, we limit the description of VIC to its relevant snow processes investigated in the study; readers can consult [Liang et al. \(1994\)](#) for more information on the model. This analysis focuses on eight parameters that are suspected to govern VIC snow model performance, discussed further in [Table 3](#): new snow density (ρ), snow roughness (η), liquid water holding capacity (λ), new snow albedo (α), two accumulation period albedo decay parameters (\exp_{A1} , \exp_{B1}), and two thaw period albedo decay parameters (\exp_{A2} , \exp_{B2}). The VIC snow model simulates snow in several forms: ground snow pack, snow in vegetation canopy, and snow on top of lake ice ([Cherkauer et al., 2003](#)). Only ground snowpack is considered in this study because data are collected in clearings without canopy cover. The VIC ground snowpack consists of two layers—a thin surface layer and a thick deep pack layer—both for which VIC solves the full mass and energy balances. Energy exchange between the atmosphere and snowpack occurs only within the surface snow layer. The total energy available for refreezing liquid water or melting snowpack over a given time step depends on the net energy exchange between the ground, snow, and atmosphere. Net radiation at the snow surface is calculated given incoming shortwave and longwave radiation. The Mountain Microclimate Simulation Model (MT-CLIM) version 4.3 ([Thornton et al., 2000](#)) was employed as a pre-processor within VIC to estimate radiation and vapor terms from the diurnal temperature range and precipitation occurrence.

Albedo governs solar radiation reflection and thus dominates the snowpack energy balance during the melt season, which typically occurs between April and July in the western U.S. VIC assumes snow albedo decays with time since last snowfall based on the U.S.

² Note that the shortened parameter names used here follow typical modeling practice and are not necessarily abbreviations.

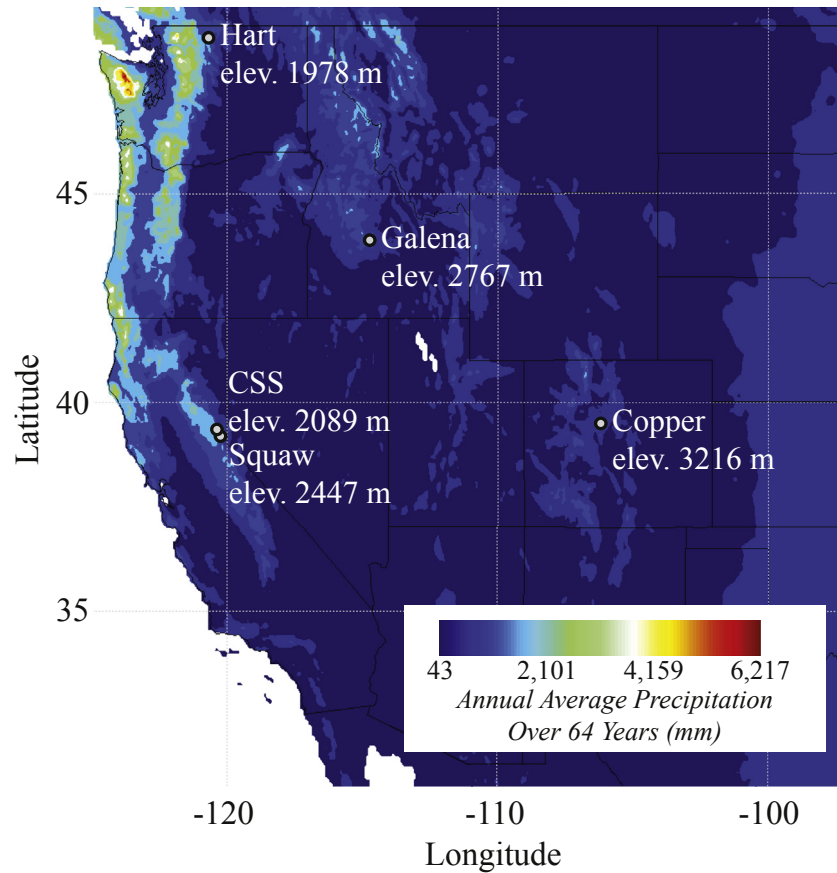


Fig. 1. Map depicting the SNOTEL study site locations across the western United States. The colors represent annual average precipitation over a 64-year period. Precipitation data is from Livneh et al. (2015). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2

SNOW-17 parameter descriptions, default values, and ranges used in the Sobol' sensitivity analysis. Values for ranges were adapted from the associated references.

Parameter	Description	Units	Default value	Range Min. Max.	Reference
SCF ^a	snow correction factor; accounts for gage catch deficiencies and losses during accumulation periods, which controls snowfall input into the model	—	1	0.70 1.40	He et al. (2011a)
UADJ	average wind function during rain-on-snow events; controls melt during rain-on-snow events	mm/mb/6 h	0.04	0.03 0.20	He et al. (2011a)
MBASE	temperature above which melt typically occurs; controls snowmelt during non-rain periods	°C	1.00	0.00 2.00	Anderson (1973)
MFMAX	maximum melt factor (June 21st); used to calculate melt factor which controls snowmelt during non-rain periods	mm/°C/6 h	1.05	0.50 2.00	He et al. (2011a)
MFMIN	minimum melt factor (Dec. 21st); used to calculate melt factor which controls snowmelt during non-rain periods	mm/°C/6 h	0.60	0.05 0.50	He et al. (2011a)
TIPM	antecedent temperature index; controls heat exchange during non-melt period	—	0.10	0.05 1.00	He et al. (2011a)
NMF	maximum negative melt factor; controls heat exchange during non-rain periods	mm/°C/6 h	0.15	0.05 0.50	He et al. (2011a)
PLWHC	percent liquid water holding capacity of a well-aged snow cover; the amount of water a snowpack can hold before output occurs	decimal percent	0.04	0.02 0.30	He et al. (2011a)

^a SCF does not have a real default value and is usually chosen to give the best estimate of the amount of water in snow cover at the beginning of the melt season. It represents an average value over all of the accumulation periods.

Army Corps of Engineers (1956) formulation. It uses two separate sets of coefficients for the accumulation and melt seasons, categorized based on the absence and presence of liquid water in the surface snow layer. If the cold content is less than 0, the melt season parameterization is used, otherwise the accumulation parameterization is used. For each season the snow albedo from the decay rate is parameterized as

$$\alpha_{\text{snow}} = \alpha_{\text{max}} \exp_A^{t^{\exp_B}} \quad (3)$$

where α_{max} is the albedo of fresh fallen snow, t is the number of days since last snowfall, and the default values for \exp_A and \exp_B are listed in Table 3.

Fig. 2 depicts how the albedo decay curves can change with

Table 3

VIC parameter descriptions, default values, and ranges used in the Sobol' sensitivity analysis. Values for ranges were adapted from the associated references. The albedo decay parameters do not have listed references because the ranges were determined in this research. There are two sets of ranges for the decay parameters because we executed two sensitivity analysis experiments. The initial experiment allowed the ranges to be as large as possible, whereas the secondary experiment explored whether limiting the snow albedo to its most likely values would greatly alter the sensitivity results.

Parameter	Description	Units	Default value	Range		Reference
				Min.	Max.	
ρ	new snow density	kg/m ³	50	25	150	McClung and Schaerer (2006)
η	snow roughness	m	0.03	0.001	0.03	Singh and Singh (2001)
λ	liquid water holding capacity	decimal percent	0.035	0.02	0.15	Singh and Singh (2001)
α_{\max}	new snow albedo	—	0.85	0.80	0.90	Singh and Singh (2001)
exp _{A1}	accumulation period albedo decay parameter	—	0.94	0.30	0.99	—
				0.82	0.99	
exp _{B1}	accumulation period albedo decay parameter	—	0.58	0.30	0.99	—
				0.00	0.46	
exp _{A2}	thaw period albedo decay parameter	—	0.82	0.10	0.99	—
				0.82	0.99	
exp _{B2}	thaw period albedo decay parameter	—	0.46	0.10	0.99	—
				0.00	0.46	

different parameter values selected here, resulting in snow albedo ranging from the albedo of fresh fallen snow to approximately 0, which is unrealistic. According to Wiscombe and Warren (1980), the shortwave albedo of snow can fall from 0.90 to 0.50 or less,

depending on the snow condition. Albedo parameterizations are expected to play a major role in model performance because albedo is a significant factor for calculating snowmelt (Etchevers et al., 2004).

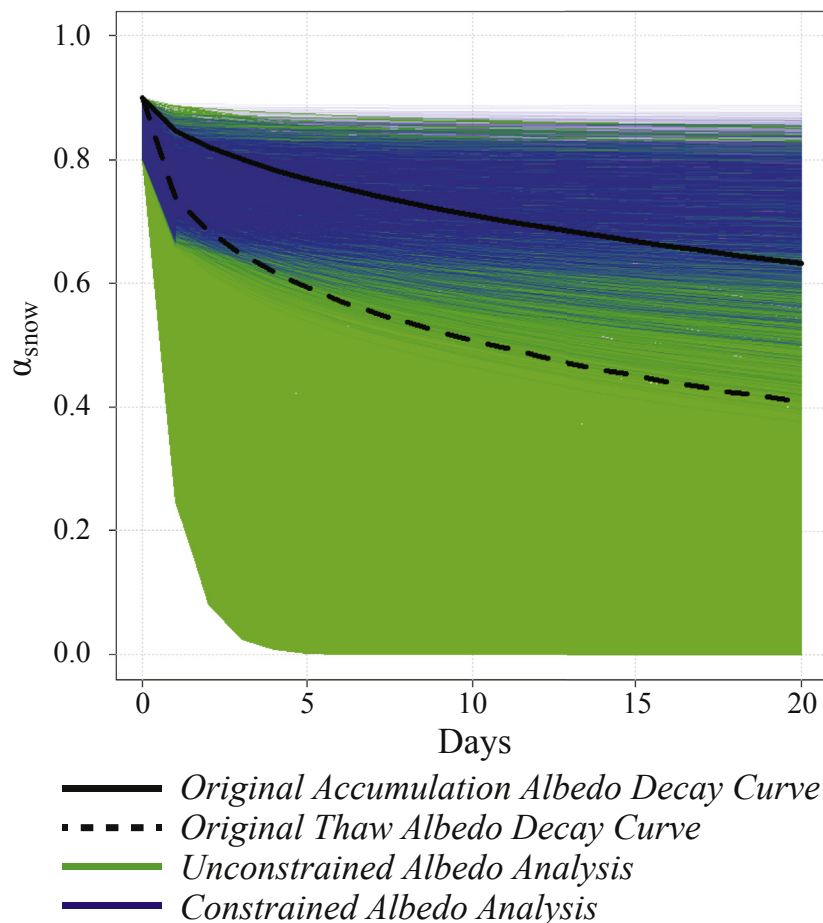


Fig. 2. Sensitivity of albedo decay curves when varying parameter values within the VIC snow model. The green area shows the range of albedo decay curves produced from the original Sobol' albedo decay curve parameter ranges (unconstrained albedo analysis). The blue area shows the range of albedo decay curves produced by narrowing the Sobol' ranges of albedo decay curve parameters to eliminate unrealistic albedo values (constrained albedo analysis). The black lines are the original accumulation and thaw albedo decay curves hard-coded into VIC. We run the constrained albedo analysis for a single site location to explore the impact of changing parameter ranges on Sobol' results and ensure parameter sensitivities are not significantly affected by this range. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.4. Model parameters

Tables 2 and 3 summarize all parameters included in the sensitivity analysis, their default values, ranges used in the analysis, and references for parameter ranges. Feasible ranges for each parameter within SNOW-17 are based on Anderson (2006) and previous studies using SNOW-17 (Franz et al., 2008; He et al., 2011a; Hogue et al., 2000). Feasible ranges for the ρ , η , λ , and α parameters within VIC are based on typical physical values found in literature (McClung and Schaerer, 2006; Singh and Singh, 2001).

Snow albedo decay parameters do not have published ranges. However, it is widely accepted that snow albedo tends to decay from a maximum value immediately following fresh snowfall to a lower value based on surface metamorphism of the snow surface during snow-free periods (e.g., Wiscombe and Warren, 1980). Therefore, the range of values for the \exp_{A1} , \exp_{B1} , \exp_{A2} , and \exp_{B2} parameters are based on the physical plausibility of their resultant albedo decay curve values (Fig. 2). In the initial experimentation, which we term the “unconstrained analysis,” the ranges were as large as possible (i.e., albedo had to be larger than zero), whereas a more limited secondary experiment, termed the “constrained analysis”, explored whether limiting snow albedo to its most likely value between approximately 0.35 and 0.9 would greatly alter sensitivity results.

For the unconstrained analysis, albedo decay parameter values were selected to allow a large range of residual albedo values after a relatively long lag time (t) of 20 days since the last snowfall. Given that snow albedo of less than 0.5 might be considered unrealistic, we explore the implications of this in a constrained albedo analysis over the CSS site, creating an additional set of decay coefficients that ensure snow albedos generally stay within the 0.35–0.9 range. The exact constraint was determined by using the minimum albedo value (0.36) that could result from the lowest possible initial albedo ($\alpha_{\max} = 0.8$) after a decay of t of 20 days, representing a low cutoff for possible albedo values.

2.5. Baseline model simulations

In order to facilitate comparisons between new model parameterizations and a default run of the model, we create a baseline model simulation for each site. Tables 2 and 3 provide baseline parameter values for all parameters in this study. The baseline VIC parameterization is taken from prior literature as discussed in the table. For SNOW-17, the value of SCF is typically chosen to give the best estimate of SWE at the beginning of the melt season (Anderson, 2006).

2.6. Objective functions and Sobol' sensitivity analysis

Model performance is evaluated using four objective functions that calculate the difference between observed and simulated SWE timeseries. Table 4 describes the objective functions that include: root mean square error (RMSE); Nash-Sutcliffe Efficiency (NSE); difference in peak SWE magnitude (PEAK); and difference in number of days that snow is on the ground, the snow cover duration (SCD). Each water year is considered separately for each objective function calculation. The worst performing value for each objective function out of the water years is selected as the representative objective value for the entire model run (a min-max approach) (Zeff et al., 2014). This method eliminates the possibility of obtaining a good objective value for a model run that performs very well for some water years, but poorly for others, and it ensures a model run is as good or better than its objective value for the worst performing water year.

The Sobol' sensitivity analysis method varies all of the model's

parameters in predefined regions to quantify the amount that each of the k parameters contributes to the variance of the output, $V(Y)$, where the output, Y , is the objective function matching the modeled SWE to the observed SWE. This method is based on variance decomposition

$$V(Y) = \sum_{i=1}^k V_i + \sum_{i < j}^k V_{ij} + \dots + V_{1,\dots,k} \quad (4)$$

where V_i is the partial variance representing the first-order effect of parameter X_i on the model output Y ; V_{ij} is the partial variance representing the second-order effect of interactions between parameters X_i and X_j ; $V_{1,\dots,k}$ is the interaction effect of all k parameters. Contributions to variance can be caused by variations in one parameter (first-order effects) or interactions of parameters with one or more other parameters. Sensitivity indices are therefore assigned to the parameters based on their relative contribution to the overall variance. For more information on the method see Sobol' (1993) and Saltelli et al. (2008). The total effects are calculated via the following equation:

$$S_{Ti} = \frac{V_i}{V} + \sum_{j \neq i} \frac{V_{ij}}{V} + \dots \quad (5)$$

in which the total order sensitivity for the i th parameter is the sum of the first order sensitivity, the second order interactions, and all other interactive effects that include this parameter. Sensitivities using equation (5) are calculated for both the SNOW-17 and VIC parameters.

The Sobol' variance decomposition method requires an ensemble of model parameterizations to be created for each model, and we employ a particular type of cross-sampling that creates a total of $2N(p+1)$ model simulations, where N is the sample size, and p is the number of parameters (Saltelli, 2002) to perform all calculations. Our analysis uses a sample size of 1,000, which results in a total of 18,000 model simulations for an analysis with eight parameters.

2.7. Temperature change analysis

The second part of this study involves applying a temperature change scenario at each site for both models to determine how errors resulting from model parameterizations propagate to portrayals of climate warming. Under future climate scenarios, GCMs consistently produce an increase in temperature, but differ in the magnitude of increase. However, most projections suggest that the variance of precipitation worldwide will increase in future climate change scenarios (Dore, 2005). In some locations in the U.S., one projection might suggest an increase in precipitation while another projection at the same location will suggest a decrease, so averaging precipitation outputs results in these varying predictions canceling each other out. As a result, we apply the simple delta-change method (Hay et al., 2000) to only downscaled temperature data to find the sensitivity of the models' projections exclusively to temperature change. This method involves computing the differences between the historical observed temperatures and future simulated temperatures and adding these differences to the observed timeseries. In order to obtain these delta-change values in our analysis, we (1) calculate the future average monthly temperature using five water years (2050 through 2054), (2) calculate the difference or delta-change between the future average monthly temperatures and historical average monthly temperatures obtained from the years in this analysis, (3) add the monthly delta-change values to the historical daily temperatures to create a

Table 4
Objective functions used to quantify model performance for the Sobol' sensitivity analysis.

Objective	Equation	Description
RMSE ^{a,b} (mm)	$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (O_t - Y_t)^2}$	Statistical metric that ranges from 0 to ∞ . The closer the value is to 0, the better the model performance.
NSE ^{b,c}	$NSE = 1 - \frac{\sum_{t=1}^n (O_t - Y_t)^2}{\sum_{t=1}^n (O_t - \bar{O}_t)^2}$	Statistical metric based on the least squares method. Values can range from $-\infty$ to 1, and an efficiency of 1 corresponds to a perfect match of modeled to observed SWE. Values are helpful because they can be compared across sites (Nash and Sutcliffe, 1970).
PEAK ^{a,b,d} (mm)	$PEAK = \max(O_t) - \max(Y_t) $	Difference in maximum SWE accumulation per season between the modeled and observed. The closer the value is to 0, the better the model performance.
SCD ^{a,b,e} (days)	$SCD = \text{days}(O_t > 1 \text{ mm}) - \text{days}(Y_t > 1 \text{ mm}) $	Difference in number of days when snow is on the ground ($> 1 \text{ mm}$) between the modeled and observed (Mizukami et al., 2013). The closer the value is to 0, the better the model performance.

^a O_t = observed SWE (mm/day).

^b Y_t = simulated SWE (mm/day).

^c \bar{O}_t = mean observed SWE (mm/day).

^d Max = maximum.

^e days = number of days.

daily timeseries of future temperatures. The minimum delta-change applied to the temperature data is 1.4 °C, and the maximum delta-change applied to temperatures is 2.4 °C.

Projected temperature data come from the World Climate Research Programme's Working Group on Coupled Modeling (Taylor et al., 2012) coordinated GCM experiments, representing the fifth phase of the Coupled Model Inter-comparison Project (CMIP5). CMIP5 produced outputs for four different climate scenarios or representative concentration pathways consisting of RCP2.6, RCP4.5, RCP6.0, and RCP8.5 (RCP8.5 is the trajectory we have been on since 2006) that provide a rough estimate of the anticipated range of radiative forcing by the year 2100. Given the coarse GCM spatial scales (e.g. ~ 100 km), downscaling techniques have been developed for projection of temperature and precipitation to regional and local scales (Crane and Hewitson, 1998; Wilby and Wigley, 1997). CMIP5 monthly climate and hydrology projections, downscaled using the Bias Corrected Statistical Disaggregation (BCSD) method, are available at 1/8° spatial resolution translations of climate projections, over the contiguous U.S. (Wood et al., 2004). We use downscaled monthly temperature data for the RCP4.5, a mid-range emissions scenario, at each site location to perform the temperature change analysis portion of this research.

3. Results and discussion

3.1. Baseline results

Fig. 3 presents baseline SWE results for one study location (Copper) using default parameter values for both models. Baseline results demonstrate that simulated SWE has some noteworthy deviations from observed SWE. For example, in Fig. 3, the VIC simulation at Copper for WY 1999 has an unusual decrease in SWE that is not present in the observations and does not appear in simulated SNOW 17 SWE. The explanation for the discrepancy is that SNOW-17 uses daily mean temperature as an input while VIC uses daily minimum and maximum temperatures as inputs. VIC estimates solar radiation based on these minimum and maximum values, so a large diurnal temperature range results in VIC over-estimating solar radiation for that particular time period.

Table 5 summarizes the objective function values for both the baseline model runs as well as the Sobol' ensembles described in the next section. Recall that our objective function values reported in the table are based on the water year with the worst match between simulated and observed values. For example, in Fig. 3, WY 2000 has the largest deviation, which yields an NSE value for VIC of 0.59. Baseline objective values for CSS, Squaw, and Hart perform better for VIC in contrast to Copper and Galena, which perform better for SNOW-17 (Copper's PEAK notwithstanding). This

suggests that VIC better simulates deeper snowpacks while SNOW-17 simulates shallower snowpacks with greater fidelity using baseline parameter values. According to Table 5, NSE values for all baseline model runs range from -0.08 to 0.89 , indicating there is significant room for improvement across both models as demonstrated in Fig. 3.

3.2. Ensemble of model results

The next step in analyzing model performance is to view the overall results from the Sobol' ensemble, to determine the range of model behavior based on parameter uncertainty. The ensembles shown in Fig. 4 depict the range of SWE outputs produced from the Sobol' parameterizations for SNOW-17 and VIC. Fig. 4 (a), (d), (g), (j), and (m) show that including SCF in the ensemble can produce modeled SWE that exceeds accumulated precipitation. However, Fig. 4 (b), (e), (h), (k), and (n) show that when setting SCF to a fixed value and eliminating it from the parameter ensemble, SNOW-17 does not produce SWE that exceeds accumulated precipitation. Recall that the SCF parameter is included in the model to account for gauge catch deficiencies, but it also implicitly includes losses during the snow accumulation period due to sublimation and redistribution of blowing snow. Our results with a fixed value of the SCF parameter show that improper values of it can be responsible for unrealistic model results. Empirical models, such as SNOW-17, often use correction factors such as SCF to overcome uncertainty in model forcings, and our findings on its effect are corroborated with prior research (He et al., 2011b). As a result, the sensitivity analysis and temperature change experiments in subsequent sections use a fixed value of 1 for the SCF parameter.

Like in the SNOW-17 ensembles, some SWE results in the VIC Hart and Squaw ensembles, depicted in Fig. 4 (l) and (o), respectively, exceed accumulated precipitation. This behavior is the result of snow not completely melting between water years, i.e. glaciation, which typically occurs in areas that have the combination of high precipitation, low energy availability (either radiation or temperature) and high albedo. We note that Harts Pass and Squaw Valley are in regions where small alpine glaciers are common and suspect that errors in meteorological data at these sites in specific years are responsible for erroneous glaciation. For example, temperatures could be too cold, or the derived shortwave radiation could be too low. We do not suspect that this is an issue with the albedo being too high in VIC, or parametric settings being askew in SNOW17.

In the initial VIC parameterization ensemble, we include albedo decay parameter values that produce unrealistic albedo ranging from approximately 0 to 0.9. In reality, snow albedo ranges from approximately 0.5 to 0.9. To explore the impact of this assumption,

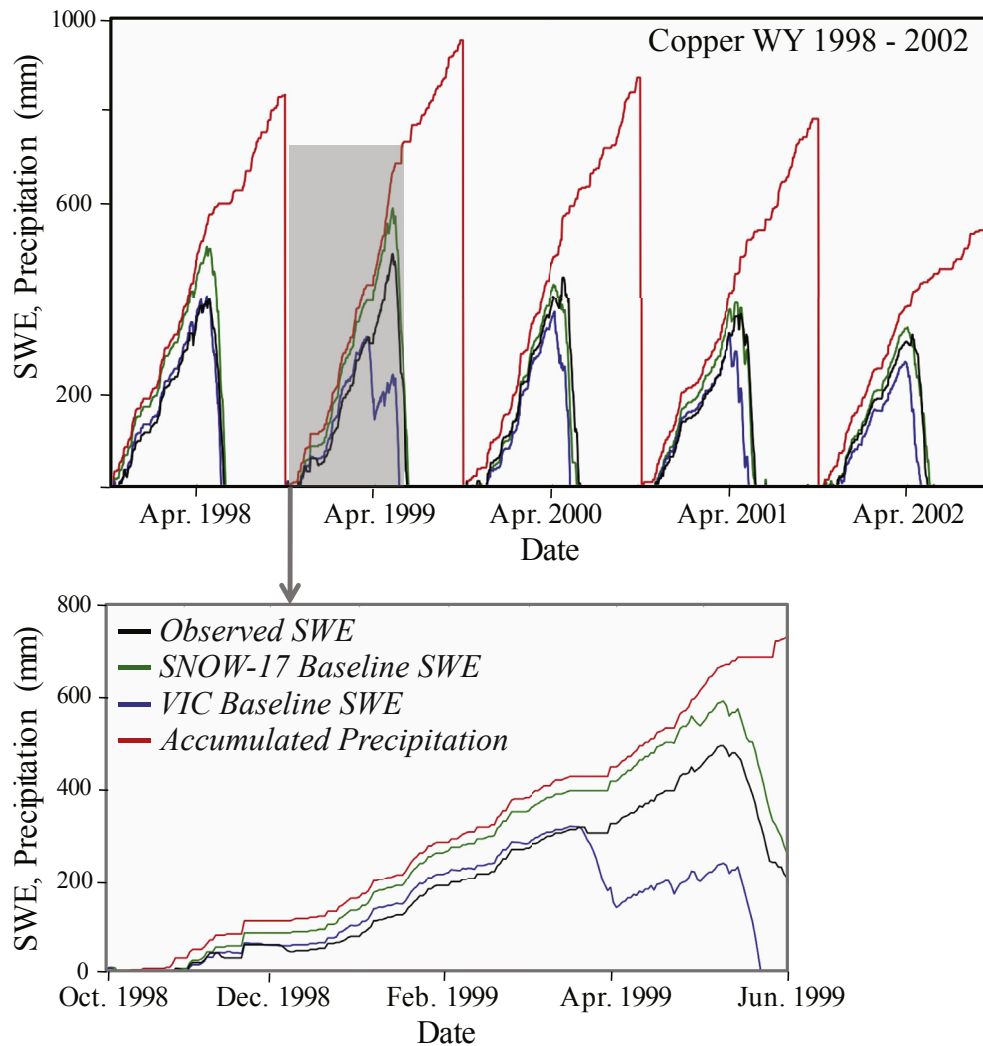


Fig. 3. Baseline SWE outputs for SNOW-17 and VIC for one site location (Copper). These outputs are produced using default model parameter sets (Tables 3 and 4). The only parameter without a default value is the SNOW-17 SCF parameter, which is fixed at a value of 1. Green lines on the plots represent SNOW-17 baseline SWE; blue lines represent VIC baseline SWE; black lines represent observed SWE; red lines represent adjusted accumulated precipitation, which is discussed in Section 2.1. Water year 1999 is zoomed in to better visualize the discrepancies between the modeled and observed SWE. All site locations and water years have varying degrees of differences between the modeled and observed SWE. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

we select a single site location (CSS) to run a follow-up analysis in which we limit the albedo parameter ranges to those that produce more realistic albedo values. Further investigation shows VIC SWE ensemble produced for CSS in the constrained albedo analysis is narrower than that produced in the unconstrained albedo analysis, but still indicates the same parameters affecting SWE output; this will be explored further in Fig. 7 below. VIC ensemble results demonstrate that SWE fidelity is dependent on fixed, potentially sub-optimal parameter values that may achieve good SWE simulations despite inappropriately calibrated parameters. Further exploration of model runs also reveals adjustment of VIC snow model parameters can correct the unrealistic melting of SWE that occurred in Copper's second water year in the baseline results.

3.3. Sobol' sensitivity results

The next step is to explore the sensitivity index results (i.e., which parameters are most responsible for producing a high variance in model output with respect to quantitative objective functions). Raster plots in Figs. 5–7 use colored squares to summarize

Table 5
Best objective function values for baseline and Sobol' results for SNOW-17 (S-17) and the VIC model.

Site	Model run	PEAK (mm)		SCD (days)		NSE		RMSE (mm)	
		S-17	VIC	S-17	VIC	S-17	VIC	S-17	VIC
Copper	Baseline	113	100	10	31	0.81	0.59	59	97
	Sobol'	56	44	8	3	0.94	0.95	48	33
CSS	Baseline	425	1381	28	24	−0.08	0.29	365	249
	Sobol'	264	587	7	12	0.82	0.86	124	103
	(unconstrained) Sobol' (constrained)	—	163	—	15	—	0.88	—	129
Galena	Baseline	58	254	18	44	0.88	0.57	47	110
	Sobol'	49	182	10	7	0.93	0.95	39	46
Hart	Baseline	425	1352	18	11	0.52	0.77	335	232
	Sobol'	311	1073	7	9	0.85	0.82	168	171
Squaw	Baseline	303	1692	18	6	0.84	0.89	327	193
	Sobol'	155	1182	13	2	0.93	0.97	184	124

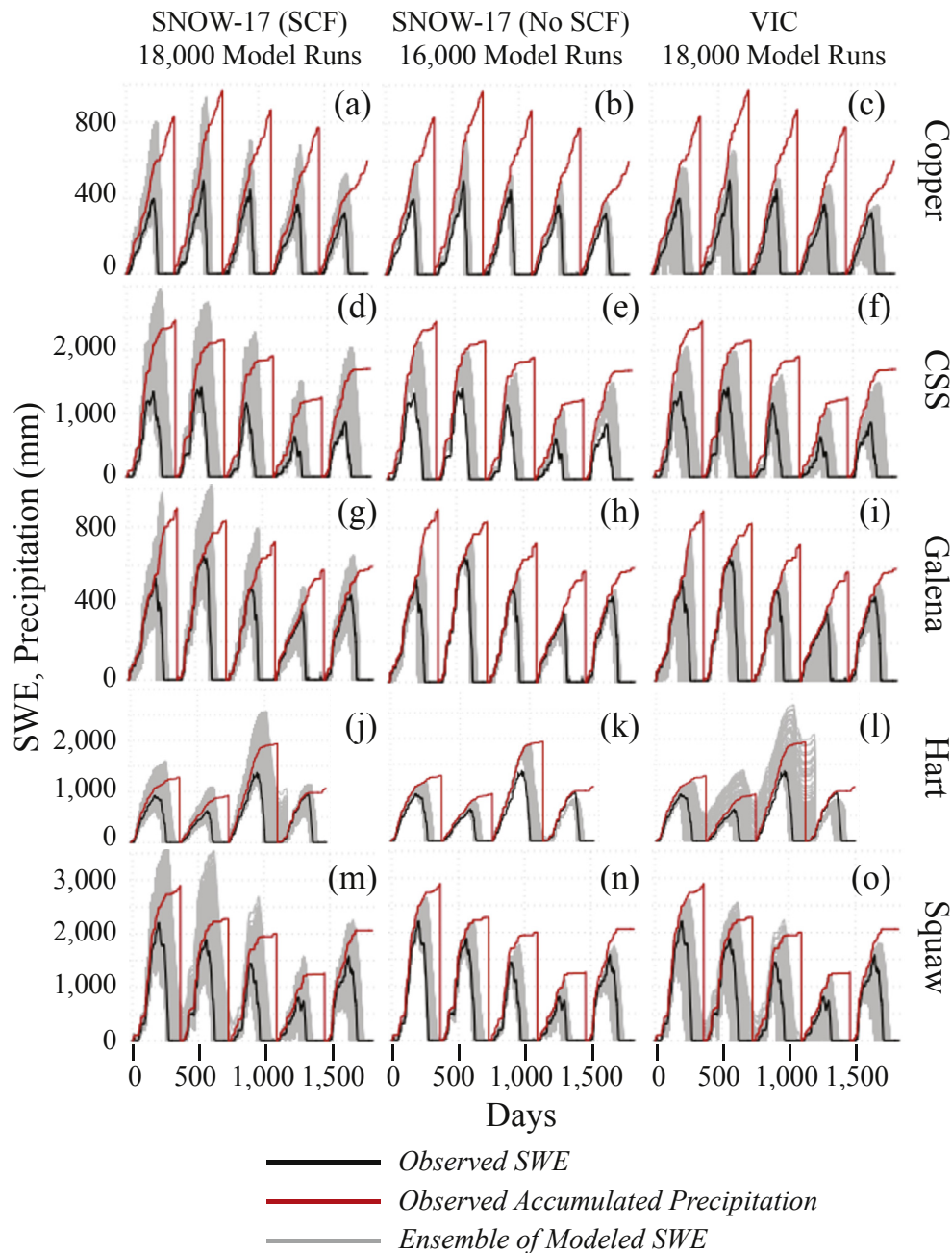


Fig. 4. Ensemble of SNOW-17 and VIC SWE outputs resulting from the ensemble of parameter sets generated in the Sobol' analysis. The black lines represent observed SWE, the red lines represent accumulated precipitation, and the gray areas represent the ensemble of model outputs. The first column of plots, including Fig. 4 (a), (d), (g), (j), and (m), represents the SWE ensembles produced when varying eight of the SNOW-17 parameters. The second column of plots, including Fig. 4 (b), (e), (h), (k), and (n), represents the SWE ensembles produced when the SCF parameter is fixed at a value of 1, and the remaining seven SNOW-17 parameters are still varied. The third column of plots, including Fig. 4 (c), (f), (i), (l), and (o), represents the SWE ensembles produced when varying eight VIC snow parameters. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the Sobol' total and first-order sensitivity indices and interactions for SNOW-17 and VIC parameters. Fig. 5 demonstrates results for SNOW-17, Fig. 6 shows results for VIC, and Fig. 7 shows an additional analysis with a modification to albedo parameter within VIC.

According to the SNOW-17 results in Fig. 5, the sensitivity of MFMAX is largest across all of the site locations and objective functions, except for the PEAK objective function where PLWHC dominates. MFMAX is the maximum melt factor, which occurs on June 21st. MFMAX and MFMIN, the minimum melt factor which occurs on December 21st, are used in an empirical equation to calculate the seasonal variation in the non-rain melt factor within

SNOW-17. However, PLWHC—representing the amount of water a snowpack can hold before output from the snowpack occurs—is the most sensitive with respect to PEAK. He et al. (2011b) also identified MFMAX as a sensitive parameter, but in their study PLWHC was deemed insensitive.

Fig. 6 shows a full sensitivity analysis for all VIC parameters across the sites, demonstrating that the \exp_{A1} and \exp_{A2} parameters exhibit high sensitivity indices across all objective functions. Subsequently, Fig. 7 demonstrates these results across both albedo decay curve parameter scenarios. These parameters are part of the accumulation and ablation albedo decay curve equations (Section

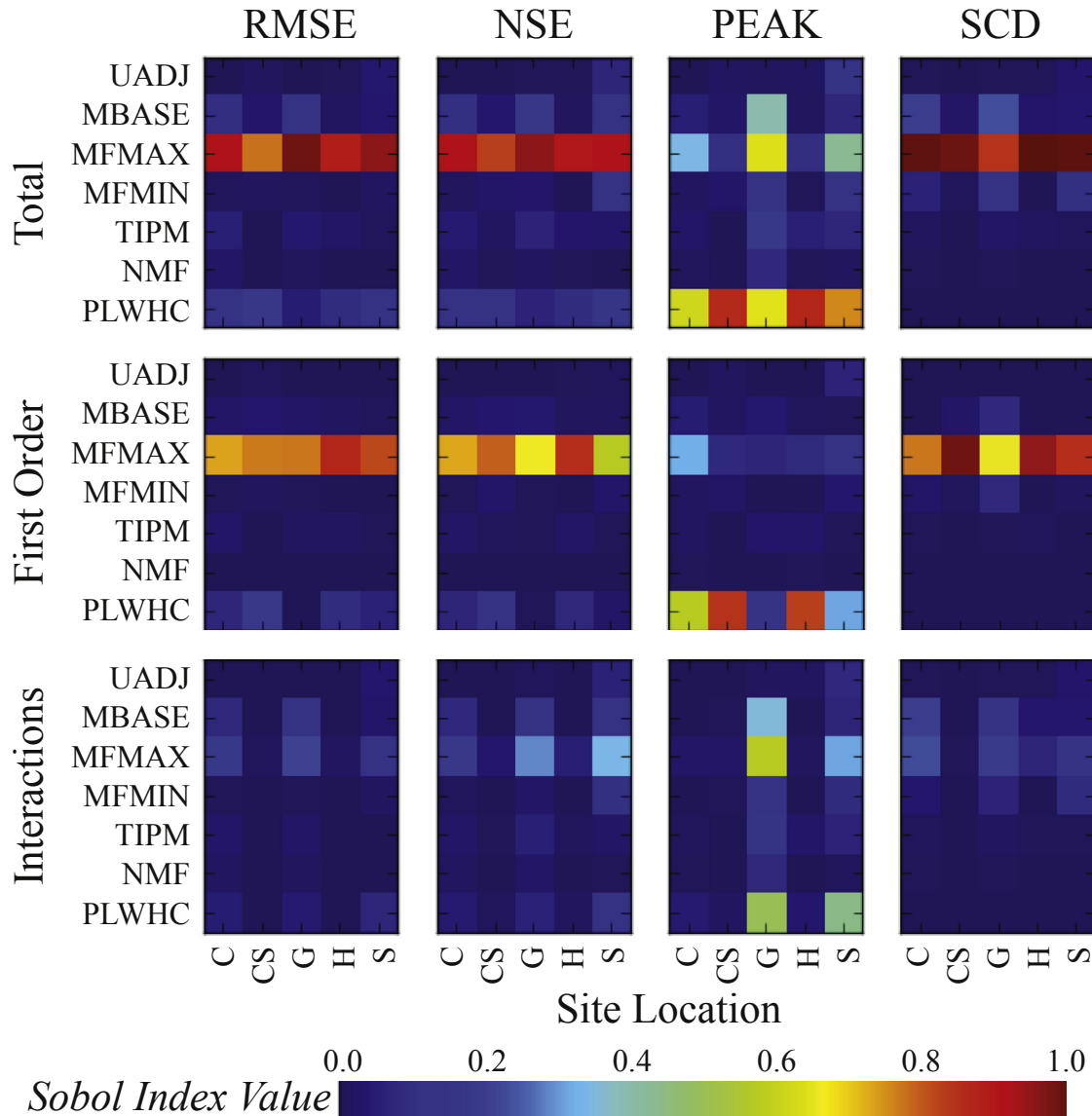


Fig. 5. Total and first-order sensitivity indices and interactions for SNOW-17 parameters. The larger the value of each box, the more sensitive the parameter is with respect to a specific objective function. The x-axis consists of site locations (C – Copper, CS – CSS, G – Galena, H – Hart, S – Squaw), and the y-axis consists of the model's parameters. Each column of rasters contains results with respect to a specific objective function.

2.3), and the default values for these parameters were originally identified based on a single study location in 1972 (Bras, 1990). It is not surprising that the albedo decay curve parameters have the highest sensitivity because albedo is a key factor for calculating snowmelt (Etchevers et al., 2004). However, the results reveal that even in complex physically-based models parameter uncertainty can greatly impact model output (Mendoza et al., 2014).

Figs. 5–7 also demonstrate different dimensions of the Sobol' sensitivity analysis, separating the results into total, first-order, and interactive terms. The first-order sensitivity index of a parameter is a measure of the effect on model output variance of varying only that parameter. Values of first-order indices can provide insight into which factors, if known, would be expected to reduce the variance the most. The high first-order sensitivity of MFMAX in Fig. 5 suggests that, if known, MFMAX would be expected to reduce the variance the most with respect to NSE, RMSE, and SCD, while PLWHC would be expected to reduce it the most for PEAK. Fig. 6 shows that parameter exp_{A2} dominates the VIC first-order sensitivities, with exp_{A1} also showing some larger first-order indices.

The total order indices are the sum of the first and interactive effects and can provide insight with respect to factor fixing. Parameters with low total order indices can be fixed anywhere within their range with minimal impact on the model output. Similarly to the first-order indices, MFMAX and PLWHC exhibit the largest total order indices in SNOW-17 and exp_{A2} and exp_{A1} exhibit the largest total order indices in VIC. According to Figs. 5 and 6, fixing SNOW-17 parameters UADJ and NMF and fixing VIC parameters λ and exp_{A2} to any value within the Sobol' range would minimally affect outputs with respect to all sites and objective functions.

Calculating the sums for the first order and total order indices gives an indication about the presence of non-linearity of interactions. The sums of the total order indices across site locations and objective functions for SNOW-17 range from 1.00 to 2.07, and the sums of the first-order indices for SNOW-17 range from 0.292 to 1.00. The sums of the total order indices for VIC range from 1.03 to 1.98, while the sums of the first-order indices for VIC range from 0.563 to 0.954. When the first-order indices are close to the total order indices, the model is considered to be linear because this

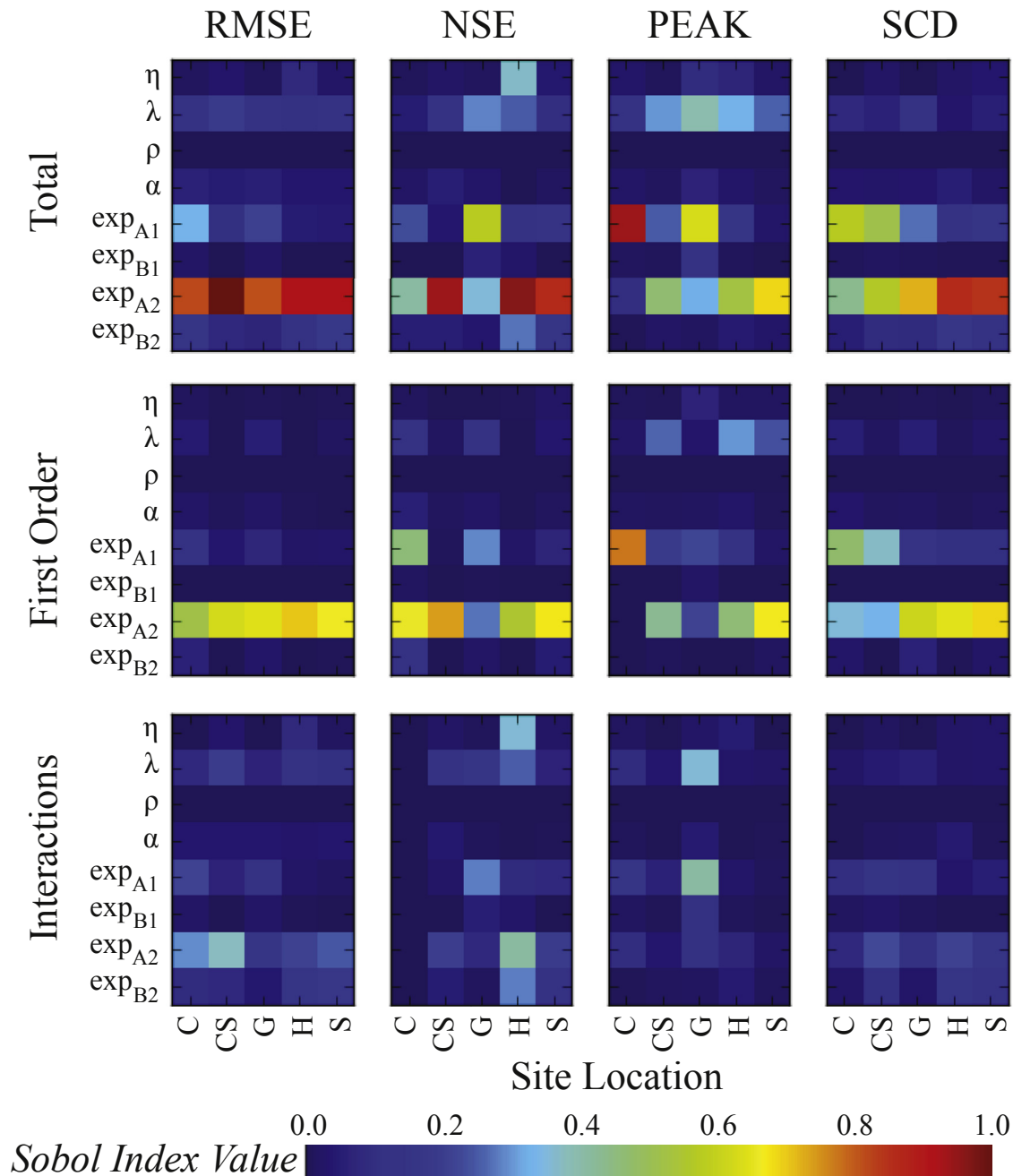


Fig. 6. Total and first-order sensitivity indices and interactions for VIC parameters. The larger the value of each box, the more sensitive the parameter is with respect to a specific objective function. The x-axis consists of site locations (C – Copper, CS – CSS, G – Galena, H – Hart, S – Squaw), and the y-axis consists of the model's parameters. Each column of rasters contains results with respect to a specific objective function.

indicates minimal interactions among parameters. The differences between SNOW-17 total and first-order indices range from 0.004 to 1.78 with a median value of 0.305, and the differences between VIC total and first-order indices range from 0.072 to 1.40 with a median value of 0.485. Generally, the two models appear to be non-linear; however, the small differences between total and first-order sums at the low end of the range indicate that at times the models act in a linear fashion, depending on objective function and site location. Similarly to the total and first-order index results, the MFMAX and PLWHC parameters in SNOW-17, and the exp_{A2} parameter in VIC have the largest values for interactions. However, Fig. 5 depicts that the MBASE, MFMIN, and TIMP parameters have larger values for

interactions than first-order indices. Fig. 6 depicts that η , λ , and exp_{B2} generally have larger values for interactions than first-order indices. The parameters with larger interactions than first-order indices contribute to output variance via their interactions with other parameters, and some of these interactions appear to be noteworthy.

We conclude our analysis with a discussion about the best objective function values that appeared in our model run ensembles, summarized in Table 5. Sampling the parameter space within the Sobol' ensemble leads to improved objective values relative to the baseline runs. Notably, VIC's default parameterization had poor NSE, which is consistent with the findings of Feng et al. (2008), who

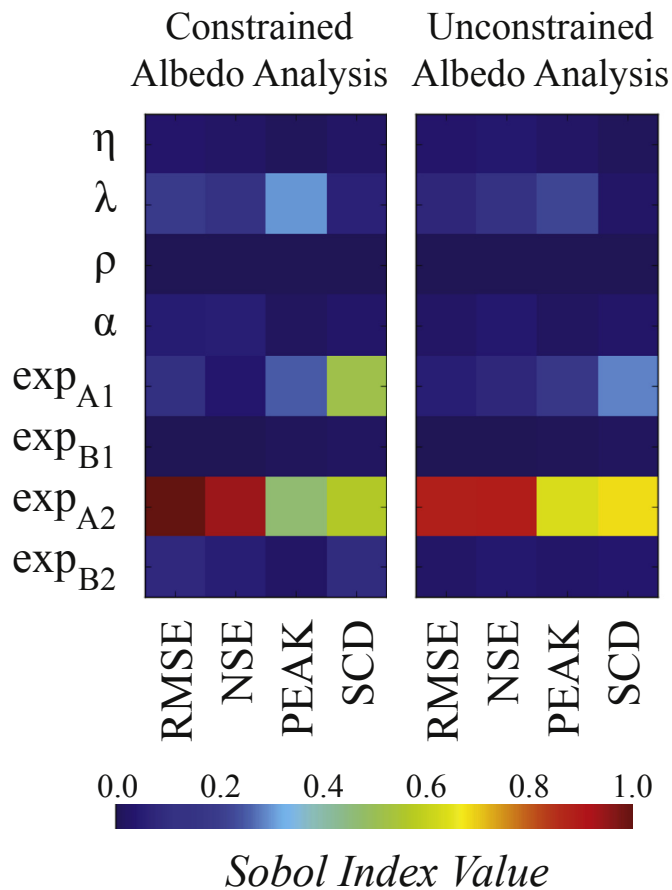


Fig. 7. Total sensitivity indices for VIC parameters at the CSS location. The constrained albedo analysis raster depicts the total indices calculated when limiting the ranges of the albedo decay curve parameters to realistic albedo values (constrained albedo analysis). The unconstrained albedo analysis raster depicts the total indices calculated from a larger range of albedo decay curve parameters that can produce unrealistic albedo values (unconstrained albedo analysis).

conclude VIC is unable to realistically model runoff timing despite water storage and refreezing effects within its physics. Table 5 shows adjusting model parameters helps improve this weakness within both models. VIC also has difficulty minimizing the PEAK objective function whereby SNOW-17 outperforms VIC with the exception of at Copper. The two models achieve similar NSE values for all sites. VIC achieves better RMSE values for CSS and Squaw, while the RMSE values for the remaining sites are similar between SNOW-17 and VIC. Additionally, the constrained albedo analysis yields the best PEAK value for CSS, which indicates VIC's limitation in minimizing the PEAK function could be an artifact of the original albedo decay parameter ranges. The constrained albedo analysis also results in improved NSE values for VIC, while worsening the SCD and RMSE values, resulting in an RMSE value close to that of SNOW-17. These results demonstrate performance of both models can be significantly improved with sampling of their parameter spaces (such as within a Sobol' analysis), and this is especially important for VIC because it demonstrates that adjusting historically fixed parameters can improve performance.

Overall, variations in results among the three climate regimes indicate that site climate characteristics can carry into the signal of model performance. For example, VIC Sobol' results demonstrate that the three climate regimes have different parameter sensitivity patterns. The exp_{A1} parameter is most sensitive with respect to PEAK for the maritime sites, whereas the exp_{A2} parameter is the

most sensitive with respect to PEAK for the continental and inter-mountain sites, as shown in Fig. 6.

3.4. Temperature change results

Fig. 8(a) through (e) shows the daily differences in SWE between SNOW-17 and VIC generated during the Sobol' analysis, and the temperature change scenarios for the same model parameterizations. The timeseries plot in each figure shows the site-specific observed SWE for the water years used in this analysis. The sets of model runs are chosen to examine how errors in parameterizations propagate to portrayals of climate change. Fig. 8 (a) through (e) show the largest differences in daily SWE occurs during peak season (June), while minimal to no daily differences in SWE occur when there is little to no snow.

For all site locations, SNOW-17 produces a similar signal for the best and worst model parameterizations with respect to NSE. For CSS, Squaw, Hart, and Copper, the worst SNOW-17 parameterizations produce peak SWE biased late relative to the best performing parameterizations. For Galena, which has the shallowest snowpack, the worst SNOW-17 parameterizations produce peak SWE with an early bias. The small variation in daily differences in SWE between the best and worst model runs for SNOW-17 suggests the SNOW-17 parameter values do not greatly affect the delta-change in daily SWE. Regardless of baseline model performance, SNOW-17 will produce a similar delta-change in SWE. Consistent SWE daily differences throughout the SNOW-17 model simulations suggest a limitation in SNOW-17 to produce realistic projections under non-stationary conditions due to dependency on model parameters that have been calibrated with respect to historical conditions. This is in accordance with Gan and Burges (1990), who determine calibrated parameters may become invalid when extended beyond hydro-climatic regime used in calibration as a result of such parameters being climate sequence dependent. Dependence on calibrated parameters versus physical processes limits the model's sensitivity to climate change.

The VIC and SNOW-17 temperature change results show the best model simulations produce similar changes in daily delta SWE values. However, the worst model runs have different distributions of daily SWE sensitivities dependent upon site location. For the maritime sites, VIC exhibits large changes in daily SWE during peak SWE months (April, May, June), and minimal changes in daily SWE otherwise. For the inter-mountain and continental sites, there is little to no change in daily SWE throughout the entire water year.

SNOW-17 produces similar signals between the best and worst performing baseline model runs, but we do not know if these sensitivities represent true climate conditions. VIC demonstrates consistent variation in the daily SWE delta change throughout different model runs. Further investigation reveals that poor parameterizations in VIC underestimate SWE because albedo is too low and does not maintain the typical SWE accumulation and ablation pattern, especially for the sites with shallow snowpack. Conversely, poor performing SNOW-17 parameterizations overestimate SWE, but maintain the typical SWE accumulation and ablation pattern throughout the water years. The lowest performing VIC parameterizations produce outputs that perform considerably poorer than the lowest performing SNOW-17 parameterizations. We attribute this behavior to the greater sensitivity of the VIC model parameters and their physical interactions.

4. Conclusions

Hydrological models are used to project future water resources conditions and guide important management decisions, such as

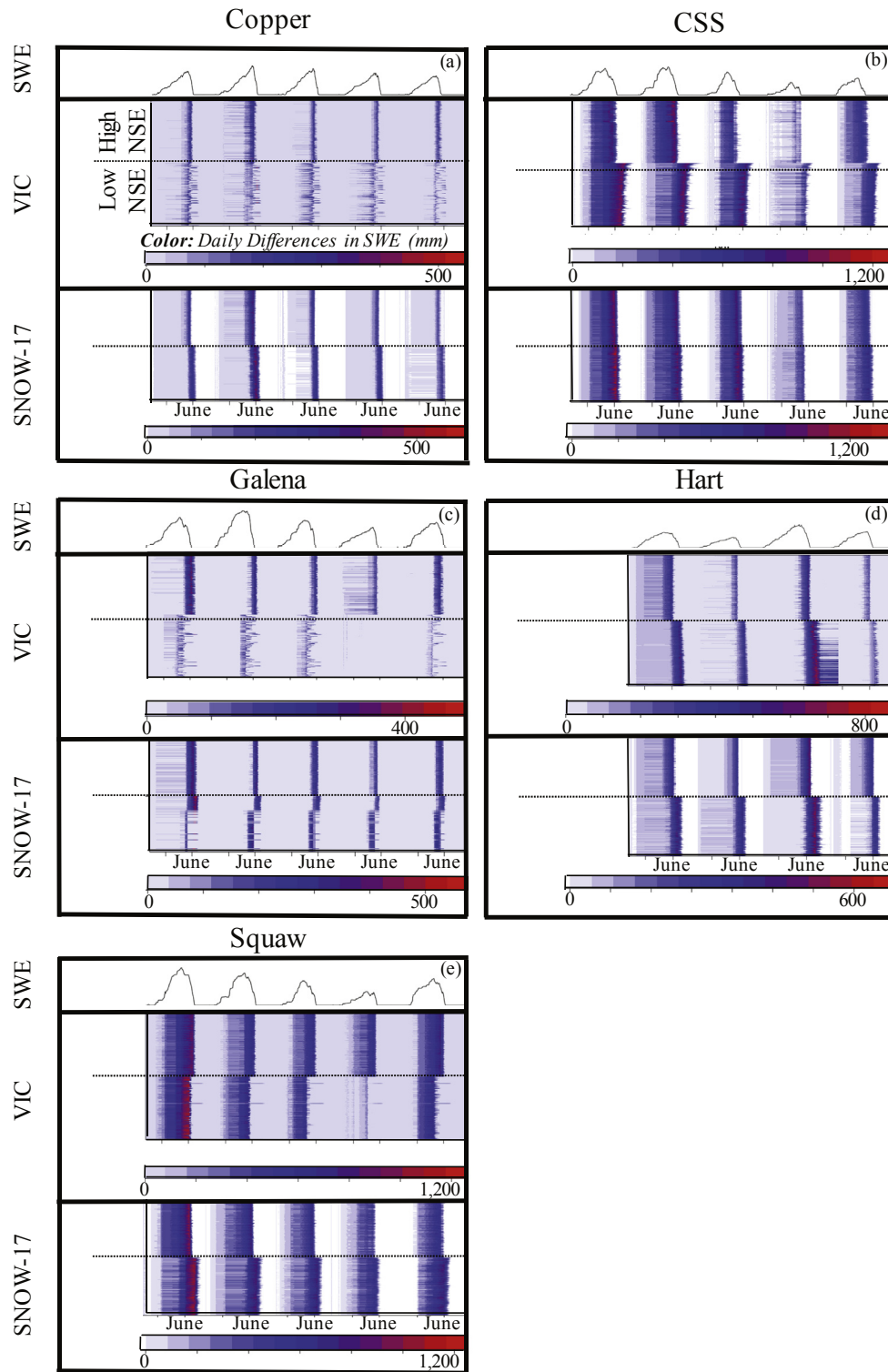


Fig. 8. Timeseries showing the daily difference in SWE between the original ensembles and the temperature change results for SNOW-17 and VIC; units are shown in plot (a). Each column in the rasters represents a day, each row represents a different model parameterization, and the color represents the daily SWE differences between the control modeled SWE and temperature change modeled SWE. The top half of the rasters for each site depict differences for the top 2% of model runs (360 runs for VIC, 320 runs for SNOW-17) with respect to NSE (High NSE) from the Sobol' analysis. The bottom half of the rasters depict differences for the worst 2% of model runs with respect to NSE (Low NSE). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

those made by the U.S. Bureau of Reclamation that have the potential to impact water systems (Kalra et al., 2014). However, hydrological modeling uncertainty ultimately propagates into model

projections, and as a result, there are implications of this uncertainty in water management decisions. Insights from our research and related studies can be used to develop and improve models

that allow for exploration of parameter uncertainties and improve modeling practices to inform more robust management decisions. Continued research on understanding and improving hydrological model performance—especially with respect to applying sophisticated diagnostic tools and integrating them into operations—is necessary to improve modeling under a changing climate. Current modeling practices do not always deliver the relevant information required for water management decisions, and model diagnostic exercises such as this one can potentially aid the models' usability.

This research contributes to the body of work focused on improving SWE projections and supports the following conclusions. Historically-fixed parameters within the VIC snow model, especially the \exp_{A1} and \exp_{A2} parameters, can significantly affect VIC SWE output. The \exp_{A2} and \exp_{A1} parameters were shown to be the first- and second-most sensitive model parameters, respectively. Furthermore, calibrating the VIC snow model, albeit not common practice, results in significant improvements in simulated SWE. Snow model calibration currently requires editing of hard-coded parameters, and our research motivates efforts for model developers to ease users' ability to adjust parameters. Consistent with previous studies, we find no clear evidence that a physically-based model outperforms a conceptual model. Detailed comparison across two snow models demonstrates baseline results of SNOW-17 generally perform better for shallower snowpacks, while VIC performs better for deep snowpacks. In both models, different climate regimes end up having different sensitivity patterns. SNOW-17 achieves better PEAK objective function values across all sites; yet, the constrained albedo analysis results suggest that the user-selected Sobol' parameter ranges dictate VIC performance with respect to PEAK. However, results indicate the two models are capable of producing similar objective function values for SCD, NSE, and RMSE. Our results have the potential to inform calibration exercises by fixing some insensitive parameters and also guiding users on which parameters in both models are most important and thus deserve close attention when using the models in an operational context. A full calibration is outside the scope of this paper, but Houle (2015) describes additional work investigating potential calibrations for the two models using multi-objective optimization techniques.

Software availability

The model diagnostic calculations utilized the MOEAFramework, which is freely available at <http://www.moeaframework.org/>. Specifically, the MOEAFramework has the capability to generate model parameterizations and calculate Sobol' sensitivity indices. The Variable Infiltration Capacity hydrologic model is available at <http://www.hydro.washington.edu/Lettenmaier/Models/VIC/Overview/ModelOverview.shtml>. The SNOW-17 implementation used in the study is available on GitHub at https://github.com/jrkasprzyk/misc_scripts/blob/master/snow17.m. Downscaled hydrologic projections are available here: http://gdo-dcp.ucllnl.org/downscaled_cmip_projections/dcpInterface.html. The data used in the study are freely available, and will be provided upon request from the corresponding author.

Acknowledgements

Funding for this research is from the University of Colorado Boulder, Civil, Environmental, and Architectural Engineering Department. This work utilized the Janus supercomputer, which is supported by the National Science Foundation (award number CNS-0821794) and the University of Colorado Boulder. The Janus supercomputer is a joint effort of the University of Colorado Boulder, the University of Colorado Denver and the National Center

for Atmospheric Research. There are no apparent conflicts of interest from affiliations or funding. We acknowledge the World Climate Research Programme's Working Group on Coupled Modeling, which is responsible for CMIP, and we thank the climate modeling groups for producing and making available their model output. For CMIP the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. We also acknowledge the constructive comments of peer reviewers as well as the careful proofreading of Joseph Ronald Kasprzyk.

References

- Anderson, E., 2006. National Weather Service River Forecast System: Snow Accumulation and Ablation Model - SNOW-17.
- Anderson, E.A., 1973. National Weather Service River Forecast System—snow Accumulation and Ablation Model, Technical Memorandum NWS HYDRO-17. November 1973. 217 P.
- Barnett, T.P., Adam, J.C., Lettenmaier, D.P., 2005. Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature* 438, 303–309. <http://dx.doi.org/10.1038/nature04141>.
- Borgonovo, E., 2006. Measuring uncertainty importance: investigation and comparison of alternative approaches. *Risk Anal.* 26, 1349–1361. <http://dx.doi.org/10.1111/j.1539-6924.2006.00806.x>.
- Bras, R.L., 1990. *Hydrology: an Introduction to Hydrological Science*. Addison-Wesley.
- Butler, M.P., Reed, P.M., Fisher-Vanden, K., Keller, K., Wagener, T., 2014. Identifying parametric controls and dependencies in integrated assessment models using global sensitivity analysis. *Environ. Model. Softw.* 59, 10–29. <http://dx.doi.org/10.1016/j.envsoft.2014.05.001>.
- Cherkauer, K.A., Bowling, L.C., Lettenmaier, D.P., 2003. Variable infiltration capacity cold land process model updates. *Glob. Planet. Change* 38, 151–159. [http://dx.doi.org/10.1016/S0921-8181\(03\)00025-0](http://dx.doi.org/10.1016/S0921-8181(03)00025-0).
- Confalonieri, R., Bellocchi, G., Bregaglio, S., Donatelli, M., Acutis, M., 2010. Comparison of sensitivity analysis techniques: a case study with the rice model WARM. *Ecol. Model.* 221, 1897–1906. <http://dx.doi.org/10.1016/j.ecolmodel.2010.04.021>.
- Cosenza, A., Mannina, G., Vanrolleghem, P.A., Neumann, M.B., 2013. Global sensitivity analysis in wastewater applications: a comprehensive comparison of different methods. *Environ. Model. Softw.* 49, 40–52. <http://dx.doi.org/10.1016/j.envsoft.2013.07.009>.
- Crane, R.G., Hewitson, B.C., 1998. Doubled CO2 precipitation changes for the susquehanna basin: downscaling from the genesis general circulation model. *Int. J. Climatol.* 18, 65–76.
- Demaria, E.M., Nijssen, B., Wagener, T., 2007. Monte Carlo sensitivity analysis of land surface parameters using the Variable Infiltration Capacity model. *J. Geophys. Res.* 112, D11113. <http://dx.doi.org/10.1029/2006JD007534>.
- Denoth, A., 2003. Structural phase changes of the liquid water component in Alpine snow. *Cold Reg. Sci. Technol.* 37, 227–232. [http://dx.doi.org/10.1016/S0165-232X\(03\)00066-1](http://dx.doi.org/10.1016/S0165-232X(03)00066-1).
- Dore, M.H.I., 2005. Climate change and changes in global precipitation patterns: what do we know? *Environ. Int.* 31, 1167–1181. <http://dx.doi.org/10.1016/j.envint.2005.03.004>.
- Essery, R., Morin, S., Lejeune, Y., B. Ménard, C., 2013. A comparison of 1701 snow models using observations from an alpine site. *Adv. Water Resour.* 55, 131–148. <http://dx.doi.org/10.1016/j.advwatres.2012.07.013>.
- Etchevers, P., Martin, E., Brown, R., Fierz, C., Lejeune, Y., Bazile, E., Boone, A., Dai, Y.-J., Essery, R., Fernandez, A., Gusev, Y., Jordan, R., Koren, V., Kowalczyk, E., Nasonova, N.O., Pyles, R.D., Schlosser, A., Shmakin, A.B., Smirnova, T.G., Strasser, U., Verseghy, D., Yamazaki, T., Yang, Z.-L., 2004. Validation of the energy budget of an alpine snowpack simulated by several snow models (SnowMIP project). *Ann. Glaciol.* 38, 150–158. <http://dx.doi.org/10.3189/172756404781814825>.
- Feng, X., Sahoo, A., Arsenault, K., Houser, P., Luo, Y., Troy, T.J., 2008. The impact of snow model complexity at three CLPX sites. *J. Hydrometeorol.* 9, 1464–1481. <http://dx.doi.org/10.1175/2008JHM860.1>.
- Feeberg, J., Jenkins, K.J., 2005. Assessing uncertainty in ecological systems using global sensitivity analyses: a case example of simulated wolf reintroduction effects on elk. *Ecol. Model.* 187, 259–280. <http://dx.doi.org/10.1016/j.ecolmodel.2005.01.042>.
- Franz, K.J., Hogue, T.S., Sorooshian, S., 2008. Operational snow modeling: addressing the challenges of an energy balance model for National Weather Service forecasts. *J. Hydrol.* 360, 48–66. <http://dx.doi.org/10.1016/j.jhydrol.2008.07.013>.
- Ganji, A., Maier, H.R., Dandy, G.C., 2016. A modified Sobol' sensitivity analysis method for decision-making in environmental problems. *Environ. Model. Softw.* 75, 15–27. <http://dx.doi.org/10.1016/j.envsoft.2015.10.001>.
- Gan, T.Y., Burges, S.J., 1990. An assessment of a conceptual rainfall-runoff model's ability to represent the dynamics of small hypothetical catchments: 2.

- Hydrologic responses for normal and extreme rainfall. *Water Resour. Res.* 26, 1605–1619. <http://dx.doi.org/10.1029/WR026i007p01605>.
- Gan, Y., Duan, Q., Gong, W., Tong, C., Sun, Y., Chu, W., Ye, A., Miao, C., Di, Z., 2014. A comprehensive evaluation of various sensitivity analysis methods: a case study with a hydrological model. *Environ. Model. Softw.* 51, 269–285. <http://dx.doi.org/10.1016/j.envsoft.2013.09.031>.
- Groisman, P.Y., Legates, D.R., 1994. The accuracy of United States precipitation data. *Bull. Am. Meteorol. Soc.* 75, 215–227. [http://dx.doi.org/10.1175/1520-0477\(1994\)075<0215:TAOUP>2.0.CO;2](http://dx.doi.org/10.1175/1520-0477(1994)075<0215:TAOUP>2.0.CO;2).
- Hall, J.W., Tarantola, S., Bates, P.D., Horritt, M.S., 2005. Distributed sensitivity analysis of flood inundation model calibration. *J. Hydraul. Eng.-ASCE* 131, 117–126. [http://dx.doi.org/10.1061/\(ASCE\)0733-9429\(2005\)131:2\(117\)](http://dx.doi.org/10.1061/(ASCE)0733-9429(2005)131:2(117)).
- Hay, L.E., Wilby, R.L., Leavesley, G.H., 2000. A comparison of delta change and downscaled GCM scenarios for three mountainous basins in the United States¹. *JAWRA J. Am. Water Resour. Assoc.* 36, 387–397. <http://dx.doi.org/10.1111/j.1752-1688.2000.tb04276.x>.
- He, M., Hogue, T.S., Franz, K.J., Margulis, S.A., Vrugt, J.A., 2011a. Characterizing parameter sensitivity and uncertainty for a snow model across hydroclimatic regimes. *Adv. Water Resour.* 34, 114–127. <http://dx.doi.org/10.1016/j.advwatres.2010.10.002>.
- He, M., Hogue, T.S., Franz, K.J., Margulis, S.A., Vrugt, J.A., 2011b. Corruption of parameter behavior and regionalization by model and forcing data errors: a Bayesian example using the SNOW17 model. *Water Resour. Res.* 47 <http://dx.doi.org/10.1029/2010WR009753> n/a–n/a.
- Hogue, T.S., Sorooshian, S., Gupta, H., Holz, A., Braatz, D., 2000. A multistep automatic calibration scheme for river forecasting models. *J. Hydrometeorol.* 1, 524–542. doi:10.1175/1525-7541(2000)001<0524:AMACSF>2.0.CO;2
- Houle, E.S., 2015. Inter-model Diagnostics for Two Snow Models and Implications for Management. M.S. Thesis. University of Colorado Boulder, Boulder, CO.
- Kalra, H., Hallegatte, S., Lempert, R., Brown, C., Fozzard, A., Gill, S., Shah, A., 2014. Agreeing on Robust Decisions: New Processes for Decision Making under Deep Uncertainty.
- Kontoravdi, C., Asprey, S.P., Pistikopoulos, E.N., Mantalaris, A., 2005. Application of global sensitivity analysis to determine goals for design of experiments: an example study on antibody-producing cell cultures. *Biotechnol. Prog.* 21, 1128–1135. <http://dx.doi.org/10.1021/bp050028k>.
- Leisenring, M., Moradkhani, H., 2011. Snow water equivalent prediction using Bayesian data assimilation methods. *Stoch. Environ. Res. Risk Assess.* 25, 253–270. <http://dx.doi.org/10.1007/s00477-010-0445-5>.
- Liang, X., Lettenmaier, D.P., Wood, E.F., Burges, S.J., 1994. A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *J. Geophys. Res. Atmos.* 99, 14415–14428. <http://dx.doi.org/10.1029/94JD00483>.
- Livneh, B., Bohn, T.J., Pierce, D.W., Munoz-Arriola, F., Nijssen, B., Vose, R., Cayan, D.R., Brekke, L., 2015. A spatially comprehensive, hydrometeorological data set for Mexico, the U.S., and Southern Canada 1950–2013. *Sci. Data* 2, 150042. <http://dx.doi.org/10.1038/sdata.2015.42>.
- Livneh, B., Deems, J.S., Schneider, D., Barsugli, J.J., Molotch, N.P., 2014. Filling in the gaps: inferring spatially distributed precipitation from gauge observations over complex terrain. *Water Resour. Res.* 50, 8589–8610. <http://dx.doi.org/10.1002/2014WR015442>.
- Livneh, B., Rosenberg, E.A., Lin, C., Nijssen, B., Mishra, V., Andreadis, K.M., Maurer, E.P., Lettenmaier, D.P., 2013. A long-term hydrologically based dataset of land surface fluxes and States for the conterminous United States: update and extensions*. *J. Clim.* 26, 9384–9392. <http://dx.doi.org/10.1175/JCLI-D-12-00508.1>.
- McClung, D., Schaerer, P.A., 2006. *The Avalanche Handbook*, third ed. The Mountaineers Books.
- Mendoza, P.A., Clark, M.P., Barlage, M., Rajagopalan, B., Samaniego, L., Abramowitz, G., Gupta, H., 2014. Are we unnecessarily constraining the agility of complex process-based models? *Water Resour. Res.* <http://dx.doi.org/10.1002/2014WR015820> n/a–n/a.
- Mizukami, N.P., Clark, M.G., Slater, A.D., Brekke, L.M., Elsner, M.R., Arnold, J., Gangopadhyay, S., 2013. Hydrologic implications of different large-scale meteorological model forcing datasets in mountainous regions. *J. Hydrometeorol.* 15, 474–488. <http://dx.doi.org/10.1175/JHM-D-13-036.1>.
- Mote, P.W., Hamlet, A.F., Clark, M.P., Lettenmaier, D.P., 2005. Declining mountain snowpack in western north America*. *Bull. Am. Meteorol. Soc.* 86, 39–49. <http://dx.doi.org/10.1175/BAMS-86-1-39>.
- Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I — a discussion of principles. *J. Hydrol.* 10, 282–290. [http://dx.doi.org/10.1016/0022-1694\(70\)90255-6](http://dx.doi.org/10.1016/0022-1694(70)90255-6).
- Neumann, M.B., 2012. Comparison of sensitivity analysis methods for pollutant degradation modelling: a case study from drinking water treatment. *Sci. Total Environ.* 433, 530–537. <http://dx.doi.org/10.1016/j.scitotenv.2012.06.026>.
- Norton, J., 2015. An introduction to sensitivity assessment of simulation models. *Environ. Model. Softw.* 69, 166–174. <http://dx.doi.org/10.1016/j.envsoft.2015.03.020>.
- Nossent, J., Elsen, P., Bauwens, W., 2011. Sobol' sensitivity analysis of a complex environmental model. *Environ. Model. Softw.* 26, 1515–1525. <http://dx.doi.org/10.1016/j.envsoft.2011.08.010>.
- Pianosi, F., Beven, K., Freer, J., Hall, J.W., Rougier, J., Stephenson, D.B., Wagener, T., 2016. Sensitivity analysis of environmental models: a systematic review with practical workflow. *Environ. Model. Softw.* 79, 214–232. <http://dx.doi.org/10.1016/j.envsoft.2016.02.008>.
- Raje, D., Krishnan, R., 2012. Bayesian parameter uncertainty modeling in a macro-scale hydrologic model and its impact on Indian river basin hydrology under climate change. *Water Resour. Res.* 48 <http://dx.doi.org/10.1029/2011WR011123> n/a–n/a.
- Rasmussen, R., Baker, B., Kochendorfer, J., Meyers, T., Landolt, S., Fischer, A.P., Black, J., Thériault, J.M., Kucera, P., Gochis, D., Smith, C., Nitu, R., Hall, M., Ikeda, K., Gutmann, E., 2012. How well are we measuring snow: the NOAA/FAA/NCAR winter precipitation test. *Bed. Bull. Am. Meteorol. Soc.* 93, 811–829. <http://dx.doi.org/10.1175/BAMS-D-11-00052.1>.
- Rosolem, R., Gupta, H.V., Shuttlesworth, W.J., Gonçalves, L.G.G., Zeng, X., 2012. Towards a comprehensive approach to parameter estimation in land surface parameterization schemes. *Hydrol. Process* 27, 2075–2097.
- Saltelli, A., 2002. Making best use of model evaluations to compute sensitivity indices. *Comput. Phys. Commun.* 145, 280–297.
- Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gattelli, D., Saisana, M., Tarantola, S., 2008. *Global Sensitivity Analysis: the Primer*. Wiley.
- Saltelli, A., Tarantola, A., Campolongo, F., Ratto, M., 2004. *Sensitivity Analysis in Practice - a Guide to Assessing Scientific Models*. John Wiley and Sons, Chichester.
- Saltelli, A., Tarantola, S., Chan, K.P.S., 1999. A quantitative model-independent method for global sensitivity analysis of model output. *Technometrics* 41, 39–56. <http://dx.doi.org/10.2307/1270993>.
- Sarrazin, F., Pianosi, F., Wagener, T., 2016. Global sensitivity analysis of environmental models: convergence and validation. *Environ. Model. Softw.* 79, 135–152. <http://dx.doi.org/10.1016/j.envsoft.2016.02.005>.
- Singh, P., Singh, V.P., 2001. *Snow and Glacier Hydrology*. Springer Science+Business Media.
- Sobol', I.M., 1993. Sensitivity estimates for nonlinear mathematical models and their Monte Carlo estimation. *Math. Model Comput. Exp.* 1, 407–417.
- 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.
- Stewart, I.T., Cayan, D.R., Dettinger, M.D., 2004. Changes in snowmelt runoff timing in western North America under a 'Business as usual' climate change scenario. *Clim. Change* 62, 217–232. <http://dx.doi.org/10.1023/b:clim.0000013702.22656.e8>.
- Tang, Y., Reed, P., Wagener, T., Van Werkhoven, K., 2007. Comparing sensitivity analysis methods to advance lumped watershed model identification and evaluation. *Hydrol. Earth Syst. Sci.* 11, 793–817.
- Taylor, K.E., Stouffer, R.J., Meehl, G.A., 2012. An overview of CMIP5 and the experiment design. *Bull. Am. Meteorol. Soc.* 93, 485–498. <http://dx.doi.org/10.1175/BAMS-D-11-00094.1>.
- Thornton, P.E., Hasenauer, H., White, M.A., 2000. Simultaneous estimation of daily solar radiation and humidity from observed temperature and precipitation: an application over complex terrain in Austria. *Agric. For. Meteorol.* 104, 255–271. [http://dx.doi.org/10.1016/S0168-1923\(00\)00170-2](http://dx.doi.org/10.1016/S0168-1923(00)00170-2).
- Tian, W., 2013. A review of sensitivity analysis methods in building energy analysis. *Renew. Sustain. Energy Rev.* 20, 411–419. <http://dx.doi.org/10.1016/j.rser.2012.12.014>.
- United States Army Corps of Engineers, 1956. *Snow Hydrology (Summary report of the Snow Investigations, US Army Corps)*.
- Van Werkhoven, K., Wagener, T., Reed, P., Tang, Y., 2008. Characterization of watershed model behavior across a hydroclimatic gradient. *Water Resour. Res.* 44 <http://dx.doi.org/10.1029/2007WR006271> W01429.
- Vezzaro, L., Mikkelsen, P.S., 2012. Application of global sensitivity analysis and uncertainty quantification in dynamic modelling of micropollutants in storm-water runoff. *Environ. Model. Softw.* 27–28, 40–51. <http://dx.doi.org/10.1016/j.envsoft.2011.09.012>.
- Wagener, T., 2003. Evaluation of catchment models. *Hydrol. Process* 17, 3375–3378.
- Wilby, R.L., Wigley, T.M.L., 1997. Downscaling general circulation model output: a review of methods and limitations. *Prog. Phys. Geogr.* 21, 530–548. <http://dx.doi.org/10.1177/030913339702100403>.
- Wiscombe, W.J., Warren, S.G., 1980. A model for the spectral albedo of snow. I Pure Snow. *J. Atmos. Sci.* 37, 2712–2733. [http://dx.doi.org/10.1175/1520-0469\(1980\)037<2712:AMFTSA>2.0.CO;2](http://dx.doi.org/10.1175/1520-0469(1980)037<2712:AMFTSA>2.0.CO;2).
- Wood, A.W., Leung, L.R., Sridhar, V., Lettenmaier, D.P., 2004. Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Clim. Change* 62, 189–216. <http://dx.doi.org/10.1023/B:CLIM.0000013685.99609.9e>.
- Xue, Y., Sellers, P.J., Kinter, J.L., Shukla, J., 1991. A simplified Biosphere model for global climate studies. *J. Clim.* 4, 345–364. [http://dx.doi.org/10.1175/1520-0442\(1991\)004<0345:ASBMFG>2.0.CO;2](http://dx.doi.org/10.1175/1520-0442(1991)004<0345:ASBMFG>2.0.CO;2).
- Zepp, H.B., Kasprzyk, J.R., Herman, J.D., Reed, P.M., Characklis, G.W., 2014. Navigating financial and supply reliability tradeoffs in regional drought management portfolios. *Water Resour. Res.* 50, 4906–4923. <http://dx.doi.org/10.1002/2013WR015126>.