Sensitivity of Snow Simulations to Different Atmospheric Forcing Data Sets in the Land Surface Model CAS-LSM

Yan Wang1,2, Zhenghui Xie1,2✉, Binghao Jia1✉, Longhuan Wang1,2, Ruichao Li1,2, Bin Liu1,2, Si Chen1,2, Jinbo Xie1✉, and Peihua Qin1✉

1State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China, 2College of Earth and Planetary Sciences, University of Chinese Academy of Sciences, Beijing, China

Abstract The quality of snow simulations in land surface models (LSMs) largely depends on the accuracy of the atmospheric forcing data, especially precipitation and air temperature. To investigate the sensitivities of snow simulations to atmospheric forcing, historical simulations from 1981–2010 were conducted using the Chinese Academy of Sciences land surface model (CAS-LSM) with four atmospheric forcing data sets: third Global Soil Wetness Projects (GSWP3), the Water and Global Change (WATCH) Forcing Data (WFD/WFDEI), the Climate Research Unit - National Centers for Environmental Prediction (CRU-NCEP), and Princeton. A sensitivity index (Ψ) is utilized to quantify the sensitivity of the simulated snow cover fraction (SCF), snow water equivalent (SWE), and snow depth (SDP) to the uncertainties in forcing data. By comparing the simulated results with satellite-based products and in situ observations, we find that CAS-LSM generally captured the spatial and seasonal variations of the SCF, SWE, and SDP. The simulation based on GSWP3 produced more reasonable estimates than the other three simulations, particularly for the SCF and SDP. The sensitivity analysis suggested that the SWE and SDP suffered the most from the uncertainties in the atmospheric forcing data sets. The sensitivities of the SCF, SWE, and SDP to precipitation uncertainties had various seasonal cycles depending on the climate regime. The highest sensitivity in boreal climates appeared during January–March, while in warm temperate climates the highest sensitivity existed during April–June. On average, the strongest precipitation sensitivity was found for temperate climates with cold, dry winters. The sensitivities to uncertainties in air temperature showed similar patterns; however, air temperature sensitivity was generally dominant over precipitation sensitivity in boreal climates, while in warm temperate climates both of them were high.

Plain Language Summary We conducted four historical simulations using the land surface model Chinese Academy of Sciences land surface model (CAS-LSM) with four state-of-the-art atmospheric forcing data sets, that is, third Global Soil Wetness Project (GSWP3), the Water and Global Change (WATCH) Forcing Data (WFD/WFDEI), the Climate Research Unit - National Centers for Environmental Prediction (CRU-NCEP), and Princeton, to investigate the sensitivities of snow simulations to different atmospheric forcing data sets. The land surface model demonstrated its capability in reproducing temporal and spatial variations of historical snowpack over the Northern Hemisphere. Compared with the other three simulations, the simulation based on GSWP3 produced more reasonable estimates, particularly for the snow cover fraction and snow depth. Sensitivity analysis suggested that snow mass and snow depth suffered the most from the uncertainties in the atmospheric forcing data sets. The simulated snowpack is more sensitive to the uncertainties in precipitation and air temperature in temperate climates with cold, dry winters. This work highlighted that reliable forcing data, especially precipitation data, are of great significance to the accuracy of snow simulations.

1. Introduction

With a large-scale distribution over the Northern Hemisphere (NH), seasonal snow is one of the most important components of hydrologic processes and climate systems (Robinson & Frei, 2000). The most recognized effects of snow on climate are related to its high albedo, which strongly influences the surface and lower atmosphere radiation and energy budgets (Cohen, 1994; Qu & Hall, 2013). Snow affects the ground thermal regime due to its high emissivity and absorptivity, low thermal conductivity, and high latent heat (Gouttevin et al., 2012; Zhang, 2005). Snow variability may also affect regional weather patterns through large-scale...
atmospheric circulation (Lemke et al., 2007; Zuo et al., 2011). On the other side, changes in climate influence
the spatial and temporal fluctuation of snow, as snow accumulation and melting processes are closely
related to air temperature and precipitation (Brown & Mote, 2009). The observed spring snow cover extent
over the NH has decreased continuously over recent decades in response to global warming (Brutel-Vuille, 2016). A warming climate leads to less snowfall, earlier spring ablation, and shorter snow seasons, which in turn alter hydrological cycles and water resource distribution (IPCC, 2014). Changes in snowmelt runoff have been shown to be one of the most significant responses to climate change (Zakharova et al., 2011). Therefore, accurate estimations of snow evolution in current land surface models (LSMs) and Earth System Models (ESMs) are important toward improving hydrological modeling, numerical weather forecasting, and climate prediction.

Snow simulations in LSMs mainly suffer from errors in the forcing data and imperfect model parameterizations. Great efforts have been made to improve model simulations via parameterization development, and the physical representations of snow in LSMs vary widely, ranging from simple slab models to multilayer models that simulate more sophisticated physical processes of snow. Toure et al. (2016) provided a comprehensive evaluation of snow simulations from the Community Land Model, Version 4 (CLM4) including the snow cover fraction (SCF), snow water equivalent (SWE), and snow depth (SDP) and attributed the modeled discrepancies to the snow parameterization and the subgrid structure in CLM4. Xie et al. (2016) recently investigated the impact of two snow cover schemes (Niu & Yang, 2007; Swenson & Lawrence, 2012) in the CLM4.5 on the snow distribution and surface energy budget over the Tibetan Plateau (TP). To comprehensively evaluate the performance of current snow schemes, several model intercomparison projects focusing on snow have been developed in recent years, notably the Snow Model Intercomparison Project (SnowMIP) Phase 1 (Etchevers et al., 2004) and Phase 2 (Essery et al., 2009; Rutter et al., 2009). These intercomparisons provided insights into the performance of snow models of different complexity and highlighted some common problems such as treatment of forest snow processes.

LSMs require extensive meteorological forcing data, including data on precipitation, temperature, shortwave and longwave radiation, humidity, surface pressure, and wind speed. Currently, a number of state-of-the-art global scale atmospheric forcing data sets are available that are usually derived from reanalysis products in combination with ground-based observations or remotely sensed data. Related forcing products include the meteorological forcing data of the third Global Soil Wetness Projects (GSWP3; Kim, 2017), the Water and Global Change (WATCH) Forcing Data (WFD/WFDEI; Weedon et al., 2011, 2014), the Climate Research Unit - National Centers for Environmental Prediction (CRU-NCEP) forcing data (Viovy, 2011), and the Princeton Global Forcing data (Sheffield et al., 2006). Several past studies have suggested that different meteorological forcing data sets result in large differences in the modeled snow processes (Mizukami et al., 2014; Müller Schmied et al., 2016; Schmucki et al., 2014; Wang et al., 2016). The results from previous studies show that uncertainties in precipitation and temperature largely affect the SWE, melting rates, and snow disappearance timing (Raleigh et al., 2015), while other forcing variables such as radiative fluxes and humidity also affect estimates of the SWE and snowpack ablation (Harpold & Brooks, 2018; Raleigh et al., 2016). In fact, wind speed has proven to be important in simulations of blowing snow and snowdrift sublimation in windy regions (Xie et al., 2018). Although there have been an increasing number of studies exploring uncertainties related to simulations of hydrological and snow processes in LSMs, these uncertainties are still not well understood.

Investigations of the impact of different atmospheric forcings are the present focus of model intercomparison studies such as the historical land simulations of the Land Surface, Snow, and Soil Moisture Model Intercomparison Project (LS3MIP; van den Hurk et al., 2016), where a number of LSMs were driven by four state-of-the-art forcing data sets and compared to observational data sets of, for example, soil moisture and snow cover. In particular, the propagation of uncertainties in atmospheric forcing is one topic to be addressed in LS3MIP. In this study, we conducted four sensitivity simulations using the LSM for Chinese Academy of Sciences (CAS-LSM) (Xie et al., 2018) driven by different atmospheric forcing data sets, that is, GSWP3, CRU-NCEP, WFD/WFDEI, and Princeton. We assessed the modeled SCF, SWE, and SDP using a set of observations, including the Moderate Resolution Imaging Spectroradiometer (MODIS)-based SCF product (Hall & Riggs, 2015), the Advanced Microwave Scanning Radiometer for EOS (AMSR-E)-derived SWE product (Tedesco et al., 2004), SWE from the Snowpack Telemetry (SNOWTET) network (Serreze et al., 1999; Sun et al., 2019; Yan et al., 2018), SDP from the China Meteorological Administration (CMA), and SDP from
the University of Arizona (UA; Broxton et al., 2019; Zeng et al., 2018). In addition to the detailed evaluation of our modeling results, our analysis allowed for a comparison between the four forcing data sets and the sensitivities of the simulated snow evolution to the forcing uncertainties in different snow climates.

The paper is organized as follows. Section 2.1 describes the model, while sections 2.2 and 2.3 describe the atmospheric forcing data sets and the validation data sets, respectively. Section 2.4 presents the experimental design, and section 2.5 explains the evaluation and analysis methods. Section 3.1 compares the four atmospheric forcing data sets, and section 3.2 presents an evaluation of the model simulations. Sections 3.3 presents the sensitivity analysis, and section 4 provides the conclusions and a discussion.

2. Data and Methods

2.1. Model Description

The CAS-LSM (Xie et al., 2018), which is the land component of the Flexible Global Ocean-Atmosphere-Land System Model for Chinese Academy of Sciences (CAS-FGOALS), has been unprecedentedly incorporated with three schemes of groundwater lateral flow (GLF) (Zeng et al., 2016), human water use (HWR) (Zeng et al., 2017; Zou et al., 2015) and frost and thaw fronts (FTFs) dynamics in soils (Gao et al., 2016, 2019), making it a potential tool for providing better understanding of land processes and land-atmosphere interactions, especially in cold and arid regions experiencing human interventions. The CAS-LSM was developed based on the CLM4.5 (Oleson et al., 2013), which simulates biogeophysical and biogeochemical processes such as soil hydrology, snow hydrology, momentum, and heat flux between the land and atmosphere, as well as the carbon and nitrogen cycles. On the basis of the groundwater model of CLM4.5, Zeng et al. (2016) derived a two-dimensional groundwater movement equation based on Darcy’s Law and the Dupuit approximation (Bear, 1972) to describe groundwater exchange between grid cells. Human water withdrawal included the processes of groundwater pumping and surface water intake. Gao et al. (2019) used the Stefan equation to obtain one-directional FTFs dynamics in a soil file (Jumikis, 1978), which is important for diagnosing the evolution of frozen soil and soil carbon-nitrogen processes.

The CAS-LSM uses a one-dimensional vertical, multilayer snow model that represents processes such as snow accumulation, melt, compaction, metamorphism, percolation, and the refreezing of water (Lawrence & Slater, 2009). The snowpack can have up to five layers depending on the total SDP. Because the processes of snowfall and snowmelt differ, Swenson and Lawrence (2012) proposed a new method that calculated the SCF separately for accumulation and ablation. When snowfall event occurs, the SCF is calculated as a function of the SWE:

\[
 f_{sno}^{n+1} = 1 - \left( 1 - \tanh \left( k_{\text{accum}} q_{\text{sno}} \Delta t \right) \left( 1 - f_{sno}^{n} \right) \right),
\]

where \( f_{sno}^{n+1} \) is the SCF from the previous time step, \( f_{sno}^{n+1} \) is the updated SCF, \( k_{\text{accum}} \) is a constant whose default value is 0.1, and \( q_{\text{sno}} \Delta t \) is the amount of newly fallen snow during the time interval \( \Delta t \). When snow ablation occurs, \( f_{sno} \) is calculated from the depletion curve:

\[
 f_{sno} = 1 - \left( \cos^{-1} \left( \frac{2R_{\text{sno}} - 1}{\pi} \right) \right)^{1/N_{\text{melt}}},
\]

where \( R_{\text{sno}} \) is the ratio of the current SWE, \( W_{\text{sno}} \), to the maximum accumulated SWE, \( W_{\text{max}} \), and \( N_{\text{melt}} \) is a parameter that controls the shape of the SCF, which is related to the topographic variability within the grid cell. The SDP in a grid cell is computed as a function of the SCF \( f_{sno} \) and the SWE \( W_{sno} \):

\[
 z_{sno}^{n+1} = z_{sno}^{n} + \Delta z_{sno},
\]

\[
 \Delta z_{sno} = \frac{q_{\text{grnd,ice}} \Delta t}{f_{sno} \rho_{sno}},
\]

where \( \rho_{sno} \) is the bulk density of the new snow (kg/m\(^3\)), which is calculated as (Anderson, 1976).
\[ \rho_{\text{snow}} = \begin{cases} 
50 + 1.7(17)^{1.5}, & T_{\text{atm}} > T_f + 2 \\
50 + 1.7(T_{\text{atm}} - T_f + 15)^{1.5}, & T_f - 15 < T_{\text{atm}} < T_f + 2 \\
50, & T_{\text{atm}} > T_f + 2 
\end{cases} \] 

(5)

where \( T_{\text{atm}} \) is the atmospheric temperature (K) and \( T_f \) is the freezing temperature of water (273.15 K).

Snow falling in forested areas is partitioned into interception by the canopies, throughfall to the ground, and drip from vegetation, which are calculated as functions of the leaf area index. Canopies shade underlying snow from both direct and diffuse solar radiation. The model uses a two-stream approximation (Dickinson, 1983; Sellers, 1985) for radiative transfer within vegetative canopies. Incoming longwave radiation to underlying snow is parametrized as the sum of radiation from the canopy and from the sky through canopy gaps. Again, this depends on leaf area index. Therefore, forest canopies modify surface energy fluxes, which, in turn, directly affect the snow distribution.

2.2. Atmospheric Forcing Data Sets

a. GSWP3

The GSWP3 provides atmospheric forcing data for the years 1901–2014 at 3-hourly resolution, which were dynamically downscaled to the T248 (~0.5°) grid from the Twentieth Century Reanalysis (20CR) (Compo et al., 2011) using a spectral nudging technique (Yoshimura & Kanamitsu, 2008) in a Global Spectral Model (GSM). Additionally, biases in the downscaled 20CR were corrected using observational data, subjected to vertically weighted damping (Hong & Chang, 2012) and single ensemble corrected (Yoshimura & Kanamitsu, 2013). Global Precipitation Climatology Centre (GPCC) Version 6 (for precipitation), Climatic Research Unit (CRU) Time Series (TS) 3.21 (for temperature), and Surface Radiation Budget project (SRB, for downward shortwave/longwave radiation) were used for bias correction. We used GSWP3, Version 1.0 (Kim, 2017) in this study.

b. WFD/WFDEI

The WATCH program created meteorological forcing data at 3-hourly, half-degree resolution covering the twentieth century (Weedon et al., 2011). The WATCH forcing data for 1958–2001 are based on the European Centre for Medium-range Weather Forecasts (ECMWF) ERA-40 reanalysis, and the data for 1901–1957 were generated using reordered ERA-40 data. The WFDEI forcing data are based on the same methodology as the WFD forcing, but using the ERA-Interim analysis data, which are available from 1979 to 2014 (Weedon et al., 2014). The 2-m air temperature, 2-m specific humidity, surface pressure, and downward longwave radiation were elevation corrected, and other fields (i.e., rainfall, snowfall, and shortwave radiation) were corrected using CRU monthly observations as well as GPCC precipitation monthly observations.

c. CRU-NCEP

The CRU-NCEP forcing data are based on a combination of two existing data sets: the 0.5° × 0.5° monthly CRU TS 3.2 (Mitchell & Jones, 2005) covering the years 1901–2002 and the 2.5° × 2.5° reanalysis products developed by the National Centers for Environmental Prediction (NCEP) covering the years 1948–2010 (Kalnay et al., 1996). For rainfall, air temperature, humidity, and solar radiation, we used the NCEP-generated diurnal and daily variability data, where the monthly means were corrected based the CRU climatology data. Other meteorological variables (i.e., surface pressure, longwave radiation, and wind speed) were directly interpolated from the NCEP data to 0.5° × 0.5° resolution. Here we used CRU-NCEP, Version 4 (Viovy, 2011).

d. The Princeton

The Princeton Global Forcing data set (Sheffield et al., 2006) was developed by combining the NCEP reanalysis with global observational data sets of precipitation, temperature, and radiation. The data currently (Version 2.2) are available from 1901 to 2014 at 1.0°, 0.5°, and 0.25° resolutions with 3-hourly time steps. The precipitation data were statistically downscaled in space based on the Global Precipitation Climatology Project (GPCP) daily product and disaggregated in time to 3-hourly intervals using the Tropical Rainfall Measuring Mission (TRMM) 3-hourly real-time data set. Other climate variables were
downscaled in space based on elevation. Precipitation was bias corrected to monthly CRU TS3.21 but was not undercatch corrected, and daily temperature was adjusted to match the CRU TS3.21 monthly data.

2.3. Validation Data

a. MODIS SCF Product

To validate the modeled SCF, we used the monthly average snow cover products of the 0.05° resolution MODIS Climate Modeling Grid (CMG), Version 6 (Hall & Riggs, 2015) product, which were derived from daily snow cover observations in the MODIS/Terra Snow Cover Daily L3 Global 0.05° CMG (MOD10C1) data set. To compensate for the impact of cloudiness, the Clear index (CI), which is defined as the percentage of clear sky, was used to filter the daily observations. Only grid cells with a CI greater than 70 were included in the average. MODIS/Terra data are available from 24 February 2000 to present, which were obtained from the National Snow and Ice Data Center (NSIDC) (https://nsidc.org/data/MOD10CM/versions/6). In this study, the data from 2004 to 2010 are used, which were aggregated to be consistent with the 0.9° × 1.25° CAS-LSM resolution.

b. AMSR-E SWE Product

We assessed the modeled SWE against the AMSR-E/Aqua monthly level-3 SWE products, Version 2 (Tedesco et al., 2004). The retrieval of the SWE was performed using the AMSR-E/Aqua L2A Global Swath Spatially Resampled Brightness Temperature unresampled data. SDP retrieval was first performed based on the method presented by Kelly et al. (2003), which used the differences of brightness temperature measured at different frequencies and which were then projected to the 25-km EASE Grid projection. The SWE was then calculated from the SDP using the global snow density monthly climatology map. The AMSR-E SWE data were obtained from the NSIDC (https://nsidc.org/data/ae_mosno/versions/2) from 19 June 2002 to 1 October 2011. In this study, we used the data from 2003 to 2010, which were aggregated to the CAS-LSM resolution of 0.9° × 1.25°.

c. SNOTEL In Situ SWE Data

As a further evaluation, we compared the modeled SWE with the SNOTEL (Serreze et al., 1999) data, which are point scale in situ SWE observations collected by the Natural Resources Conservation Service (NRCS) in the western United States, including Alaska. Pacific Northwest National Laboratory provided BCQC SNOTEL data (Sun et al., 2019; Yan et al., 2018), consisting of bias-corrected and quality-controlled daily SWE data up through 30 September 2018 for 829 active stations located in the western United States and Alaska. The data are available from the Pacific Northwest National Laboratory (https://dhsvm.pnnl.gov/bcqc_snotel_data.stm). In this study, a total number of 515 stations were screened with the criteria that each station has data available for at least 90% of the months from 2001–2010 and used over the 10-year study period.

d. CMA In Situ SDP Data

To evaluate the modeled SDP, the daily SDP data from 756 stations during January 1961 to December 2010 over China were used, as provided by the CMA (http://data.cma.cn/). We selected 559 stations for the years 2001–2010 after performing quality control procedures. The daily observations were averaged to monthly values. The SDP values of the model grid cells were interpolated to each station’s location using bilinear interpolation.

e. UA Gridded SDP data

We then used the University of Arizona (UA) daily 4-km SDP data over the conterminous United States (CONUS) from 1981 to 2017 for further evaluation of SDP (Broxton et al., 2019; Zeng et al., 2018). The data were created by assimilating in situ SDP observations from thousands of sites from the National Weather Service’s Cooperative Observer Program (COOP) network and SWE observations from the National Resources Conservation Service’s SNOTEL networks with modeled, gridded temperature and precipitation data from the Parameter elevation Regression on Independent Slopes Model (PRISM). The details of the methodology and the robustness and accuracy of the data set have been reported (Broxton, Dawson, & Zeng, 2016; Broxton, Zeng, & Dawson, 2016; Dawson et al., 2017, 2018). The data are available via NSIDC.
In this study, we used the data from 2001 to 2010, which were aggregated to the CAS-LSM resolution of 0.9° × 1.25°.

2.4. Experimental Design

The CAS-LSM was driven by four atmospheric forcing data sets, that is, GSWP3, WFD/WFDEI, CRU-NCEP, and Princeton, for the years 1901–2010 at a spatial resolution of 0.9° × 1.25°. Hereafter, the four simulations are referred to as S-GSWP3, S-WFDEI, S-CRUNCEP, and S-Princeton. We first recycled the climate mean and variability from two decades of forcing data (1901–1920) as the spin-up to achieve equilibrium. During spin-up, the land use, CO₂, and other forcings were held constant at the 1,850 levels. For the historical simulations, we used the Land Use Harmonized Version 2 (LUH2) transient data sets produced by the University of Maryland and the University of New Hampshire research groups, which are the product of the Land Use Model Intercomparison Project (LUMIP; https://cmip.ucar.edu/lumip) as part of the Coupled Model Intercomparison Project 6 (CMIP6). The land units in the LUH2 data were translated to the corresponding PFT mosaic in the CAS-LSM using the methods described by Lawrence et al. (2012). Notably, we only activated the HWR and FTFs modules in this work, since the GLF module needed rather high resolution. The model outputs for the years from 1981 to 2010 were used for result analyses.

2.5. Analysis Methods

We computed the mean bias, root-mean-square error (RMSE), and Pearson’s correlation coefficient (R) to compare the simulated SCF, SWE, and SDP with the satellite-based data and in situ observations. To evaluate the relative performance of the four simulations, we took the bias, RMSE, and R into account comprehensively. Based on the comparisons of the NH, we ranked the bias, RMSE, and R for four simulations from 1 to 4, where 1 represents the simulation with the lowest bias and RMSE, or the highest R, and 4 represents the largest bias and RMSE, or the lowest R. We then averaged the ranking values of bias, RMSE, and R, and the lowest (highest) value represents the best (worst) performance among the four simulations. The ranking algorithm was used in Wang and Zeng (2012) and Wang et al. (2016). We computed the climatological seasonal cycles of the SCF, SWE, and SDP to evaluate temporal variations. Because the quality of satellite-derived snow products tends to be affected by the presence of forests (Arsenault et al., 2014), our comparisons were further stratified by grid cells primarily covered or not covered by forests. Based on the PFTs fraction in the surface data derived from LUH2 transient data sets that were used to run the CAS-LSM, we stipulated that if the sum of the percentages of needleleaf forest and broadleaf forest in the grid cell was greater than the sum of the percentages of other land cover types (e.g., barren, shrublands, and grasslands), the grid cell was classified as a “forested” grid cell. Otherwise, the grid cell was classified as a “non-forested” grid cell (Figure S1 in the supporting information).

A sensitivity index (Ψ) was defined to quantify the sensitivity of the snow simulations to the atmospheric forcing uncertainties as follows:

\[ \psi_p = \frac{U_s}{U_p}, \quad \psi_t = \frac{U_s}{U_t}, \]

\[ U_x = \frac{1}{\Omega_x}, \]

\[ \Omega_x = \frac{m\sigma_x^2 - \bar{x}^2}{(m - 1)\sigma^2} \]

where \( U \) is an uncertainty index derived from the similarity index (\( \Omega \)) used by Koster et al. (2000). The subscript \( x \) corresponds to \( s \) (SCF, SWE, and SDP, respectively), \( p \) (precipitation), or \( t \) (air temperature); \( m \) is the number of ensemble members (\( m = 4 \) in this work); and \( \bar{x} \) denotes the ensemble average time series of the SCF, SWE, SDP, precipitation, or air temperature, which is computed as follows:

\[ \bar{x} = \frac{1}{m} \sum_{i=1}^{m} x_{ni}, \]

where \( n \) denotes the time period; \( \sigma_x^2 \) represents the variance of SCF, SWE, SDP, precipitation, or air temperature; and \( \sigma^2_\bar{x} \) denotes the variance of the ensemble average time series. Essentially, \( \Omega \) measures the...
“degree of similarity” in the time series of the considered snow variable between the members of the corresponding ensemble. If each member produces exactly the same time series, then $\sigma^2_x$ will be equal to $\sigma^2_x$; however, if the time series is completely independent, then $\sigma^2_b$ is expected to be approximately $\sigma^2_x/m$. Therefore, the value of $\Omega$ ranges from 0 to 1, where a value of 0 indicates that the four time series differ widely, and a value of 1 indicates that the four time series are identical to each other. Additionally, we adopted the reciprocal of the similarity index to quantify the uncertainties in the simulated SCF, SWE, and SDP due to different atmospheric forcing data sets.

The sensitivity index $\Psi$ measures the uncertainties in the ensemble of model simulations for snow variables (i.e., SCF, SWE, and SDP) to the uncertainties in the corresponding atmospheric forcing variables (i.e., precipitation and air temperature), which can be an indicator of the sensitivity of snow simulations to the atmospheric forcing data set. A similar approach was used to quantify the sensitivity of soil wetness to uncertainties in the used precipitation and radiation forcing data by Wei et al. (2008); however, we use $U$ here instead of the standard deviation.

3. Results

3.1. Intercomparison of the Four Atmospheric Forcing Data Sets

The CAS-LSM uses 0°C as a temperature threshold to partition the total precipitation into rainfall and snowfall; hence, both the total precipitation and the air temperature affect snowfall as well as snow simulation. Table 1 presents the mean annual air temperature, total precipitation, and snowfall of the four atmospheric forcing data sets during 1981–2010. The average temperatures over the NH were very similar for all forcings, since all of them were bias corrected to different versions of the CRU TS product. However, the total precipitation and snowfall of the four forcing data sets differed from each other. The largest mean annual precipitation and snowfall averaged over the NH came from WFDEI, followed by GSWP3, Princeton, and then CRU-NCEP. Figure 1 shows the differences of the mean annual precipitation and snowfall between GSWP3 and the other three data sets. Over most land areas, the total precipitation in WFDEI was slightly larger than that in GSWP3, except for Alaska, parts of South America and southern Asia. In contrast, the total precipitation in CRU-NCEP and Princeton was lower than that of GSWP3 globally, except for Greenland, central Asia, and parts of South America. Notably, the precipitation of GSWP3 and WFDEI was bias corrected to different versions of GPCC observations with different approaches to undercatch correction. The precipitation of CRU-NCEP and Princeton was bias corrected to monthly CRU TS3.21, but they were not undercatch corrected. In terms of snowfall, the annual mean difference between WFDEI and GSWP3 was positive over the NH, especially over Greenland and western parts of the TP. CRU-NCEP and Princeton show similar distributions with negative differences over middle-high latitudes over the NH, and the difference was greater between CRU-NCEP and GSWP3.

3.2. Evaluation of Model Simulations

a. SCF

Figure 2 shows the seasonal distributions of the SCF differences between MODIS and the four simulations forced by GSWP3, WFDEI, CRU-NCEP, and Princeton for the NH between 2004 and 2010. During the buildup phase (October–December, OND), the SCF was negatively biased in most of the land areas, except for at high latitudes of the eastern United States and northern Europe. Large negative biases were also seen in the TP. During the maximum snow accumulation period (January–March, JFM), there was no significant difference between MODIS and the simulated SCF at the high latitudes of Eurasia. The middle latitudes of eastern Asia, western United States, and Canada were dominated by negative deviations, while there were
larger positive biases at high latitudes in mid-eastern United States and northern Europe. From April to June, positive biases existed at high latitudes of Eurasia and Canada, and negative biases persisted in central Asia, the TP, and Alaska. Overall, the four simulations all tended to underestimate the SCF for most of the NH during the accumulation period and overestimate the SCF during the ablation period. Notably, the results show underestimations in areas with high topographic complexity for all seasons. The spatial distributions of the SCF differences between MODIS and the four simulations resembled each other, but...
differences were seen in mountainous regions with complex terrain, for example, western United States and the TP.

Table 2 displays the mean biases, RMSEs, and $R$ values for different regions of the NH, including Alaska, CONUS, Europe, the TP, western Siberia, and eastern Siberia (Figure 3). For the entire NH poleward of 30°N, the modeled SCF from the four simulations all agreed well with the MODIS-derived SCF, with an
average bias of $-1.29\%$ and a correlation coefficient of 0.77. Compared with the MODIS products, S-GSWP3 had a low negative bias of $-0.50\%$, while S-WFDEI had a low positive bias of 0.08\%. S-CRUNCEP and S-Princeton had relatively high biases of $-2.66\%$ and $-2.09\%$, respectively. In general, the four simulations showed good performances in the CONUS, Europe, western Siberia, and eastern Siberia. Large negative biases were found in Alaska, except for S-GSWP3, which has strong agreement with the MODIS product with a correlation coefficient of 0.94, accompanied by a low bias of $-0.95\%$. The model had poor performance in the TP, with an average bias (RMSE) of $-13.67\%$ ($29.60\%$) and the lowest correlation coefficient of 0.32. The large biases in these areas were attributed to the deficiencies of the snow process parameterization and uncertainties in the atmospheric forcing data sets. Snow redistribution due to blowing wind often occurs in mountains, which can result in large variations of snow cover; however, the CAS-LSM neglects any blowing snow processes. Also, the atmospheric forcing data were too coarse to resolve any subgrid effects (Toure et al., 2016). There was little difference between the modeled SCF from the four simulations over the NH, CONUS, Europe, and western Siberia. However, in Alaska and the TP, the results from the four simulations differed widely, especially in Alaska, since the total precipitation from the four atmospheric forcing data contains considerable differences in these regions. Generally, the SCF was more underestimated by S-CRUNCEP and S-Princeton than by S-GSWP3 and S-WFDEI.

In addition, we compared the simulated and MODIS SCF under different elevations and forest cover conditions. Figure 4 shows that elevation exerts little influence on the simulation of SCF at low altitudes, whereas the underestimations of SCF increase at high altitudes. The strongest correlation between underestimations

| Region       | Bias (%)  | RMSE (%) | R     |
|--------------|-----------|----------|-------|
| NH           | S-GSWP3   | S-WFDEI  | S-CRUNCEP | S-Princeton |
| Alaska       | $-0.50$   | 0.08     | $-2.66$ | $-2.09$ | 16.26 | 16.51 | 16.60 | 16.66 | 0.78 | 0.78 | 0.76 | 0.76 |
| CONUS        | 0.01      | 2.23     | $-0.25$ | $-0.58$ | 13.59 | 14.51 | 13.65 | 14.06 | 0.63 | 0.65 | 0.62 | 0.60 |
| Europe       | 4.21      | 2.97     | $-0.01$ | 1.81   | 19.92 | 18.15 | 17.95 | 18.28 | 0.70 | 0.73 | 0.72 | 0.70 |
| TP           | $-15.13$  | $-13.49$ | $-13.21$ | $-12.85$ | 26.39 | 30.41 | 30.47 | 31.11 | 0.38 | 0.32 | 0.28 | 0.31 |
| Western Siberia | 1.46    | 3.00     | 0.76   | 0.66   | 14.34 | 15.32 | 14.28 | 13.52 | 0.95 | 0.95 | 0.95 | 0.96 |
| Eastern Siberia | $-2.33$    | $-0.25$  | $-3.18$ | $-3.65$ | 13.97 | 13.85 | 15.00 | 15.02 | 0.95 | 0.96 | 0.95 | 0.95 |

Note. The NH includes land poleward of 30°N, excluding Greenland.
and elevations was seen in eastern Siberia (mean $R = 0.55$). We computed biases, RMSEs, and $R$ values between the MODIS and simulated SCF for grid cells dominantly covered by forests and for grid cells not dominantly covered by forests, respectively, over the NH (Table S1). Forested regions had larger negative biases than nonforested regions. Such results were seen from S-GSWP3, S-CRUNCEP, and S-Princeton. However, taking the larger RMSEs and lower $R$ values into consideration, it is not necessarily evident that the simulation of snow cover is worse in forested areas.

The mean seasonal cycles of the MODIS-based and the modeled SCF were compared for the years 2004–2010 (Figure 5a). Overall, the four simulations captured the seasonal variations of MODIS SCF well with peaks of SCF occurring in February. The models tended to underestimate the SCF in the fall, and the maximum differences exist in October. Compared to the SCF variations from S-CRUNCEP and S-Princeton, the seasonal variations from S-GSWP3 and S-WFDEI were closer to the MODIS observations. The average climatology of SCF from S-WFDEI (31.44%) was slightly larger than that seen in the MODIS observations (31.29%), while the results from S-GSWP3 (30.85%), S-CRU-NCEP (28.72%), and S-Princeton (29.29%) showed underestimations.

b. SWE

Figure 6 shows the seasonal distributions of the SWE differences between AMSR-E and the four simulations forced by GSWP3, WFDEI, CRU-NCEP, and Princeton for the NH between 2003 and 2010. During the snow accumulation period, the differences implied good agreement between AMSR-E and the models over the NH, with a slight overestimation in areas including western Siberia and northern Canada. The TP was dominated by negative biases. From January to March, the SWE was still positively biased in Canada and northwestern parts of Eurasia, where the average bias exceeded 60 mm. Large negative biases were found in Alaska and eastern Eurasia. During the ablation season, large positive biases of the SWE persisted in the
northern parts of Eurasia and Canada, but there was no significant difference between AMSR-E and the simulated SWE over most land areas except for the TP. Generally, for the entire NH, there was fairly good agreement between the simulated SWE and the AMSR-E product, with a mean correlation coefficient of 0.68, accompanied by a mean RMSE of 48.68 mm (Table 3). The positive biases of the four simulations suggested that, overall, the model tended to overestimate SWE. AMSR-E and CAS-LSM had the strongest correlations in western Siberia and eastern Siberia, though large biases were found in western Siberia. The TP had the lowest skill (mean $R = 0.37$), and S-WFDEI had the highest skill among the four simulations. However, the four simulations had a high correlation in Alaska (mean $R = 0.80$). The CONUS had modest performance overall (mean $R = 0.53$). Better performance was seen in Europe, with a mean correlation coefficient of 0.62. There was no significant difference between the modeled SWE from the four simulations over the NH, CONUS, Europe, and the TP. However, in Alaska and western Siberia, the values from S-GSWP3 and S-WFDEI were much larger than those from S-CRUNCEP and S-Princeton, which was consistent with the idea that the four atmospheric forcing data had considerable differences in their precipitation values in these two regions (Figure 1).

Figure 5b depicts the seasonal cycles of the AMSR-E-based and simulated SWE averaged over the NH for the period 2003–2010. On average, the model reproduced the seasonal variability of SWE well, though the four simulations showed overestimations. The biases of the four simulations were low during OND and increase rapidly from January to April. From June to September, the SWE from AMSR-E was nearly zero, but the simulated SWE was above 30 mm. The reason for this was that in the observational data, when snow was detected but is shallow, the SDP, from which the SWE was estimated, was set to 50 mm. Among the four simulations, the seasonal variations of SWE from the S-CRUNCEP simulation were closest to those seen in the AMSR-E products, since CRU-NCEP had the lowest values of precipitation.
It is worth noting that AMSR-E SWE performs poorly in forested regions (Foster et al., 2005). Although its algorithm compensates for forest effects and reduces the underestimation of the SWE in forests (Chang & Rango, 2000), the algorithm still cannot generate reasonable SWE values in forested areas. Dawson et al. (2018) demonstrated that the SWE from AMSR-E is underestimated in forested regions, whereas the SWE from nonforested areas is better. Table 4 shows that the four simulations had higher positive biases.

Figure 6. Differences in the seasonal average SWE (unit: mm) over the NH between AMSR-E and (a, b, c) S-GSWP3, (d, e, f) S-WFDEI, (g, h, i) S-CRUNCEP, and (j, k, l) S-Princeton. The left column (a, d, g, j) is the average for OND, the middle column (b, e, h, k) is the average for JFM, and the right column (c, f, i, l) is the average for AMJ. The averages are for the years 2003–2010.
and RMSEs in forested grid cells than nonforested grid cells. Though the AMSR-E SWE is not always reliable in forested areas, the results indicate that CAS-LSM suffers from uncertainties in forest snow processes. Our results indicate that the SWE from the four simulations was further underestimated at high elevations in mountainous regions (Figure S2). This may be partially attributed to the challenges of retrieving the SWE for AMSR-E in mountains.

Considering the limitations of the satellite-based SWE product described above, we also compared the simulated SWE with the point-scale SNOTEL SWE data. SNOTEL sites are located in open areas. As such, they are usually free from vegetation canopy. The mean seasonal cycles of the modeled SWE and those averaged over 515 SNOTEL stations were calculated for the 2001–2010 study period. Figure 5c shows large negative biases in the simulated SWE, especially from March to May. The S-GSWP3, S-CRUNCEP, and S-Princeton showed large underestimations, whereas the SWE from S-WFDEI was closer to that of the SNOTEL observations. The average bias and RMSE (Table 5) were −30.46 and 89.12 mm for S-WFDEI, respectively. These large biases are partially attributed to the heterogeneous distribution of snow within a grid cell with a resolution of 1° in mountainous areas, since the average elevation of the 515 SNOTEL stations was 2,152 m. Overall, the S-WFDEI and S-CRUNCEP performed well, with a correlation coefficient of 0.63, while S-GSWP3 (R = 0.59) and S-Princeton (R = 0.55) had lower correlations.

c. SDP

Next, we compared the modeled SDP against the monthly SDP observations in China. Figure 7 displays the RMSEs and R values between the observed and simulated SDP from S-GSWP3, S-WFDEI, S-CRUNCEP, and S-Princeton at each site in China based on the climatology of the SDP, where the colored patterns denoted the mean values of the RMSE and R within each SDP level. Generally, the R values increased as the SDP level increased, which meant that the model had better performance in areas with a large SDP. S-GSWP3 showed good performance with high R values and low RMSE values. S-CRUNCEP had the lowest skill, with an average R value of all 559 stations of 0.50 and an average RMSE of 2.42 cm (Table 6). Better performance was seen from the results of S-Princeton, which had the lowest RMSE value of 1.38 cm.

### Table 3

| Bias (mm) | RMSE (mm) | R |
|-----------|-----------|---|
|           | S-GSWP3   | S-WFDEI | S-CRUNCEP | S-Princeton | S-GSWP3   | S-WFDEI | S-CRUNCEP | S-Princeton |
| NH        | 26.82     | 29.67    | 8.58      | 20.40       | 54.88     | 57.71    | 35.60      | 46.56       |
| Alaska    | 17.48     | 7.52     | −24.00    | −11.90      | 53.78     | 64.48    | 41.96      | 45.66       |
| CONUS     | 9.93      | 14.90    | 7.43      | 7.71        | 24.26     | 32.14    | 20.14      | 21.00       |
| Europe    | 26.97     | 25.39    | 14.54     | 17.94       | 51.53     | 47.68    | 31.05      | 36.85       |
| TP        | −16.41    | −14.60   | −19.75    | −15.84      | 33.98     | 34.16    | 30.68      | 31.21       |
| Western Siberia | 58.95 | 59.80 | 31.52 | 34.44 | 89.33 | 89.64 | 51.29 | 56.86 |
| Eastern Siberia | 10.05 | 13.81 | −7.51 | 1.54 | 45.23 | 44.59 | 36.14 | 39.18 |

Note. The NH includes land poleward of 30°N, excluding Greenland.

### Table 4

| Bias (mm) | RMSE (mm) | R |
|-----------|-----------|---|
|           | Forest    | Nonforest | Forest | Nonforest | Forest | Nonforest |
| S-GSWP3   | 34.77     | 23.63      | 65.34  | 50.67      | 0.77   | 0.64       |
| S-WFDEI   | 40.38     | 25.37      | 72.76  | 51.67      | 0.77   | 0.65       |
| S-CRUNCEP | 20.77     | 3.70       | 48.49  | 30.43      | 0.77   | 0.63       |
| S-Princeton | 24.16 | 18.90 | 51.78  | 44.46      | 0.76   | 0.65       |
Table 5
Comparison of the Simulated SWE and SNOTEL Observations (Average of 515 Stations) for the Western United States and Alaska From 2001–2010

|        | Bias (mm) | RMSE (mm) | R        |
|--------|-----------|-----------|----------|
| S-GSWP3 | −48.58    | 97.25     | 0.59(503) |
| S-WFDEI | −30.46    | 89.12     | 0.63(498) |
| S-CRUNCEP | −50.96  | 92.13     | 0.63(492) |
| S-Princeton | −55.88 | 96.99     | 0.55(499) |

*Note.* The values in the parentheses are the number of stations with a correlation coefficient that is statistically significant at the $p = 0.05$ level.

Figure 7. Scatterplots of the RMSEs and $R$ values between the observed and simulated SDP from (a) S-GSWP3, (b) S-WFDEI, (c) S-CRUNCEP and (d) S-Princeton, at 559 stations in China for 2001–2010. The scatterplots are categorized based on the average SDP from 2001–2010.
Table 6
Summary of the Statistics Between the Simulated SDP Versus CMA Observations at 559 Stations in China and the Simulated SDP Versus UA Product for CONUS From 2001–2010

|                   | CMA (China)          | UA (CONUS)         | UA (western US) | UA (mid-eastern US) |
|-------------------|----------------------|--------------------|----------------|---------------------|
|                   | Bias (cm) | RMSE (cm) | R     | Bias (cm) | RMSE (cm) | R     | Bias (cm) | RMSE (cm) | R     | Bias (cm) | RMSE (cm) | R     |
| S-GSWP3           | 0.50      | 1.43      | 0.64 (513) | 1.80    | 7.97      | 0.74 | 2.62   | 8.47       | 0.68 | 4.15    | 9.01      | 0.66 |
| S-WFDEI           | 0.75      | 2.25      | 0.62 (516) | 1.80    | 7.97      | 0.74 | 3.4   | 9.95       | 0.69 | 3.28    | 7.15      | 0.77 |
| S-CRUNCEP         | 0.04      | 2.42      | 0.50 (510) | -0.19   | 5.43      | 0.66 | 1.99   | 9.13       | 0.63 | 0.96   | 3.34      | 0.68 |
| S-Princeton       | 0.63      | 1.38      | 0.55 (502) | -0.31   | 5.91      | 0.60 | 2.88   | 8.80       | 0.60 | 1.34   | 4.38      | 0.62 |

*Western US = western United States; mid-eastern US = mid-eastern United States. The values in the parentheses are the number of stations with a correlation coefficient that is statistically significant at $p = 0.05$. 

Figure 8. Percentage differences in the annual average (a, d, g) SCF, (b, e, h) SWE, and (c, f, i) SDP over the NH between (a–c) S-WFDEI and S-GSWP3, (d–f) S-CRUNCEP and S-GSWP3, and (g–i) S-Princeton and S-GSWP3, respectively. The averages are for the years 1981–2010.
While the seasonal variability of the simulated SDP resembles that of the CMA, the average climatology of the simulated SDP was larger than that of the observations. The mean biases averaged over the 559 stations are 0.50, 0.75, 0.04, and 0.63 cm for the four simulations, respectively (Table 6). For S-GSWP3, S-WFDEI, and S-Princeton, overestimations mainly occurred during the accumulation season from October to March, and the peak differences occurred in January. For S-CRUNCEP, the model overestimated the SDP during OND, with the maximum difference occurring in November, while underestimating it from JFM.

In addition, we compared the simulated SDP to the UA SDP product over CONUS. The mean seasonal cycles of the modeled SDP and those from the UA were displayed for the 2001–2010 study period (Figure 5e). Overall, the four simulations reproduced the seasonal variations of the SDP well, with peaks of the SDP occurring in February. The SDP from S-GSWP3, S-CRUNCEP, and S-Princeton was close to that of UA, with a bias of −0.14, −0.19, and −0.31 cm, respectively (Table 6). During the snow accumulation period, the three simulations overestimated the SDP, while underestimating it during the melting season. The four simulations tended to underestimate the SDP in the western United States with lower correlations, where simulations of snow in complex terrain are not accurately represented in the model (Table 6). Similar to the SCF, the SDP was more underestimated at high altitudes (Figure S3). We also compared the UA and simulated SDP for forested and nonforested areas (Table S2), and the results showed that the SDP is underestimated in forested regions, whereas the SDP from nonforested areas is better.

**3.3. Sensitivity of the Snow Simulations to the Different Atmospheric Forcing Data Sets**

Differences in the SCF, SWE, and SDP between S-GSWP3 and the other three simulations can help determine the simulation sensitivity to the selection of forcing data sets. Figure 8 shows the percentage differences of the annual average SCF, SWE, and SDP (1981–2010) between S-GSWP3 and three other simulations. Compared with S-GSWP3, S-CRUNCEP and S-Princeton underestimated the SCF over most of the land areas at middle latitudes. This is consistent with the negative differences of snowfall seen over middle-high latitudes over the NH (Figure 1). S-WFDEI overestimated the SCF over CONUS, with underestimations seen over middle latitudes of Eurasia; these trends are consistent with the spatial patterns of the snowfall in Figure 1 indicating that the annual mean difference between WFDEI and GSWP3 was positive over the most lands of the NH, with the exception over middle latitudes of Eurasia and Alaska. The differences of the SWE and SDP between S-WFDEI and S-GSWP3 mean showed similar patterns with the differences of the SCF, but S-WFDEI showed more overestimations of the SWE and SDP at high latitudes. In contrast, the S-CRUNCEP and S-Princeton means underestimated the SWE and SDP over most land areas, with larger differences seen over middle-low latitudes. On average, the difference of SCF between S-GSWP3 and other three simulations ranged from −7.29% (S-CRUNCEP) to 1.65% (S-WFDEI), while the difference of SWE had a range of −41.42% (S-CRUNCEP) to 14.91% (S-WFDEI), and the difference of SDP had a range of −29.86% (S-CRUNCEP) to 11.62% (S-WFDEI).
To estimate the uncertainties in the simulated SCF, SWE, and SDP among the four simulations, we further computed the uncertainty index ($U$) of the monthly mean SCF, SWE, and SDP (Figure 9). Low values of the uncertainty index of SCF over high latitudes demonstrated high similarity between the SCF time series from the four simulations. There were strong inconsistencies in low latitudes regions where snow cover was sparse, especially in eastern Asia, but the values of $U$ in parts of China were relatively low. The uncertainty index of SWE showed strong spatial heterogeneity, and its relatively high values indicated great disparities in the simulated SWE from the four simulations. Large inconsistencies were found in middle-low latitudes regions, as well as the TP and Alaska. The uncertainty index of SDP showed a similar pattern with that of the SWE, but it had lower values. The differences between the $U$ values of the SCF, SWE, and SDP indicated that the uncertainties in the atmospheric forcing data sets propagated differently through the land-only simulations.

Figure 10. Sensitivities of the (a–c) SCF, (d–f) SWE, and (g–i) SDP to uncertainties in precipitation in three seasons. The left column (a, d, g) is computed for OND, the middle column (b, e, h) is computed for JFM, and the right column (c, f, i) is computed for AMJ.
The differences in the simulated SCF, SWE, and SDP were mainly related to differences in the total precipitation and air temperature, which determined the snowfall, and in turn directly affected the snow simulations. Figure 10 depicts the sensitivities of SCF, SWE, and SDP to the uncertainties in the precipitation from OND, JFM, and AMJ. It can be seen that the sensitivity of the snow to the precipitation uncertainties had various seasonal cycles, depending on the climate regime. During the snow buildup phase (OND), high sensitivity of the SCF to the precipitation uncertainties mainly existed in warm temperate climates, such as in eastern China and central United States. This is related to their relatively dry winter climates. Uncertainties in the precipitation had little effect on the SCF at middle-high latitudes, whereas the SWE and SDP showed relatively strong sensitivity in southern Russia. During the season of maximum snow cover (JFM), the SWE and SDP exhibited strong sensitivity to precipitation uncertainties in cold, dry climates, particularly in eastern Russia. Notably, uncertainties in the precipitation led to larger variability in the SWE and

Figure 11. The same as Figure 10 but for sensitivities to uncertainties in air temperature.
SDP than SCF. In the snow melting season (AMJ), the SCF, SWE, and SDP all showed high sensitivities to uncertainties in the precipitation at middle latitudes except for western parts of the TP, where the uncertainties in the precipitation and simulated snow cover were all very high, and the response of snow in mountainous regions can be complicated by vertical gradients in precipitation and temperature (Brown & Mote, 2009). The reinforced sensitivity in the spring period may be related to stronger albedo feedback (Déry & Brown, 2007).

The sensitivities of the SCF, SWE, and SDP to uncertainties in the air temperature in OND, JFM, and AMJ are also shown in Figure 11. During the snow accumulating phase (OND), the stronger sensitivities of the SCF, SWE, and SDP to temperature uncertainties in warm temperate climates indicated that small changes in temperature generated large changes in the partitioning of precipitation into rainfall and snowfall. In the season of maximum snow cover (JFM), strong sensitivity to air temperature was seen in the NH, particularly in Siberia, implying the importance of air temperature in the formation and maintenance of snow cover. The stronger sensitivity in parts of eastern Siberia can be in part attributed to the complex topography of the mountainous areas. Snow cover in drier, colder continental interior regions tended to display lower temperature sensitivity. However, tundra snow-climate regions, such as northern Eurasia, also showed high sensitivity. During AMJ, the air temperature at high latitudes with thick snowpack was not high enough to generate snowmelt-off; hence, they displayed relatively low sensitivity. However, the SCF, SWE, and SDP were all very sensitive to uncertainties in the air temperature at middle latitudes with shallow snowpack, because snowmelt was mainly controlled by energy availability. Generally, the SWE and SDP were more sensitive to uncertainties in the precipitation and air temperature than the SCF, which were more closely controlled by snowfall amount. Hence, by comparing the relative strength of the precipitation and air temperature sensitivity, we found that sensitivity to air temperature uncertainties was dominant in relatively cold, humid climates, such as in western Eurasia, while in temperate climates with cold, dry winters, both of them were high.

4. Conclusions and Discussion

In this study, the sensitivity of snow simulations to different atmospheric forcing data sets was investigated. Four atmospheric forcing data sets, that is, GSWP3, WFDEI, CRU-NCEP, and Princeton, were used to drive the CAS-LSM. Apart from CRU-NCEP, the other three forcing data sets had nearly identical mean air temperatures. The differences in precipitation and snowfall among the four data sets varied globally and were particularly evident over some regions. On average, WFDEI had the largest precipitation and snowfall, followed by GSWP3, Princeton, and CRU-NCEP.

Compared with the MODIS SCF product, the CAS-LSM simulations captured the spatial and seasonal variation of the SCF well over the NH. Nevertheless, the model tended to slightly underestimate the SCF over most land areas. This disparity was alleviated significantly when the model was driven by GSWP3 and WFDEI. The CAS-LSM SWE agreed reasonably well with the AMSR-E product, although there were relatively large discrepancies over high latitudes, especially during the period of maximum snow cover. Compared with the SNOWTEL in situ data, the model had a tendency to underestimate the SWE in mountainous areas. Comparisons of in situ SDP observations and UA products showed that the model tended to overestimate the SDP in China and CONUS. Overall, the simulation based on GSWP3 produced more accurate results than the other simulations.

### Table 7

|     | SCF | SWE | SDP | SCF | SWE | SDP | SCF | SWE | SDP | SCF | SWE | SDP | Average |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|---------|
| S-GSWP3 | 2 | 3 | 1.5 | 1 | 3 | 2 | 1 | 2 | 1.5 | 1.33 | 2.67 | 1.67 |
| S-WFDEI | 1 | 4 | 4 | 2 | 4 | 3.5 | 2 | 1 | 1.5 | 1.67 | 3 | 3 |
| S-CRUNCEP | 4 | 1 | 1.5 | 3 | 1 | 2.5 | 3 | 4 | 3.5 | 3.33 | 2 | 2.5 |
| S-Princeton | 3 | 2 | 3 | 4 | 2 | 2 | 4 | 3 | 3.5 | 3.67 | 2.33 | 2.83 |

Note: The rankings for SCF and SWE are based on the evaluations over the NH (Tables 2 and 3), and the rankings for SDP are the average of the rankings based on the evaluations over China and CONUS (Table 6). The values of the last three columns are the average of bias, RMSE, and \( R \) rankings for SCF, SWE, and SDP, respectively. The lowest value represents the best performance among the four simulations in each column.
reasonable estimates than the ones based on WFDEI, CRU-NCEP, and Princeton, particularly for the SCF and SDP (Table 7). The mean differences in the SCF, SWE, and SDP varied regionally between the GSWP3-based simulation and the other three simulations. Compared with the GSWP3-based simulation, CRUNCEP-based and Princeton-based simulations underestimated the SCF, SWE, and SDP over most land areas in middle latitudes, while the WFDEI-based simulation overestimated the SCF, SWE, and SDP over North America and underestimated these values over middle latitudes of Eurasia. The SCF over high latitudes demonstrated small disparities between the SCF time series from the four simulations, while the SWE and SDP showed greater diversity due to the uncertainties in the atmospheric forcing data sets.

The sensitivities of the SCF, SWE, and SDP to uncertainties in the precipitation and air temperature had various seasonal cycles, depending on the climate regime. The highest sensitivity in boreal climates appeared during the season of maximum snow cover (JFM), while the highest sensitivity in warm temperate climates existed in the snow melting season (AMJ). On average, regions with cold, dry winters showed higher sensitivity to uncertainties in precipitation, such as eastern Eurasia. The sensitivities of the SCF, SWE, and SDP to uncertainties in air temperature showed similar patterns with those to the precipitation uncertainties; however, sensitivity to air temperature uncertainties was dominant in cold, humid regions with extensive winter snowfall, while in regions with cold, dry winters, both of them were high. Considering that the actual uncertainties in precipitation are much greater (Figure 1 and Table 1), our study suggests that the meteorological observational network for precipitation should be strengthened to improve the quality of precipitation data.

There are a few limitations of this study that should be mentioned. The index, $U$, which was derived from the similarity index ($\Omega$), puts emphasis on temporal coherency more than the mean and temporal variances of the considered time series (Wang et al., 2007; Wei et al., 2012), which meant that indices that emphasize different aspects might lead to different sensitivity analysis results to some extent. We focused mainly on the precipitation and air temperature in this study; however, uncertainties in snow simulation are also affected by uncertainties in other forcing fields, such as radiation fluxes, humidity, and wind speed, which remain to be addressed. Additionally, the results presented here were based on one particular LSM, and other models may produce different responses. Nevertheless, our results are useful to help understand the impact of uncertainties in atmospheric forcing data on the diversities of snow simulations and highlighted that reliable forcing data, especially precipitation data, are of great significance to the accuracy of model outputs.

Data Availability Statement

The MODIS snow cover fraction data were downloaded from https://nsidc.org/data/MOD10CM/versions/6 website. The AMSR-E snow water equivalent data were downloaded from https://nsidc.org/data/ae_mosno/versions/2 website. The snow depth data of the University of Arizona were obtained from https://nsidc.org/data/NSIDC-0719/versions/1 website. The bias-corrected and quality-controlled snow water equivalent data were obtained from the Pacific Northwest National Laboratory (https://dhsvm.pnl.gov/bcqc_snater_data_stm). The snow depth in situ observations in China are available online (http://data.lasg.ac.cn/wyan/CMA/). We would like to thank the two reviewers for their helpful comments on the manuscript.

References

Anderson, E. A. (1976). A point energy and mass balance model of a snow cover. NOAA Technical Report NWS 19, Office of Hydrology, National Weather Service, Silver Spring, MD. Arsenault, K. R., Houser, P. R., & De Lannoy, G. J. M. (2014). Evaluation of the MODIS snow cover fraction product. Hydrological Processes, 28(3), 980–998. https://doi.org/10.1002/hyp.9636 Bear, J. (1972). Dynamics of fluids in porous media (pp. 361–373). N.Y: American Elsevier. Brown, R. D., & Mote, P. W. (2009). The response of Northern Hemisphere snow cover to a changing climate. Journal of Climate, 22(8), 2124–2145. https://doi.org/10.1175/2008jcli2665.1 Broxton, P., Dawson, N., & Zeng, X. (2016). Linking snowfall and snow accumulation to generate spatial maps of SWE and snow depth. Earth and Space Science, 3, 246–256. https://doi.org/10.1002/2016EA000174 Broxton, P., Zeng, X., & Dawson, N. (2019). Why do global reanalyses and land data assimilation products underestimate snow water equivalent? Journal of Hydrometeorology, 17(11), 2743–2761. https://doi.org/10.1175/JHM-D-16-0056.1 Broxton, P., Zeng, X., & Dawson, N. (2019). Daily 4 km gridded SWE and snow depth from assimilated in-situ and modeled data over the conterminous US, version 1. [Indicate subset used]. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. https://doi.org/10.5067/0GGPB220EX6A. [Date Accessed] Brutel-Vuilmet, C., Ménégoz, M., & Krinner, G. (2013). An analysis of present and future seasonal Northern Hemisphere land snow cover simulated by CMIP5 coupled climate models. The Cryosphere, 7(1), 67–80. https://doi.org/10.5194/tc-7-67-2013
Wang, M. S., Lundquist, J. D., & Clark, M. P. (2015). Exploring the impact of forcing error characteristics on physically based snow simulations within a global sensitivity analysis framework. *Hydrology and Earth System Sciences*, 19(7), 3153–3179. https://doi.org/10.5194/hess-19-3153-2015

Robinson, D., & Frei, A. (2000). Seasonal variability of Northern Hemisphere snow extent using visible satellite data. *The Professional Geographer*, 52(2), 307–315. https://doi.org/10.1111/0033-0124.00226

Rutter, N., Essery, R., Pomeroy, J., Altimir, N., Andreadis, K., Baker, I., et al. (2009). Evaluation of forest snow processes models (SnowMIP). *Journal of Geophysical Research*, 114, D06111. https://doi.org/10.1029/2008jd011063

Schmucki, E., Marty, C., Fierz, C., & Lehning, M. (2014). Evaluation of modelled snow depth and snow water equivalent at three contrasting sites in Switzerland using SNOWPACK simulations driven by different meteorological data input. *Cold Regions Science and Technology*, 99, 27–37. https://doi.org/10.1016/j.coldregions.2013.12.004

Sellers, P. J. (1985). Canopy reflectance, photosynthesis and transpiration. *International Journal of Remote Sensing*, 6(8), 1335–1372. https://doi.org/10.1080/01431168508948283

Serreze, M. C., Clark, M. P., Armstrong, R. L., McGinnis, D. A., & Prowse, R. S. (1999). Characteristics of the western United States snowpack from snowpack telemetry (SNOWTEL) data. *Water Resources Research*, 35(7), 2145–2160. https://doi.org/10.1029/1999WR90090

Sheffield, J., Goteti, G., & Wood, E. F. (2006). Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. *Journal of Climate*, 19(13), 3088–3111. https://doi.org/10.1175/jcli3790.1

Sun, N., Yan, H., Wigmosta, M. S., Leung, L. R., Skaggs, R., & Hou, Z. (2019). Regional snow parameters estimation for large-scale land hydrological applications in the Western United States. *Journal of Geophysical Research: Atmospheres*, 124, 5296–5313. https://doi.org/10.1029/2018JD030140

Swenson, S. C., & Lawrence, D. M. (2012). A new fractional snow-covered area parameterization for the Community Land Model and its effect on the surface energy balance. *Journal of Geophysical Research*, 117, D21107. https://doi.org/10.1029/2012jd018178

Tedesco, M., Kelly, R., Foster, J. L., & Chang, A. T. (2004). AMSR-E/aqua monthly L3 global snow water equivalent EASE-grids, version 2. [Subset optical祾eous](https://doi.org/10.5067/AMSR-E/A/E_MCSNO.002). Accessed [Date Accessed].

Toure, A. M., Rodell, M., Yang, Z. L., Beaudouin, H., Kim, E., Zhang, Y. F., & Kwon, Y. W. (2016). Evaluation of the snow simulations from the Community Land Model, version 4 (CLM4.5). *Journal of Hydrometeorology*, 17(1), 153–170. https://doi.org/10.1175/JHM-D-14-0165.1

van den Hurk, B., Kim, H., Kriener, G., Seneviratne, S. I., Derksen, C., Okie, T., et al. (2016). L3SMP (v1.0) contribution to CMIP6: The land surface, snow and soil moisture model Intercomparison project—Aims, setup and expected outcome. *Geoscientific Model Development*, 9(8), 2809–2832. https://doi.org/10.5194/gmd-9-2809-2016

Vivyan, N. (2011). CRUNCEP dataset. Available at: ftp://nacp.orst.gov/synthesis/2009/frescati/temp/land_use_change/original/README.htm

Wang, A., & Zeng, X. (2012). Evaluation of multireanalysis products with in situ observations over the Tibetan Plateau. *Journal of Geophysical Research: Atmospheres*, 117, D05102. https://doi.org/10.1029/2011JD016553

Wang, A., Zeng, X., & Guo, D. (2016). Estimates of global surface hydrology and heat fluxes from the Community Land Model (CLM4.5) with four atmospheric forcing datasets. *Journal of Hydrometeorology*, 17(9), 2493–2510. https://doi.org/10.1175/jhm-d-16-0041.1

Wang, G., Kim, Y., & Wang, D. (2007). Quantifying the strength of soil moisture–precipitation coupling and its sensitivity to changes in surface water budget. *Journal of Hydrometeorology*, 8(5), 551–570. https://doi.org/10.1175/jhm-d-07-0133.1

Weedon, G. P., Balsamo, G., Beluésin, N., Gomes, S., Best, M. J., & Viterbo, P. (2014). The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data. *Water Resources Research*, 50, 7505–7534. https://doi.org/10.1002/2014wr015838

Wendron, G. P., Gomes, S., Viterbo, P., Shuttleworth, W. J., Blyth, E., Osterle, H., et al. (2011). Creation of the WATCH Forcing Data System for use to assess global and regional reference crop evaporation over land during the twentieth century. *Journal of Hydrometeorology*, 12(5), 823–848. https://doi.org/10.1175/2011jhm1369.1

Wei, J., Dimeyer, P. A., & Guo, Z. (2008). Sensitivities of soil wetness simulation to uncertainties in precipitation and radiation. *Geophysical Research Letters*, 35, L15703. https://doi.org/10.1029/2008GL034494

Wei, J., Dimeyer, P. A., Guo, Z., & Zhang, L. (2012). Impact of atmospheric variability on soil moisture-precipitation coupling. In Abdel Hannachi Climate variability – some aspects, challenges and prospects (pp. 17–36). InTechOpen. https://doi.org/10.5772/29443

Xie, Z., Hu, Z., Xie, Z., Jia, B., Sun, G., Du, Y., & Song, H. (2016). Impact of the snow cover scheme on snow distribution and energy budget modeling over the Tibetan Plateau. *Theoretical and Applied Climatology*, 131(3–4), 951–965. https://doi.org/10.1007/s00704-016-2020-6

Xie, Z., Liu, S., Zeng, Y., Gao, J., Qin, P., Jia, B., et al. (2018). A high-resolution land model with groundwater lateral flow, water use, and soil freeze-thaw front dynamics and its applications in an endorheic basin. *Journal of Geophysical Research: Atmospheres*, 123, 7204–7222. https://doi.org/10.1002/2018JD028369

Yan, H., Sun, N., Wigmosta, M., Skaggs, R., Hou, Z., & Leung, R. (2018). Next-generation intensity-duration-frequency curves for hydrologic design in snow-dominated environments. *Water Resources Research*, 54, 1093–1108. https://doi.org/10.1002/2017WR021290

Yoshimura, K., & Kanamitsu, M. (2008). Dynamical global downscaling of global reanalysis. *Monthly Weather Review*, 136(6), 2893–2998. https://doi.org/10.1175/2008wmrr2281.1

Yoshimura, K., & Kanamitsu, M. (2013). Incremental correction for the dynamical downscaling of ensemble mean atmospheric fields. *Monthly Weather Review*, 141(9), 3087–3101. https://doi.org/10.1175/mwr-d-12-00271.1

Zakharova, E. A., Kouraev, A. V., Biancamaria, S., Kolmakova, M. V., Mognard, N. M., Zemtsov, V. A., et al. (2011). Snow cover and spring flood in the northern part of Western Siberia (the Poluy, Nadym, Pur, and Taz Rivers). *Journal of Hydrometeorology*, 12(6), 1498–1511. https://doi.org/10.1175/jhm-d-11-0171.1

Zeng, X., Broxton, P., & Dawson, N. (2018). Snowpack change from 1982 to 2016 over Conterminous United States. *Geophysical Research Letters*, 45, 12,940–12,947. https://doi.org/10.1002/2018GL079621

Zeng, Y., Xie, Z., Yu, Y., Liu, S., Wang, L., Zou, J., et al. (2016). Effects of anthropogenic water regulation and groundwater lateral flow on land processes. *Journal of Advances in Modeling Earth Systems*, 8, 1106–1113. https://doi.org/10.1002/2016ms006646

Zeng, Y., Xie, Z., & Zou, J. (2017). Hydrologic and climatic responses to global anthropogenic groundwater extraction. *Journal of Climate*, 30(1), 71–90. https://doi.org/10.1175/Jcli-d-16-0209.1

Zhang, T. (2005). Influence of the seasonal snow cover on the ground thermal regime: An overview. *Reviews of Geophysics*, 43, RG4002. https://doi.org/10.1029/2004rg000157
Zou, J., Xie, Z., Zhan, C., Qin, P., Sun, Q., Jia, B., & Xia, J. (2015). Effects of anthropogenic groundwater exploitation on land surface processes: A case study of the Haihe River Basin, northern China. *Journal of Hydrology*, 524, 625–641. https://doi.org/10.1016/j.jhydrol.2015.03.026

Zuo, Z., Zhang, R., Wu, B., & Rong, X. (2011). Decadal variability in springtime snow over Eurasia: Relation with circulation and possible influence on springtime rainfall over China. *International Journal of Climatology*, 31(9), 1336–1345. https://doi.org/10.1002/joc.2355