Evaluation of the Snow Simulations from the Community Land Model, Version 4 (CLM4)

ALLY M. TOURE,*,+ MATTHEW RODELL,* ZONG-LIANG YANG,# HIROKO BEAUDOING,*,@ EDWARD KIM,* YONGFEI ZHANG,# AND YONGHWAN KWON#

* *Hydrological Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, Maryland*
+ *Universities Space Research Association, Columbia, Maryland*
# *Department of Geological Sciences, Jackson School of Geosciences, The University of Texas at Austin, Austin, Texas*
@ *Earth System Science Interdisciplinary Center, University of Maryland, College Park, College Park, Maryland*

(Manuscript received 8 September 2014, in final form 17 August 2015)

ABSTRACT

This paper evaluates the simulation of snow by the Community Land Model, version 4 (CLM4), the land model component of the Community Earth System Model, version 1.0.4 (CESM1.0.4). CLM4 was run in an offline mode forced with the corrected land-only replay of the Modern-Era Retrospective Analysis for Research and Applications (MERRA-Land) and the output was evaluated for the period from January 2001 to January 2011 over the Northern Hemisphere poleward of 30°N. Simulated snow-cover fraction (SCF), snow depth, and snow water equivalent (SWE) were compared against a set of observations including the Moderate Resolution Imaging Spectroradiometer (MODIS) SCF, the Interactive Multisensor Snow and Ice Mapping System (IMS) snow cover, the Canadian Meteorological Centre (CMC) daily snow analysis products, snow depth from the National Weather Service Cooperative Observer (COOP) program, and Snowpack Telemetry (SNOTEL) SWE observations. CLM4 SCF was converted into snow-cover extent (SCE) to compare with MODIS SCE. It showed good agreement, with a correlation coefficient of 0.91 and an average bias of -1.54×10^2 km^2. Overall, CLM4 agreed well with IMS snow cover, with the percentage of correctly modeled snow–no snow being 94%. CLM4 snow depth and SWE agreed reasonably well with the CMC product, with the average bias (RMSE) of snow depth and SWE being 0.044 m (0.19 m) and −0.010 m (0.04 m), respectively. CLM4 underestimated SNOTEL SWE and COOP snow depth. This study demonstrates the need to improve the CLM4 snow estimates and constitutes a benchmark against which improvement of the model through data assimilation can be measured.

1. Introduction

Seasonal snow plays an important role in the earth's climate system and hydrological processes (Dewey 1977; Kukla 1981; Cohen and Rind 1991; Cohen 1994). About 98% of the seasonal snow is located in the Northern Hemisphere (NH), and about 42% of the land in the NH is covered by snow for a significant duration (Dingman 2002). On such a large scale, snow influences the surface energy and water balances (Karl et al. 1993; Bamzai and Shukla 1999). On the regional scale, the snow variability in space and time affects weather patterns (Cohen and Entekhabi 1999; Chapin et al. 2005; Turner and Slingo 2011). Snow also constitutes a valuable water resource for hundreds of millions of people. Accurate estimation of snow-cover states is crucial for improving initial conditions for numerical weather forecasts, climate prediction model simulation, water resources management, flood preparedness, and mitigation. It is also an important part of the investigation of snow albedo feedbacks and climate change.

Clearly, it is essential that models used to simulate global climate and hydrological processes accurately describe the evolution of seasonal snowpack. Increasingly complex climate models incorporate sophisticated snow schemes to improve the simulation of interactions between the land and the atmosphere. The Community Earth System Model (CESM; Vertenstein et al. 2012; Gent et al. 2011), developed at the National Center for Atmospheric Research (NCAR), is a prominent example. CESM consists of five separate models, which

Corresponding author address: Matthew Rodell, Hydrological Sciences Laboratory, NASA Goddard Space Flight Center, Code 617, 8800 Greenbelt Rd., Greenbelt, MD 20771.
E-mail: matthew.rodell@nasa.gov

DOI: 10.1175/JHM-D-14-0165.1

© 2016 American Meteorological Society

simultaneously simulate the earth's atmosphere, ocean, land, land ice, and sea ice. A central coupler coordinates the models and allows information and fluxes to be exchanged between them.

Land surface models (LSMs) suffer from one of or a combination of the following: 1) errors in forcing data, 2) improper model parameters, 3) simplified physical processes, and 4) simplified numerical solutions or methods. The performance of the snow scheme in the Community Land Model, version 4 (CLM4), which is the land component of CESM, is unknown because there has not been a comprehensive validation against observations. A few studies have investigated some aspects of the CLM4 snow model, including the snow-cover fraction (SCF) estimates, the snow albedo parameterizations, and the water budget. Swenson and Lawrence (2012) demonstrated that the parameterization used to determine SCF-based snow depth in CLM4 exhibits a bias toward early melt when compared to satellite-observed SCF, and they proposed a new SCF parameterization as a function of snow water equivalent (SWE) instead. The results showed an improvement of the surface energy budget in snow-covered areas. Thackeray et al. (2014) recently showed that the weak simulated snow albedo feedback over the boreal forest was attributable to a poor parameterization of the CLM4 mechanism of snow removal from the forest canopy. CLM4 was also assessed against its older version (CLM3.5) (Lawrence et al. 2011). Improvements in soil water storage, evapotranspiration, surface albedo, permafrost, and land air temperature were observed. Li et al. (2011) evaluated the performance of CLM4 runoff simulation at the catchment scale and found the temporal variability of modeled runoff to be abnormally large. They proposed increasing the spatial resolution and adjustments to the model parameter values. While it is possible to resolve some of the model's problems by increasing the resolution and complexity, assimilation of multiple datasets can be a more optimal way to improve model prediction.

The purpose of this study is to evaluate CLM4 snow output in preparation for a subsequent study on merging snow-related observations from multiple satellite sensors within CLM4 using data assimilation. We developed and implemented a technique for optimally merging snow-related observations from multiple observing systems, based on multivariate data assimilation within the CLM4. The Moderate Resolution Imaging Spectroradiometer (MODIS) SCF, the Advanced Microwave Scanning Radiometer for Earth Observing System (AMSR-E) radiance, and the Gravity Recovery and Climate Experiment (GRACE) terrestrial water storage changes observations will be assimilated simultaneously to generate spatially and temporally continuous offline global SWE fields, at resolutions as high as 1/8°. Each satellite observation will be assimilated where it is more likely to have the best impact in improving CLM4.

We ran CLM4 in an offline (uncoupled to the atmosphere) mode at ~0.5° resolution forced by the meteorological data from the land-only replay of the Modern-Era Retrospective Analysis for Research and Applications (MERRA-Land; Reichle et al. 2011; Reichle 2012) from January 2001 to January 2011 over the NH poleward of 30°N. We assessed the model SCF, snow depth, and SWE using various observational data, including the MODIS snow cover (Hall et al. 2002), the Canadian Meteorological Centre (CMC) daily snow analysis products (Brasnett 1999; Brown et al. 2003; Brown and Brasnett 2010), the Interactive Multisensor Snow and Ice Mapping System (IMS) snow-cover product (Ramsay 1998; Helfrich et al. 2007), snow depth from the National Weather Service (NWS) Cooperative Observer (COOP) program, and the Snowpack Telemetry (SNOTEL) snow depth and SWE. Our analysis will serve as a benchmark for evaluating multisensor data assimilation results, and it will identify locations where data assimilation may potentially improve modeled snow.

The paper is organized as follows. Section 2a describes CLM4, section 2b explains the experimental design, and section 2c describes the validation datasets. Section 3 presents the results and the analysis, and section 4 is dedicated to the major conclusions and discussion for future studies.

2. Data and methods

a. Model description

CLM4 describes the physical, chemical, biological, and hydrological processes by which terrestrial ecosystems interact with climate across a variety of spatial and temporal scales. Subgrid heterogeneity is represented by fractional covers of glacier, lake, wetland, urban, and vegetated land. The vegetated portion of a grid cell is further subdivided into a mosaic of plant functional types (pfts), each with its own leaf and stem area index and canopy height. Modeled hydrological processes include canopy interception; throughfall; stemflow; evapotranspiration; infiltration; runoff; and soil moisture, groundwater, and snowpack evolution (Oleson et al. 2008). CLM4 has a one-dimensional vertical, multilayer snow model (Dai et al. 2003; Kluzek 2012) based on the parameterizations developed by Anderson (1976), Jordan (1991), and Dai and Zeng (1997). Using grid-averaged atmospheric forcing, CLM4 simulates processes such as snow accumulation, depletion, densification, metamorphism, and percolation and refreezing of water (Lawrence and Slater 2010).

TABLE 1. Summary of the datasets and land surface models used in the study.

Dataset	Description	Resolution	Domain	Parameters
CLM4	Snow model with up to five layers	0.5° × 0.67° converted into CMC resolution for comparison	NH	SCF, snow depth, and SWE
MODIS	*Terra* MODIS CMG level-3 snow-cover products (MOD10C1)	0.05° × 0.05° aggregated into 0.5° × 0.67°	NH	SCF
IMS	Binary snow product based on satellite observations	~0.25° × 0.25°	NH	SCF
CMC	Snow depth reanalysis SWE estimates based on Sturm et al. (2010) parameterization	~0.25° × 0.25°	NH	Snow depth and SWE
MERRA-Land	Three-layer snow model	0.5° × 0.67° converted into CMC resolution for comparison	NH	SCF, snow depth, and SWE
SNOTEL	Point-scale in situ SWE, most at 2000–3000 m altitude	702 stations data used	United States	SWE
COOP	Point-scale in situ snow depth most at low altitude	1797 stations data used	United States	Snow depth

The number of snow layers can reach five depending on the depth of the snowpack. Each layer is characterized by its thickness, ice mass, liquid water content, temperature, and effective grain radius. Snow can also exist in the model without being represented by explicit snow layers, which occurs when the snow depth is less than 0.01 m. A snow layer is initialized when the snow depth is greater than or equal to 0.01 m and, depending on whether there is accumulation or depletion of snow, snow layers are combined or subdivided. Liquid water from rain or snowmelt remains within a snow layer until the layer's water holding capacity is reached and the excess water is allowed to percolate to the underlying layer. The percolating water at the base of the snowpack becomes runoff or infiltration depending on the permeability of the soil (see Beven and Kirkby 1979; Niu et al. 2005; Niu and Yang 2006). Snow densification is a result of processes including destructive and constructive metamorphism, such as overburden pressure (see Flanner and Zender 2006) and melt–freeze cycle due to the presence of liquid water in the snow (see Brun 1989). One of the major new capabilities of the snow model is the representation of the evolution of snow microphysical structure (effective grain size) as a function of temperature, layer temperature gradient, density, and initial size distribution (see Flanner and Zender 2006). Snow effective grain size is critical for determining snow albedo, the depth of radiative penetration, and the influence of impurities, such as black carbon and mineral dust on the albedo. Snow albedo is calculated as a function of snow aging, effective grain size, solar angle, and pollution (Warren and Wiscombe 1980; Flanner and Zender 2005, 2006; Flanner et al. 2007). SCF is calculated as a function of snow density and snow depth as follows (Niu and Yang 2007):

$$f_{\mathrm{sno}} = \tanh\left\{\frac{Z_{\mathrm{sno}}}{2.5Z_{0m,g}[\min(\rho_{\mathrm{sno}}, 800)/\rho_{\mathrm{new}}]^{m}}\right\}, \quad (1)$$

where Z_{sno} is the depth of snow (meters), $Z_{0m,g} = 0.01$ is the momentum roughness length for soil (meters), $\rho_{\mathrm{new}} = 100\,\mathrm{kg\,m^{-3}}$ is the density of new snow, and $m = 1$ is used for global applications. Snow density is calculated as follows:

$$\rho_{\mathrm{sno}} = W_{\mathrm{sno}}/Z_{\mathrm{sno}}, \quad (2)$$

where W_{sno} is the SWE ($\mathrm{kg\,m^{-2}}$).

b. *Experimental design*

The CLM4 simulation was forced with atmospheric data from MERRA-Land (Reichle et al. 2011; Reichle 2012). To create the initial conditions, we ran CLM4 for 600 years driven by repeating 25-yr (1979–2004) MERRA meteorological forcing, with the satellite-based plant phenology, land use, aerosol and nitrogen deposition, and the CO_2 level for year 2000 conditions as a reference case created at NCAR. The resulting restart file was then used in the subsequent 1979–2012 simulation using the same initial atmospheric and satellite-derived phenological composition set as for the spinup process.

c. *Validation datasets*

Table 1 is a summary of the models and datasets used in this study.

1) MODIS SNOW-COVER OBSERVATIONS

To assess the CLM4 SCF output, we used the 0.05° *Terra* MODIS climate-modeling grid (CMG) level-3 snow-cover products (MOD10C1; Hall et al. 2006a,b). The MOD10C1 data are created by assembling daily tiles of 500-m resolution (MOD10A1) and binning them

into a CMG cell. The MOD10A1 tile snow-cover data are based on a snow-mapping algorithm using the Normalized Difference Snow Index (NDSI) and other criteria tests (see Hall et al. 2006a). The CMG product includes a daily confidence index (CI) field, which is defined as the percentage of clear-sky land observed for the day. The MOD10C1 snow-cover products have been evaluated extensively against independent satellite and ground-based datasets for a wide variety of landscapes and have shown good agreement with the in situ observations (Bitner et al. 2002; Klein and Barnett 2003; Maurer et al. 2003; Rango et al. 2003; Brubaker et al. 2005; Hall and Riggs 2007; Riggs and Hall 2012). The products are available from 24 February 2000 to present via the National Snow and Ice Data Center (NSIDC) website. In this study, we used the data poleward of 30°N from 2001 to 2010 (10 years). We aggregated MOD10C1 to coincide with 0.5° × 0.67° (latitude × longitude) CLM4 grid resolution, taking into account only CMG cells with a daily percentage of snow greater than 0 and CI greater than 20%. The daily SCF are then averaged into monthly SCF in order to minimize the cloud-cover effect while accommodating data gaps.

2) CMC DAILY SNOW ANALYSIS PRODUCT

CMC snow analysis consists of the NH snow depth data (Brasnett 1999; Brown et al. 2010). The snow depth is generated based on a 6-hourly optimal interpolation of an extensive in situ snow depth report from the World Meteorological Organization (WMO) information system. The first-guess field is obtained through a simple snow model driven with 6-hourly meteorological forcing from the Canadian forecast model (Brasnett 1999). CMC data have a horizontal resolution of approximately 24-km in a polar stereographic grid. The grid closely approximates what is used by NOAA for the 24-km daily IMS snow product (see description below). The observations are available via NSIDC from 1998 through 2014. In this study, we used daily data for the period from 1 January 2001 to 31 December 2010. The CMC snow depth data were converted to SWE estimates (see Forman et al. 2012) using Sturm et al. (2010) snow densities. Considered to be the best available snow observations for the NH, CMC snow depth and SWE data are frequently relied upon to evaluate modeled snow (e.g., Su et al. 2010; Reichle et al. 2011; Forman et al. 2012).

3) IMS SNOW-COVER PRODUCT

The IMS is a binary snow (snow–no snow) product of NOAA's National Environmental Satellite, Data, and Information Service (NESDIS; Ramsay 1998; Helfrich et al. 2007). IMS provides estimates of daily NH snow cover at an approximate 24-km resolution in a polar stereographic projection. The IMS snow product uses a temporally variable and complementary array of data such as 1) satellite imagery including the Polar Operational Environmental Satellites (POES), the Advanced Very High Resolution Radiometer (AVHRR), the Geostationary Operational Environmental Satellite (GOES), the Special Sensor Microwave Imager (SSM/I), the Geostationary Meteorological Satellite (GMS), and the European Weather Satellite (Meteosat); 2) derived snow products generated from the Advanced Microwave Sounding Unit (AMSU), the AMSR-E, the SSM/I, the National Centers for Environmental Prediction (NCEP) models, and the United States Air Force Snow and Ice Analysis; and 3) in situ observations. These data are subject to extensive quality control during the generation of IMS product though intercomparison of these multiple sources of data. A quality-control module in the data processing algorithm is aimed at reducing uncertainties due to the limitations of satellite data collection such as the dense cloud cover, thick forest canopy, or the difficulties of discriminating snow cover from other similar-looking features such as cloud. MODIS is also used, but was not part of the original satellite observation used as input source data for IMS. MODIS is used as auxiliary data to enhance IMS starting in February 2004, and its use only involved the MODIS channel 1, while the MOD10C1 snow-cover retrieval is based on the NDSI, which uses channels 4 and 6 of MODIS observations. Hence, IMS and MODIS SCF are considered as independent in this study. IMS is used operationally as input data in numerical weather forecast models run by the NWS as well as several other governmental agencies. The data are available via the NSIDC website from 1997 to present. In this study, we used the data from 1 January 2001 to 31 December 2010.

4) IN SITU DATA

(i) SNOTEL data

The SNOTEL (Serreze et al. 1999) data are point-scale in situ snow depth and SWE observations collected by the Natural Resources Conservation Service (NRCS) from a network of about 813 automated snow stations in 13 states in the western United States, including Alaska. Most sites are located in 2000–3000-m elevation regions. All SNOTEL sites measure SWE using a pressure-sensing snow pillow and report on a daily basis at 0800 UTC. The data are available from the NSIDC website. For a detailed description of the stations and the data, see Serreze et al. (1999). Although there are scale and spatial sampling issues with SNOTEL network data, their data remain the only reliable in situ SWE observations available for model validations. SNOTEL has been

successfully used for regional-scale SWE analysis (Simpson et al. 1998; Serreze et al. 1999; Peterson et al. 2000; Moser et al. 2011) and validation of LSMs (Pan et al. 2003; Clow et al. 2012). In this study, a total number of 702 stations were screened and used over the 10-yr study period.

(ii) COOP

The COOP weather- and climate-monitoring network consists of more than 10 000 stations across every state and territory in the United States and provides measurements made by volunteers over a wide range of landscapes, including national parks, coastal, mountainous, and agricultural areas, as well as urban and suburban areas. At nearly 8000 of those sites, daily surface temperature, precipitation, snowfall, and snow depth are measured. The data are subject to comprehensive quality-control methods (see Durre et al. 2010) to ensure consistency between snow and other meteorological observations. The data are available via NOAA's National Centers for Environmental Information (NCEI) website (http://www1.ncdc.noaa.gov). The data were screened (as described in section 2d) and a total of 1808 stations for the 10-yr period of 2001–10 were used in this study.

5) LAND SURFACE DATA PRODUCT OF MERRA-LAND

The MERRA (Rienecker et al. 2011) forcing data are generated by the NASA Global Modeling and Assimilation Office (GMAO) using the Goddard Earth Observing System, version 5 (GEOS-5), and it is associated with the Data Assimilation System (DAS), version 5.2.0 (Rienecker et al. 2008). The forcing fields are available hourly at horizontal resolution of 0.5° × 0.67° (latitude × longitude) and 72 vertical levels from the surface to 0.01 hPa (Rienecker et al. 2011) and include precipitation, solar radiation, pressure, specific humidity, air temperature, and wind speed. The MERRA precipitation field used in this study is corrected with the NOAA Climate Prediction Center Unified (CPCU) daily precipitation product, which was obtained from the MERRA-Land dataset [for more details, see Reichle (2012)].

MERRA-Land uses the latest version of Catchment land surface model (Koster et al. 2000; Ducharne et al. 2000). The Catchment model includes a three-layer snow module that incorporates snow physics including densification, snowmelt, and refreeze (Stieglitz et al. 2001). When snowfall occurs on bare soil, the snow is initially spread out over a fraction of the catchment so that at any given point the SWE is equal to a constant value known as minimum water equivalent (WEMIN). In the MERRA-Land GEOS, version 5.7.2 (GEOS-5.7.2), WEMIN was changed from 13 to 26 mm. Other parameters also changed [see Reichle et al. (2011), their Table 2], including the capacity of canopy interception reservoir, the areal fraction of canopy leaves onto which large-scale precipitation falls, the areal fraction of canopy leaves onto which convective precipitation falls, and the maximum depth of the uppermost snow layer. Detailed explanations on the improvements are found in Reichle et al. (2011). The data are available online through the Goddard Earth Sciences (GES) Data and Information Services Center (DISC) (http://disc.sci.gsfc.nasa.gov/mdisc) from 1 January 1980 to present. In this study, we used daily averages of snow outputs (snow depth and SWE) at a resolution of 0.5° × 0.67° (latitude × longitude) from 1 January 2001 to 31 December 2010.

d. Validation metrics

We computed several metrics to assess CLM4 snow outputs using the data from independent observations described above. We excluded the summer months (July–September) from the computation of the metrics and required a minimum 200 days across the 10-yr daily time series (~5%). Standard statistical techniques including mean bias, regular and normalized root-mean-square error (RMSE), Pearson correlation coefficient R, and Nash–Sutcliffe model efficiency (NSCE) were used to evaluate the CLM4 SCE against MODIS observations. To assess CLM4 snow depth and SWE using CMC, SNOTEL, and COOP data, we computed the mean bias, RMSE, and climatological seasonal cycles over the 10-yr period. Anomaly time series were also calculated by subtracting from each time series its mean seasonal cycle. We then computed the anomaly correlation between the model anomaly time series and corresponding observation anomaly time series. Correlation between MERRA-Land and CMC was also computed and statistical significance of the difference between them was computed using the Fisher r-to-z transformation test. The standard probabilistic verification scores (Wilks 2006) were used to compare IMS SCF against the model. Because IMS snow cover is a binary (snow–no snow) type product, the model fractional snow cover was first converted into binary snow cover. Frequency bias B, probability of detection (POD), false alarm rate (FA), and proportion correct (PC) were calculated using the contingency table elements (a, b, c, and d) defined in Table 2.

The frequency bias estimates compare the frequency of the model estimates of presence or absence of snow to the observation and do not measure the model's accuracy. The proportion correct provides information about the accuracy of binary estimate. It ranges from 0% to 100%, with a value equal to 100% representing the perfect score. The probability of detection represents a

TABLE 2. Contingency table of observations vs model showing the count of snow–no snow events pairs, where *a* represents the hits, *b* represents the false alarms, *c* represents the misses, and *d* represents the correct negative or negative hits.

		MODIS observations	
		Snow	No snow
Model	Snow	*a*	*b*
	No snow	*c*	*d*

ratio of snow presence prediction to the total number of snow observation. As in PC, POD ranges from 0% to 100%, with 100% being the perfect score:

$$B = (a + b)/(a + c), \quad (3)$$

$$\mathrm{POD} = a/(a + c), \quad (4)$$

$$\mathrm{FA} = b/(b + d), \quad \text{and} \quad (5)$$

$$\mathrm{PC} = (a + d)/(a + b + c + d). \quad (6)$$

3. Results

a. Comparison against MODIS product

Figure 1 shows the time series of CLM4 and MODIS SCE observations for the entire Northern Hemisphere. Table 3 summarizes the statistical comparison between CLM4 and MODIS SCE for the entire NH and selected regions with complex topography where CLM4 tends to underestimate SCE. The values are averaged over the 10-yr study period (2001–11). CLM4 SCE exhibits relative bias [relative bias = (CLM4 SCE/MODIS SCE) − 1] of −0.07 and normalized RMSE (NRMSE = RMSE/range of SCE) of 0.60. CLM4 SCE correlates well with MODIS SCE with *R* of 0.91 and NSCE of 0.93. The low values of bias and NRMSE combined with the high values of *R* and NSCE, and skill score >0.5 suggests that overall the model is able to quantify SCE accurately. During the snow accumulation period, the model closely resembles the MODIS observations with a slight underestimation. Overall, other than during the months of December and January when SCE is overestimated, CLM4 underestimates SCE, particularly during the melting season (Figs. 1b, 2b) because of early snowmelt. The mean seasonal cycles of CLM4 and MODIS SCE (Fig. 2) are almost identical with peak SCE occurring in February. The maximum SCE value is about $38 \times 10^6\,\mathrm{km}^2$, which is smaller than what has been reported in Clifford (2010) for MODIS. An explanation could be that MODIS grid cells with confidence index less than 20 were not taken into consideration for the computation of SCF. This filtering was done to avoid cases where the retrieval algorithm could be mistaking cloud cover for snow or vice versa. We computed the metrics for areas where the model tends to underperform (Table 3) and found that the continental United States has the largest relative bias (−0.58), while the smallest relative bias was in western Siberia (−0.13). Time series of relative bias (figure not shown here) indicate that in areas characterized by complex topography, the model consistently

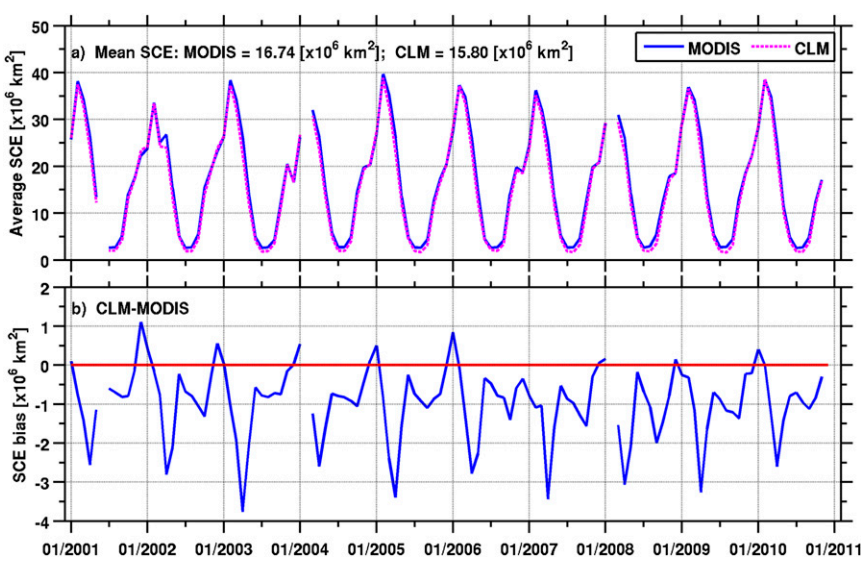

FIG. 1. Time series of monthly estimates of simulated CLM4 and observed MODIS snow area extent ($10^6\,\mathrm{km}^2$) for the NH poleward of 30°N. (a) The blue solid line is the MODIS estimate, and the dashed magenta line is the model estimate. (b) The blue line represents the difference between CLM4 and MODIS SCE.

TABLE 3. Comparison of MODIS and CLM4 SCE for the entire NH and selected regions of the globe.

	Bias (10^3 km^2)	Relative bias	RMSE (10^3 km^2)	NRMSE (to data range)	NSCE	Skill score	R
NH	−1.54	−0.07	15.62	0.60	0.83	0.60	0.91
CONUS	−2.89	−0.54	10.63	1.46	0.29	0.19	0.55
Midwestern United States	−1.71	−0.51	7.43	1.68	0.29	0.20	0.63
Rocky Mountains, United States	−4.95	−0.58	13.08	1.44	0.13	0.08	0.50
Western Siberia	−0.61	−0.13	8.68	0.82	−0.10	0.06	0.53

underestimates SCE during the ablation season. Relative bias averaged over the entire NH tends to oscillate between positive and negative, with negative relative bias occurring during the melting season.

To further the understanding of the spatial distribution of the CLM4 error against MODIS SCF and the model behavior during the snow buildup, the maximum snow accumulation, and the depletion phase (Fig. 3), we computed the difference between CLM4 and MODIS maximum snow-cover extent for October–December (OND), January–March (JFM), and April–June (AMJ). During the snow buildup phase (OND), SCF is positively biased in high latitudes in the eastern United States, the northern part of the Rocky Mountains, the southern edge of the snow line in Eurasia, and the Tibetan Plateau (TP), while the southern part of the Rocky Mountains and the Great Plains are dominated by negative bias. During the period of maximum snow cover (JFM), the snow cover in the Arctic is at 100% and unbiased. Positive biases persist in the areas with relatively high topographic variability (the Rocky Mountains and southeastern Siberia). During the melting season (AMJ), negative biases dominate the snow cover except for eastern Europe and the Rocky Mountains. Overall, in the fall the model tends to overestimate SCF while in the spring the model underestimates SCF. Swenson and Lawrence (2012) investigated this model behavior and found a problem in the parameterization of the model SCF. The CMC snow analysis data used to establish Eq. (1) tend to melt early compared to MODIS observations. Furthermore, the parameterization of SCF was derived from the AVHRR observations and the CMC snow depth analysis, both at 1° × 1° spatial resolution and monthly temporal resolution. A 1-month time period is too short compared to a typical snow accumulation period and too long to accurately describe the snowmelt. The accumulation phase of snow is longer (from the beginning of fall to late January in many areas) than the ablation period, which lasts only for a few weeks (from April to May). The same parameterization used for accumulation is also used for the ablation even though it does not reflect the hysteresis in SCF for a given snow depth between accumulation and melt phases. All of these factors contribute to CLM4 SCF overestimating snow amount in fall and underestimating it in the spring. The discrepancies between MODIS and model SCF could in part come from the averaging of model gridcell SCF used in the comparison. The averaging does not capture the subgrid spatial variability of the SCF. Subgrid distribution of snow is difficult to model accurately.

Forest regions tend to be more error prone in satellite-derived snow products (Arsenault et al. 2014). Thus, we used global forest-cover fraction (ff) data derived from MODIS observations as a mask, and computed relative bias, NRMSE, and R between CLM4 and MODIS SCE for areas with different forest-cover fractions. In areas where ff $<$ 0.25, the relative bias (NRMSE) and R are −0.23 (1.34) and 0.76, respectively. Forest-cover increase (0.25 $<$ ff $<$ 0.5) leads to increase in relative bias (NRMSE) to −0.30 (1.36) and decrease in R to 0.68. For ff $>$ 0.75, R decreases to just 0.57 with a relative bias of −0.36. However, it is not necessarily evident that the performance of CLM4

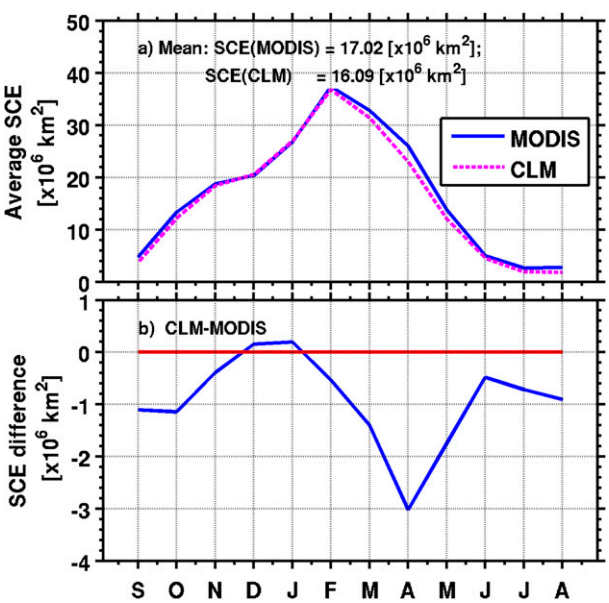

FIG. 2. For the NH poleward of 30°N (a) average SCE (10^6 km^2) from CLM4 and MODIS (2001–10) and (b) SCE difference (CLM4 − MODIS; 10^6 km^2). The red solid line in (b) indicates CLM − MODIS = 0.

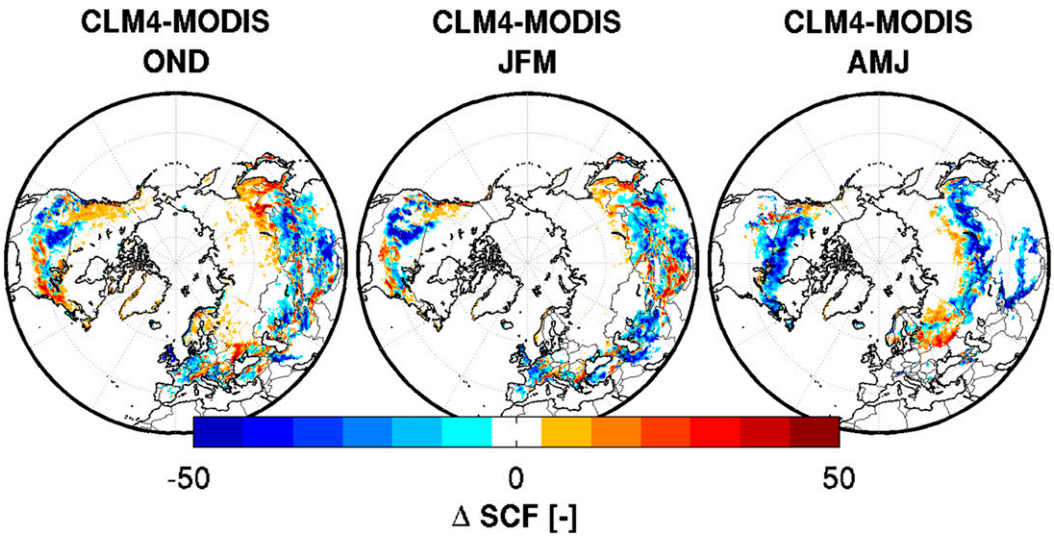

FIG. 3. Spatial distribution of the difference (CLM4 − MODIS) for SCF in percentage for OND, JFM, and AMJ from 2001 to 2010.

worsens when ff increases. While it is possible that the MODIS snow-cover fraction is not always reliable in areas with dense forest cover, the extent and nature of this error is not known quantitatively.

b. *Comparison against IMS product*

Because cloud cover affects MODIS daily snow-cover observations, it is very rare to have a daily MODIS snow map with a complete global coverage. IMS snow product constitutes an alternative that has the advantage of globally complete coverage. To compare IMS with CLM4 SCF, a threshold of 25% SCF was used to convert CLM4 SFC into binary data. The choice of the threshold was arbitrary. The value was varied between 50% and 25% and the changes did not significantly affect the result of the conversion. We applied Eqs. (3)–(6) to compute frequency bias, POD, FA, and PC of IMS versus CLM4 SCF. The computed statistics are presented in Fig. 4. The mean frequency bias (Fig. 4a) of 85% indicates that the model underestimates the NH snow cover compared to IMS. As for the temporal distribution of frequency bias, the model overestimates snow cover in October, agrees well from November to January, and underestimates from February to September. The PC (see Fig. 4b) shows high accuracy of CLM4 prediction with a seasonal mean of 94%. The higher values are found in the summer because of inclusion of hit negative (d in Table 2) in the computation of PC. Unlike PC, POD (Fig. 4c) does not include d in the calculation [Eq. (4)]. In general, there is a strong agreement between CLM4 and IMS (a seasonal average of 76%). The POD is stronger (>0.80) during the winter and weaker during the spring. FA has a negative orientation where FA = 0% represents a perfect score. Figure 4d shows that the FA is very low (a seasonal average of 3%); it increases during the accumulation phase, reaches the maximum in February at ~6%, and decreases during the depletion phase. The temporal evolution of FA is consistent with that of PC, with the lower value of FA corresponding to the higher value of PC. The low values of FA are an indication of the fact that the model underestimates IMS.

Figures 5a–h are examples of the spatial distribution pattern of hit, FA, hit negative, and misses of the model estimates compared to IMS snow cover for the snow season 2006/07. They represent the snow difference on the first day of the month of October (Fig. 5a), November (Fig. 5b), December (Fig. 5c), January (Fig. 5d), and February–May (Figs. 5e–h). Overall, the model agrees reasonably well with IMS. During the snow accumulation phase, FA (error of commission) tends to dominate and misses (error of omission) tend to occur in the Rocky Mountains, in parts of the U.S. Great Plains, and over the TP (Figs. 5a–c). From December to the end of the snow season (Figs. 5c–h), FA tends to dominate. This is because the model snow line moves earlier compared to IMS, which could be an indication of model tendency toward early snowmelt. IMS product was also found to detect snow when MODIS does not, especially during the melting season (Frei and Lee 2010). The misses in Figs. 5e–h could be in part explained by the bias in IMS SCF. In fact, Frei and Lee (2010) found that in June IMS detects 40% more snow than MODIS, especially in northern Canada (Nunavut and Northwest Territories). The most pronounced discrepancies among CLM and MODIS, IMS, and CMC occur in AMJ during the ablation season. The results are consistent with Frei and Lee (2010) when they

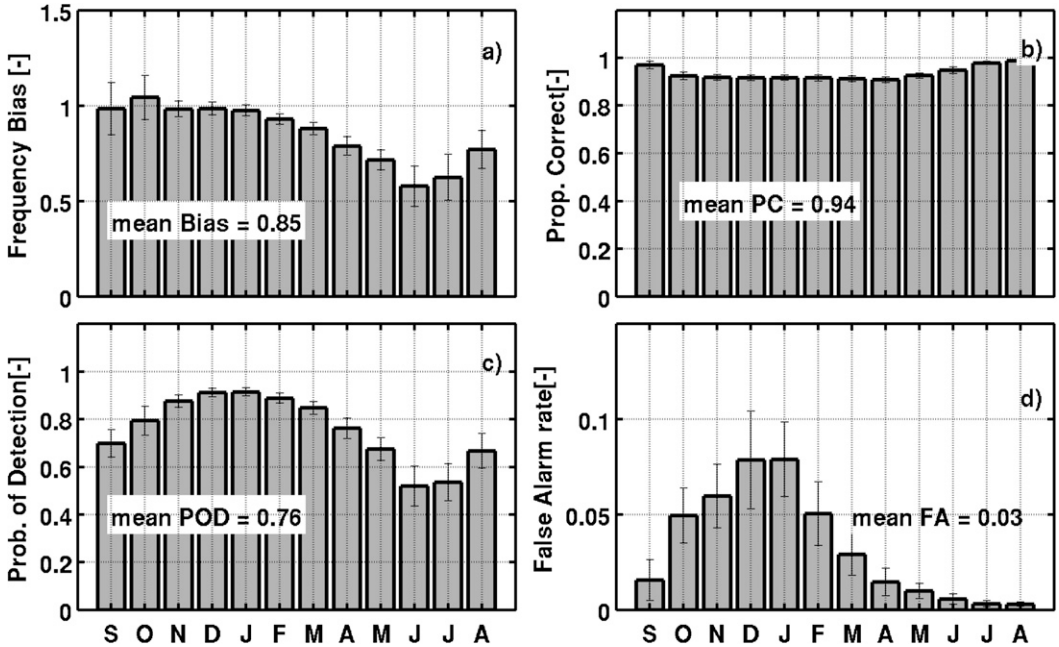

FIG. 4. Comparison of CLM4 binary SCF estimates (snow–no snow) and IMS for the NH poleward of 30°N (2001–10): (a) frequency bias, (b) PC, (c) POD, and (d) FA.

compared IMS to MODIS and CMC snow cover during AMJ over North America. The results were also confirmed by Brown et al. (2010) in a pan-Arctic study of the SCE involving IMS, MODIS, and CMC. They showed that the largest SCE in May was given by IMS, followed by MODIS and then CMC. The use of CMC monthly snow depth in the parameterization of CLM4 SCF makes the model also susceptible to the same kind of error as in CMC, as discussed in section 3a.

COMPARISON AGAINST CMC PRODUCT

We assessed CLM4 snow depth and SWE against the CMC snow products. Figure 6 represents the station distribution of bias, RMSE, and anomaly correlation. Positive high biases are found in the high latitudes and in the Rocky Mountains. Relatively deep snow cover and long cold winters characterize these areas. The model snow depth estimates agree reasonably well with the CMC reanalysis data in the southern part of the Canadian Prairies as well as in the U.S. Great Plains, western Europe, and southern Russia. Large negative biases are found in the TP. The area-weighted-average bias is -0.044 m. The same pattern also can be seen for RMSE (Fig. 6b) with an average RMSE of 0.19 (Table 4). The average anomaly correlation of snow depth is 0.47. Areas of northern Canada, Siberia, and the TP have very low skill ($R < 0.1$) as shown in Fig. 6c. Many factors could explain these discrepancies, among them the scarcity of observations used in the production of CPCU precipitation. Also, the quality of the CMC product depends on the availability of in situ snow stations used in the CMC snow depth analysis where there are fewer snow stations (mostly located at airports) at higher latitudes. The model agrees well with CMC in areas such as the United States, southern Canada, and western Europe, where a dense network of WMO stations is used to condition the CMC data (see Reichle et al. 2011). The statistics are computed for subsets of the NH domain (see Table 4). The Canadian province of Quebec is known to have a strong north–south gradient in snow depth (Brown 2010) and the vegetation changes northward from mixed forest to boreal forest and taiga–tundra. The computed statistics (Table 4) show a rapid increase of positive bias (RMSE) relative to CMC from south to north. The southern part with dense broadleaf vegetation has the highest anomaly correlation of 0.68, followed by the boreal region with 0.52 and the taiga–tundra area with just 0.30. Because of the presence of a dense network of snow and precipitation measurement stations in the south, the agreement between the model and CMC is high despite the presence of dense vegetation in the south.

Areas with complex terrain were also analyzed further, including Alaska, the Rocky Mountains, western Siberia, and the TP. There is a low positive bias in Alaska, the Rocky Mountains, and western Siberia. The TP has a large negative bias and the anomaly correlation is close to zero. The low correlation in these areas can be

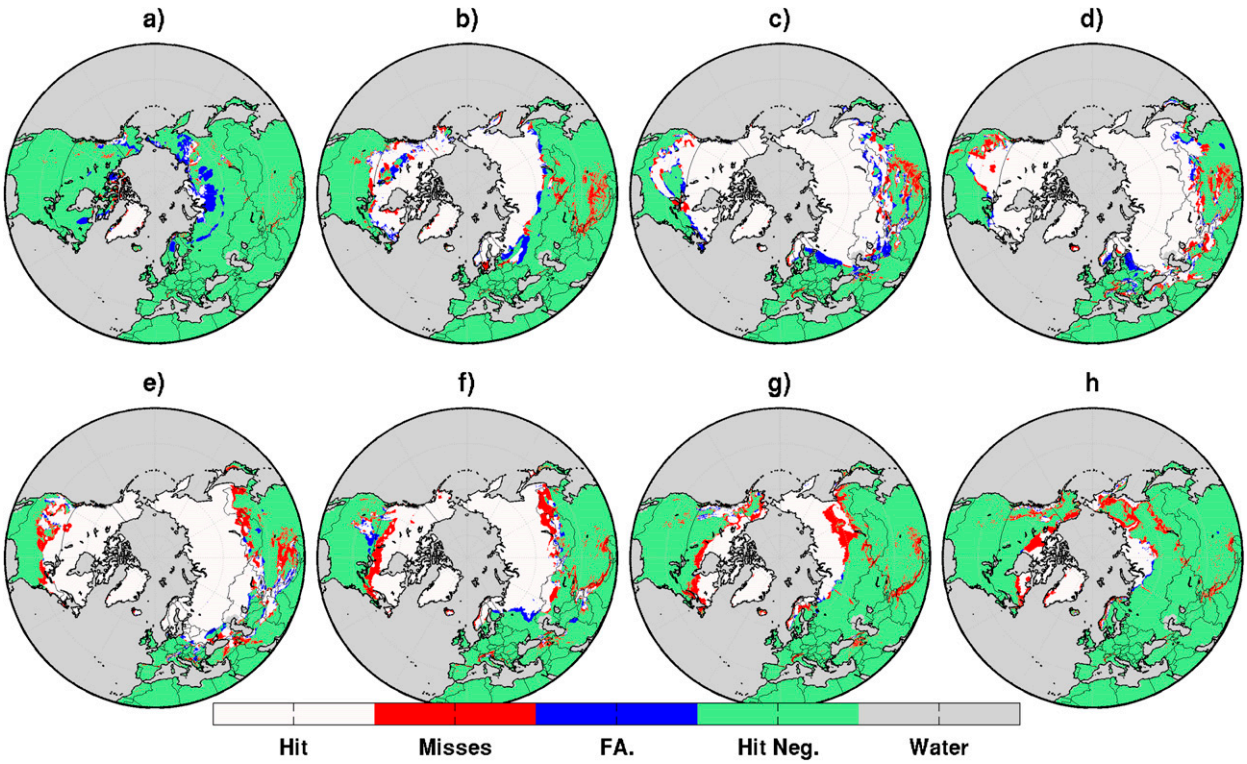

FIG. 5. Spatial distribution of hit, FA, hit negative, and misses of the model estimates compared to IMS snow cover for the snow season 2006/07. They represent the snow difference on the first day of the month of (a) October, (b) November, (c) December, (d) January, and (e)–(h) February–May.

attributed in part to the imperfect parameterization of the snow cover and the snowmelt model (Swenson and Lawrence 2012), but also to the atmospheric forcing being too coarse to resolve subgrid topographical effects. Further, CLM4 computes snow depth on the assumption that the snow is evenly spread across the grid cell, which often results in unrealistically thin snow depths where SCF is less than 1, when in reality snow may be distributed in patches.

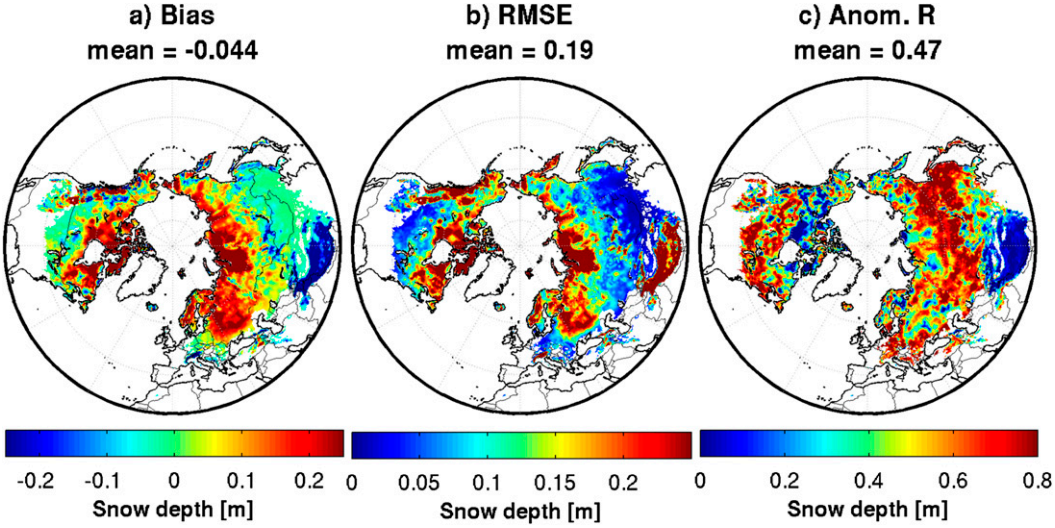

FIG. 6. Computed statistics of simulated snow depth relative to CMC snow product for the NH (2001–10): (a) bias, (b) RMSE, and (c) anomaly *R*.

TABLE 4. Summary of the statistics between CLM4 vs CMC and MERRA-Land vs CMC: bias, RMSE, and anomaly R are area-weighted average over the NH grid cells and selected areas for 2001–10.

	Bias (vs CMC) (m)				RMSE (vs CMC) (m)				Anomaly R (vs CMC)			
	Snow depth		SWE		Snow depth		SWE		Snow depth		SWE	
	CLM4	MERRA-Land	CLM4	MERRA-Land	CLM4	MERRA-Land	CLM4	MERRA-Land	CLM4	MERRA-Land	CLM4	MERRA-Land
NH	−0.044	0.001	−0.01	−0.015	0.19	0.18	0.046	0.038	0.47	0.38	0.43	0.53
Mixed forest, Quebec	0.092	−0.021	0.009	−0.03	0.14	0.14	0.141	0.043	0.68	0.58	0.58	0.6
Boreal, Quebec	0.11	−0.045	0.035	−0.013	0.203	0.17	0.057	0.047	0.52	0.55	0.46	0.54
Taiga–tundra, Quebec	0.2	0.09	0.054	0.027	0.26	0.19	0.073	0.057	0.3	0.41	0.27	0.39
Alaska	0.016	−0.005	0.019	−0.003	0.33	0.25	0.08	0.066	0.33	0.35	0.3	0.34
Rocky Mountains, United States	0.024	0.09	0.005	−0.006	0.104	0.15	0.029	0.023	0.49	0.2392	0.43	0.51
Western Siberia	0.114	0.055	0.027	0.011	0.16	0.13	0.04	0.03	0.6	0.6693	0.54	0.67
TP	−0.57	−0.43	−0.13	−0.131	0.57	0.54	0.134	0.132	0.08	0.0935	0.08	0.16
Farming prairies, Canada	0.017	0.092	0	−0.003	0.081	0.12	0.021	0.017	0.47	0.1885	0.42	0.49
Nunavut and Northwest Territories, Canada	0.16	0.095	0.034	0.018	0.19	0.15	0.044	0.0382	0.32	0.29	0.28	0.33
Mackenzie River basin	0.062	0.025	0.015	−0.003	0.21	0.19	0.057	0.0502	0.35	0.28	0.31	0.34

TABLE 5. Summary of the metrics of CLM4 vs SNOTEL SWE at 702 stations and CLM4 snow depth vs COOP at 1797 stations (from January 2001 to January 2011).

	Bias (m)	RMSE (m)	R
SNOTEL SWE	−0.22	0.23	0.45
COOP snow depth	−0.034	0.10	0.52

Areas with ephemeral snow, such as the northern part of the Canadian Prairies, also have low skill, as do Nunavut, the Northwest Territories, and the Mackenzie River basin (Fig. 6, Table 4). These areas have limited in situ observations available for correcting forcing data as well as for conditioning the CMC reanalysis product (Reichle et al. 2011). In these areas, CMC data are considered unreliable. Another source of uncertainty could be model underestimation of snow sublimation. At high latitudes, snow sublimation can be a significant component of snow ablation. Snow sublimation from on-ground and canopy-intercepted snow is a complex process that is not yet fully understood.

c. *Comparison against in situ data*

Because of the lack of reliable spatial and temporal large-scale SWE observations, point-scale SNOTEL SWE data were the only available data for the validation of CLM4 SWE. In this section, we compared CLM4 SWE and snow depth against SNOTEL SWE and COOP snow depth, respectively. A summary of the statistics is presented in Table 5.

1) CLM4 VERSUS SNOTEL SWE

The mean seasonal cycles of CLM4 SWE and those averaged over 702 SNOTEL stations were calculated for the 2001–10 study period. Figure 7 shows a large negative bias in simulated SWE. The mean SNOTEL SWE was approximately 6 times greater than that of CLM4. The mean bias and RMSE (Table 5) were −0.22 and 0.23 m, respectively. The large bias is not surprising because of the inherent variability of SWE within a grid cell of 0.5° resolution. The average elevation of the 702 stations used for the study was 2120 m. To evaluate the model's ability to simulate temporal variability of SWE, which is often more important than its absolute magnitude, the anomaly correlation was computed after subtracting the study period's mean seasonal cycle from each time series. The resulting anomaly correlation of 0.45 was close to the correlation between CLM4 and gridded CMC data (0.43). It is notable that while SNOTEL stations are located in the mountains, they are all placed on flat surfaces where deeper snow accumulation is likely, such that the spatial average may have a high bias.

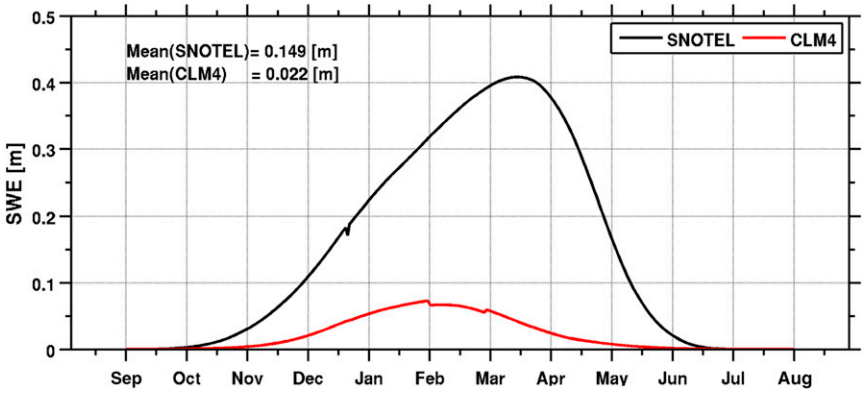

FIG. 7. Mean seasonal cycle of SWE (2001–10); black is the SNOTEL observations and red is the CLM4 estimates.

2) CLM4 VERSUS COOP SNOW DEPTH

COOP is an extensive climate-monitoring network of more than 11 000 volunteers and institutions across the United States, collecting and reporting daily hydrometeorological data including snow depth. After preprocessing and screening as described in section 2d, a total of 1797 COOP stations were considered in this analysis. Figure 8 depicts the evolution of CLM4 and COOP snow climatology. While the seasonal variability of CLM4 snow depth resembles that of COOP, the average climatology of CLM4 snow depth is ~36% lower than that of COOP observations. Similar bias and RMSE characteristics seen with SNOTEL data are also evident. CLM4 underestimates with a bias of −0.034 m and an RMSE of 0.10 m, with the larger bias occurring during the month of March. The scale issue of grid versus point is the only obvious explanation for the discrepancies between SNOTEL and CLM4. The bias is lower compared to the case with SNOTEL, presumably because COOP stations are mostly located at lower altitudes. The anomaly correlation of 0.52 is higher than in both cases with CMC and SNOTEL, but is significantly close to the anomaly correlation of CLM4 versus CMC when mountainous areas (the Rocky Mountains and the TP) are excluded from the analysis.

d. *Comparison against MERRA-Land*

We used MERRA-Land meteorological fields to force CLM4. This provides an opportunity for an intercomparison of CLM4 with the Catchment LSM. MERRA-Land snow depth and SWE were compared to CMC; the summary of the statistics is presented in Table 4. The NH area-weighted-average bias of MERRA-Land snow depth is lower at −0.001 compared to −0.044 m of CLM4, while RMSEs were 0.18 and 0.19 m, respectively. The correlation coefficient for MERRA-Land snow depth is lower at 0.38 compared to 0.47 for CLM4. The Fisher's r-to-z transformation test of the two correlation coefficients yielded a z value of −4.83 (with probability $P = 0.00$). Since the P value is less than 0.05, we concluded that the two correlation coefficients are significantly different from each other. One explanation of the lower skill of MERRA-Land compared to CLM4 could be that in the MERRA LSM, when snowfall occurs on bare soil, snow is initially spread out over a fraction of the catchment so that at any given point the SWE is equal to a constant value (WEMIN = 26 mm). As a consequence, the model tends to pile up the first snowfall vertically before spreading the snow horizontally in the catchment. Further, before the maximum SCF is reached, the model output snow depth is the depth of the snow within the snow-covered gridcell area.

The difference between CLM4 and MERRA-Land RMSE and square of correlation coefficient R^2 against

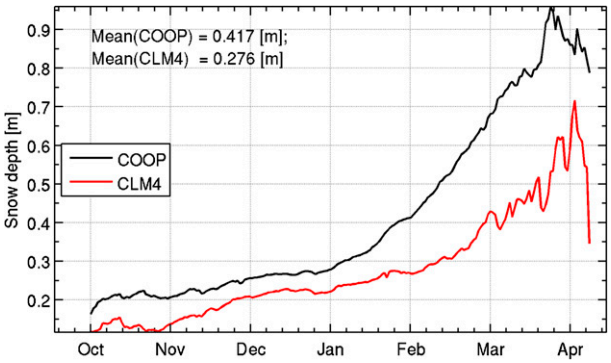

FIG. 8. CLM4 and COOP station-average snow depth climatology (average of 1797 stations). The black line represents the COOP stations signal, and the red line represents the CLM4 simulation.

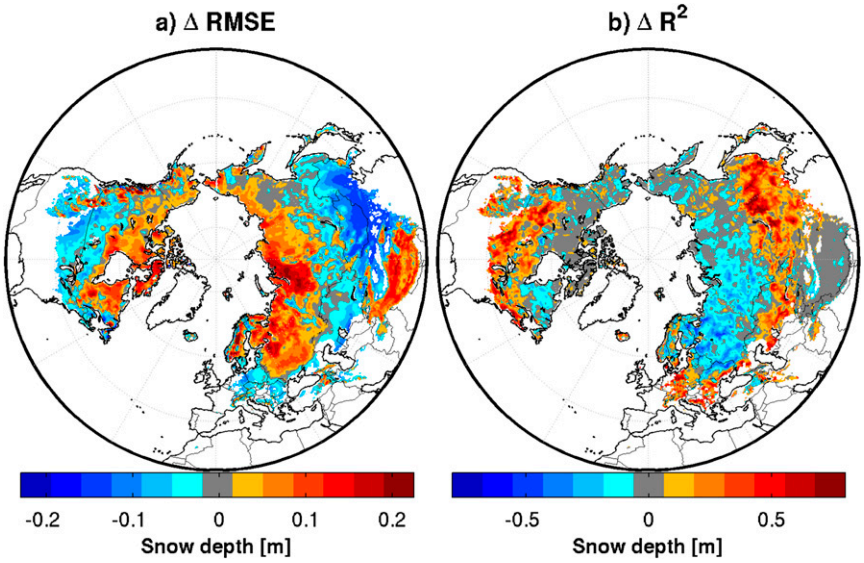

FIG. 9. Difference of (a) RMSE and (b) skill (R^2) of CLM4 and MERRA-Land compared to CMC snow depth.

CMC are presented in Fig. 9. CLM4 has higher biases in Northern Canada, Quebec, northern Eurasia, and the TP compared to MERRA-Land. Lower biases are found in southern Eurasia, the United States, and the Canadian Prairies (Fig. 9a). The models show similar spatial distribution of R^2 in areas over western Siberia, the TP, and northwestern Canada (Fig. 9b). In high altitudes, including the TP, there is a limited number of snow depth observations used to condition the CMC product. These areas correspond to areas where limited numbers of radiosonde observations are available for correction of MERRA forcing data [see Reichle et al. (2011), their Fig. 9c]; therefore, both MERRA-Land and CLM4 appear to perform worse in areas where observations of snow depth are scarce (Table 4). SWE also exhibits a similar spatial pattern (not shown), but MERRA-Land R^2 is significantly higher (0.53) than that of CLM4 (0.43). When MERRA forcing without CPCU correction is used to run CLM4, R^2 of CLM4 snow depth increases from 0.53 to 0.56 and that of SWE increases from 0.43 to 0.49. A possible explanation for this increasing skill could be that CPCU overcorrects the precipitation field, especially at high altitudes where fewer observation stations are available to condition the forcing data.

Figure 10 represents the annual cycle (2001/02 season) of fraction of land covered by snow (LCF) based on CLM4, and snow density of CLM4, MERRA, and CMC. The CMC snow density is estimated using an empirical model developed by Sturm et al. (2010), based on data from the United States, Canada, and Switzerland. The bulk of snow density is calculated as a function of snow depth h_i and the day of year (DOY) as follows:

$$\rho_{h_i,\mathrm{DOY}_i} = (\rho_{\max} - \rho_0)[1 - \exp(-k_1 h_1 - k_2 \mathrm{DOY}_i)] + \rho_0, \tag{7}$$

where $0 < \rho_0 < \rho_{\max} < 1\,\mathrm{g\,cm^{-3}}$. The variables k_1 and k_2 are the densification parameters for the depth and DOY, $\rho_{\max}$ is the maximum bulk density, and ρ_0 is the initial density of the individual snow layer. Values of these parameters as a function of the snow classes are found in Sturm et al. (2010).

We calculated CLM4 and MERRA-Land snow density for each snow class as defined in Sturm et al. (1995). Figures 10a–f represent the densities of tundra, taiga, maritime, ephemeral, prairie, and alpine snow climate classes compared to CMC snow density. The time series show that the CMC agrees reasonably well with CLM4 and MERRA-Land for all snow climate classes with the exception of ephemeral (Fig. 10d) and prairie (Fig. 10e). Tundra, taiga, maritime, and alpine snow climate classes all have high winter LCF (>70). Ephemeral and prairie classes are characterized by shallow snow and low winter LCF. Figure 11 shows the correlation coefficients between CLM4 or MERRA-Land and CMC. There is a strong correlation between simulated and observed density over all classes except for ephemeral and prairie. The poor performance of both CLM4 and MERRA in the Great Plains and in the southern part of the United States where snow is predominantly classified as ephemeral is due to the snow densities used in the models. In

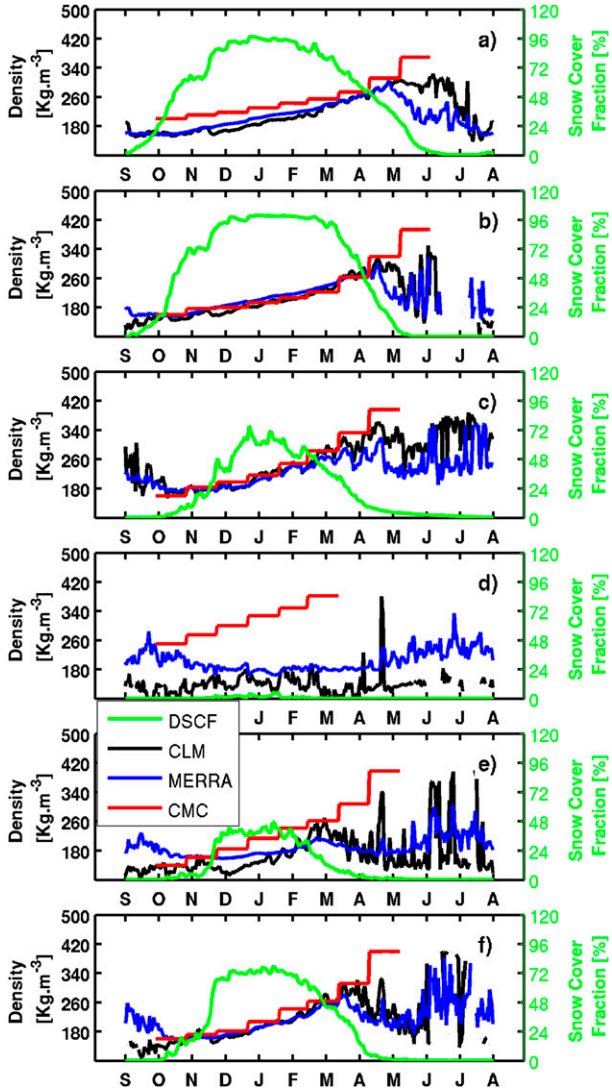

FIG. 10. Time series of LCF based on CLM4, and snow density of CLM4, MERRA-Land, and CMC for each snow class as defined in Sturm et al. (1995): (a)–(f) the densities of tundra, taiga, maritime, ephemeral, prairie, and alpine snow climate classes, respectively.

fact, Sturm et al. (2010) explicitly cautions against the use of simple parameterizations to determine density in areas with ephemeral and shallow snow because of the lack of historic climatological and snow density data and the complex nature of snow metamorphism in these areas. The CMC snow density in these areas also may be unreliable. Discrepancies in snow density between CLM4 and MERRA-Land could in part be explained by the difference in the value of new snow density, which is set to 100 and 150 $kg\,m^{-3}$ for CLM4 and MERRA-Land, respectively.

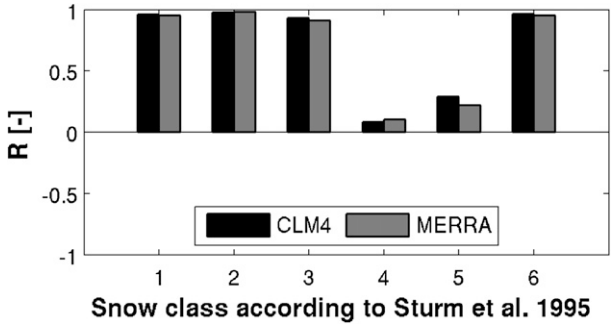

FIG. 11. Correlation coefficient of CLM4 and MERRA-Land snow density against CMC for each snow climate class as defined in Sturm et al. (1995): 1, tundra; 2, taiga; 3, maritime; 4, ephemeral; 5, prairie; and 6, alpine.

4. Discussion and conclusions

We ran CLM4 in offline mode with the MERRA-Land atmospheric forcing data. The snow model outputs were compared with a set of observation data including MODIS SCF, IMS snow-cover data, SNOTEL SWE, and COOP snow depth. The performance of CLM4 against CMC data was also compared to that of the MERRA-Land snow model.

CLM4 snow outputs agree reasonably well with MODIS SCF, IMS, and CMC observations but not with point-scale in situ SNOTEL and COOP data. The largest departure from the validation data appears in areas with complex topography.

Comparisons with MODIS and IMS show that CLM4 underestimates SCF in the U.S. Prairies and southern Russia in part because of the fast model snow ablation resulting in the snow line moving northward faster than the observation. Many processes affect the snow spatial distribution at the subgrid level, including snow interception by canopy, heterogeneous snow redistribution by wind blowing (Liston 2004), and spatially heterogeneous processes such as snowfall distribution and melting rate.

Equation (1) used to compute CLM4 SCF was derived from AVHRR observations and the CMC snow depth analysis, both at $1° \times 1°$ spatial resolution and at monthly temporal resolution.

This is problematic in that 1) the spatial resolution of the data used is too coarse to resolve subgrid SCF heterogeneity, 2) it has been demonstrated that CMC data tend to melt earlier compared to satellite observation such as MODIS, 3) the time period of monthly averaging is generally shorter than the accumulation period (which can be a few months) and longer than the snowmelt period (which may last only a few weeks), and 4) the model does not reflect the hysteresis in SCF for a given snow depth between accumulation and melt phases.

Snowmelt is controlled by meteorological conditions such as radiation and wind and topographical features of the terrain such as the slope and aspect. In CLM4, emitted longwave radiation, sensible heat, latent heat, and ground heat flux are all calculated based on the assumption of uniform snow cover, whereas in reality snow cover is more patchy. This could contribute to simulated snow melting too fast in the spring, as the model does not represent the deeper accumulations, which are often shaded and protected by local topography, allowing them to persist much longer into the warm season. In the fall, the uniformity of simulated snow cover increases the albedo, reducing net radiation and thus surface temperature, allowing snow to accumulate faster than it does in the real world. These issues were investigated by Swenson and Lawrence (2012), and they proposed two new parameterizations of SCF, one for the accumulation phase of the snow and the other for the depletion period. Testing the new parameterizations against our dataset would be of interest, but it is beyond the scope of this paper.

Even though the CLM4 binary snow cover compared very well with IMS data, the map of snow cover reveals misses in the southern United States and southern Russia, confirming the findings from comparisons of the CLM with MODIS SCF that the model has a bias toward early snowmelt. It is impossible to evaluate exactly how well CLM4 performs against IMS as IMS was found to systematically overestimate SCF when compared against MODIS observations.

SNOTEL represents the best available in situ SWE data in the United States; however, it is not representative of the snow distribution in the average CLM4 grid cell within which the SNOTEL station is located because of the complex nature of the terrain, which leads to high variability in snow deposition. One may argue that these data are not suitable for gridded model validation. We sought to investigate the temporal variations, as they are the most important in model-based application. To evaluate the temporal variations, we computed time series of anomalies for the model and SNOTEL by subtracting their climatological mean seasonal cycles before calculating the correlation. The anomaly correlation (0.45) is relatively skillful, given that of CLM4 and CMC is 0.43, and these relationships can be applied for bias correction of the SNOTEL data for a wide range of subsequent applications.

Both CLM4 and MERRA-Land show good agreement with CMC in the midlatitudes and weak skills in high latitudes such as Siberia, the TP, Alaska, and the western coast of North America (Fig. 9b). In a multimodel intercomparison, Brown and Mote (2009) found similar results in those areas when CMC data were compared to model outputs. ERA-Interim, a reanalysis product of the European Centre for Medium-Range Weather Forecasts (Simmons et al. 2007), also showed a similar distribution of error compared to CMC, and Reichle et al. (2011) reported that the MERRA-Land and ERA-Interim models perform the best in areas where in situ data are used to condition the forcing data. The maximum snow depth in CMC was set to 120 cm. This may explain the disagreement with CLM4 in areas where the snow depth is more likely to surpass the threshold.

CLM4 snow depth is computed on the assumption that precipitation events uniformly distribute snowfall across the grid cell. For SCF less than 1, this results in a thinner snowpack than would be calculated if the snow was spread in patches. This parameterization of snow depth modifies the heat flux through the snowpack and therefore influences the metamorphism of the snow, which is responsible for the compaction of the snow and evolution of the snow density.

Because of the scarcity of snow depth observations at high latitudes and in China, CMC snow depth analysis is generated entirely by the background field, but it is considered to be unreliable (Brown et al. 2010). In areas of shallow snow, CMC analysis has a tendency to underestimate snow depth because most of the snow depth observations used are from stations located in clearings at low altitudes. This also causes CMC data to have an early snowmelt in the spring (Brown et al. 2010); therefore, it is impossible to tell which of the models and CMC are most accurate in those areas. Errors in the modeled snow depth result in part from imperfect precipitation forcing. The density of precipitation gauges in high latitudes and in some mountainous regions is not high enough to reproduce the spatial variability of the precipitation (Adler et al. 2003). Another problem is precipitation undercatch due to wind-induced turbulence at the gauge orifice, wetting losses on the internal walls of the gauge, and evaporation losses in precipitation measurements (Groisman and Legates 1994). Bias corrections are usually applied to account for the undercatch, but these are imperfect and can lead to errors in the precipitation forcing (e.g., Swenson 2010). Uncertainty in the snow density parameterizations is another possible source of error contributing to the model's disagreement with CMC.

Because of the limited number of in situ snow density observations, the Sturm et al. (2010) density parameterization was used to convert CMC snow depth to SWE. The parameterization is based on the Sturm et al. (1995) seasonal snow classification (alpine, maritime, prairie, tundra, and taiga), snow depth, and the day of the year. It computes snow bulk density using a linear

density–time relation to approximate the snow densification rate. The parameterization is empirical and neglects interannual variability in densification rates. Furthermore, because the parameterization was not designed for a daily time step, it does not capture the changes in the microphysics of snow that lead to local variability in snow density. Thus, it may not be suitable to convert time series of snow depth into SWE, especially in areas with shallow and ephemeral snow.

Finally, CLM4 mass conservation is computed using the following equation:

$$\text{imbalance} = (P - \text{ET} - Q)\Delta t - (\Delta\text{can} + \Delta\text{SWE} + \Delta\text{SM} + \Delta\text{WA}), \quad (8)$$

where P is precipitation (rainfall plus snowfall), ET is evapotranspiration, and Q is runoff (all in mm day^{-1}). ET is a sum of vegetation transpiration, canopy evaporation, and soil evaporation and snow sublimation, and Q is composed of surface and subsurface runoff over land and surface runoff at glaciers, wetlands, and lake. The latter terms in Eq. (8) refer to the storage changes, where Δ denotes changes in canopy-intercepted water (can), SWE, soil moisture (SM), and aquifer water storage (WA). The modeled water budget is balanced, that is, there is no significant residual whereas there is likely to be a residual in an assimilation model water budget. We will reserve this discussion in the follow-on paper on data assimilation.

These findings lay a foundation for a follow-up study on using data assimilation to improve the simulation of snow. MODIS data only indicate the presence or absence of snow and cannot provide information on snow mass. The data can be used to improve the model prediction in areas with shallow snow (<5 cm). Model prediction in areas with moderate snow depth (<1 m) can be improved through a radiance assimilation scheme using the brightness temperature measured by AMSR-E on board the *Aqua* satellite. GRACE data can be used to improve SWE prediction in areas with deep snow, mountainous terrain, and dense vegetation cover.

Acknowledgments. This study is funded by NASA's Science of *Terra* and *Aqua* program. Resources supporting this work were provided by the NASA High-End Computing (HEC) Program through the NASA Center for Climate Simulation (NCCS) at Goddard Space Flight Center. Thanks go to K. R. Arsenault for preprocessing SNOTEL and COOP data. Thanks also go to Sam Lewis and Keith Oleson at NCAR for their help in setting up and running CESM on NASA Center for Climate Simulation (NCCS) computers.

REFERENCES

Adler, R. F., and Coauthors, 2003: The version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present). *J. Hydrometeor.*, **4**, 1147–1167, doi:10.1175/1525-7541(2003)004<1147:TVGPCP>2.0.CO;2.

Anderson, E. A., 1976: A point energy and mass balance model of a snow cover. NOAA Tech. Rep. NWS 19, 150 pp. [Available online at http://amazon.nws.noaa.gov/articles/HRL_Pubs_PDF_May12_2009/HRL_PUBS_51-100/81_A_POINT_ENERGY_AND_MASS.pdf.]

Arsenault, K. R., P. R. Houser, and G. J. M. De Lannoy, 2014: Evaluation of the MODIS snow cover fraction product. *Hydrol. Processes*, **28**, 980–998, doi:10.1002/hyp.9636.

Bamzai, A. S., and J. Shukla, 1999: Relation between Eurasian snow cover, snow depth, and the Indian summer monsoon: An observational study. *J. Climate*, **12**, 3117–3132, doi:10.1175/1520-0442(1999)012<3117:RBESCS>2.0.CO;2.

Barrett, A., 2003: National Operational Hydrologic Remote Sensing Center Snow Data Assimilation System (SNODAS) products at NSIDC. NSIDC Special Rep. 11, 19 pp. [Available online at https://nsidc.org/pubs/documents/special/nsidc_special_report_11.pdf.]

Beven, K. J., and M. J. Kirkby, 1979: A physically based variable contributing area model of basin hydrology. *Hydrol. Sci. Bull.*, **24**, 43–69, doi:10.1080/02626667909491834.

Bitner, D., T. Carroll, D. Cline, and P. Romanov, 2002: An assessment of the differences between three satellite snow cover mapping techniques. *Hydrol. Processes*, **16**, 3723–3733, doi:10.1002/hyp.1231.

Brasnett, B., 1999: A global analysis of snow depth for numerical weather prediction. *J. Appl. Meteor.*, **38**, 726–740, doi:10.1175/1520-0450(1999)038<0726:AGAOSD>2.0.CO;2.

Brown, R. D., 2010: Analysis of snow cover variability and change in Québec, 1948–2005. *Hydrol. Processes*, **24**, 1929–1954, doi:10.1002/hyp.7565.

——, and P. Mote, 2009: The response of Northern Hemisphere snow cover to a changing climate. *J. Climate*, **22**, 2124–2145, doi:10.1175/2008JCLI2665.1.

——, and B. Brasnett, 2010: Canadian Meteorological Centre (CMC) daily snow depth analysis data. Environment Canada, National Snow and Ice Data Center, Boulder, CO, digital media, updated annually, accessed 1 October 2012. [Available online at https://nsidc.org/data/docs/daac/nsidc0447_CMC_snow_depth/.]

——, ——, and D. Robinson, 2003: Gridded North American monthly snow depth and snow water equivalent for GCM evaluation. *Atmos.–Ocean*, **41**, 1–14, doi:10.3137/ao.410101.

——, C. Derksen, and L. Wang, 2010: A multi-data set analysis of variability and change in Arctic spring snow cover extent, 1967–2008. *J. Geophys. Res.*, **115**, D16111, doi:10.1029/2010JD013975.

Brubaker, K. L., R. T. Pinker, and E. Deviatova, 2005: Evaluation and comparison of MODIS and IMS snow-cover estimates for the continental United States using station data. *J. Hydrometeor.*, **6**, 1002–1017, doi:10.1175/JHM447.1.

Brun, E., 1989: Investigation of wet-snow metamorphism in respect of liquid water content. *Ann. Glaciol.*, **13**, 22–26.

Chapin, F. S., and Coauthors, 2005: Role of land-surface changes in Arctic summer warming. *Science*, **310**, 657–660, doi:10.1126/science.1117368.

Clifford, D., 2010: Global estimates of snow water equivalent from passive microwave instruments: History, challenges and future developments. *Int. J. Remote Sens.*, **31**, 3707–3726, doi:10.1080/01431161.2010.483482.

Clow, D. W., L. Nanus, K. L. Verdin, and J. Schmidt, 2012: Evaluation of SNODAS snow depth and snow water equivalent estimates for the Colorado Rocky Mountains, USA. *Hydrol. Processes*, **26**, 2583–2591, doi:10.1002/hyp.9385.

Cohen, J., 1994: Snow cover and climate. *Weather*, **49**, 150–156, doi:10.1002/j.1477-8696.1994.tb05997.x.

——, and D. Rind, 1991: The effect of snow cover on climate. *J. Climate*, **4**, 689–706, doi:10.1175/1520-0442(1991)004<0689:TEOSCO>2.0.CO;2.

——, and D. Entekhabi, 1999: Eurasian snow cover variability and Northern Hemisphere climate predictability. *Geophys. Res. Lett.*, **26**, 345–348, doi:10.1029/1998GL900321.

Dai, Y., and Q. Zeng, 1997: A land surface model (IAP94) for climate studies. Part I: Formulation and validation in off-line experiments. *Adv. Atmos. Sci.*, **14**, 433–460, doi:10.1007/s00376-997-0063-4.

——, and Coauthors, 2003: The Common Land Model. *Bull. Amer. Meteor. Soc.*, **84**, 1013–1023, doi:10.1175/BAMS-84-8-1013.

Dewey, K. F., 1977: Daily maximum and minimum temperature forecasts and the influence of snow cover. *Mon. Wea. Rev.*, **105**, 1594–1597, doi:10.1175/1520-0493(1977)105<1594:DMAMTF>2.0.CO;2.

Dingman, S. L., 2002: *Physical Hydrology*. 2nd ed., Prentice Hall, 646 pp.

Ducharne, A., R. D. Koster, M. J. Suarez, M. Stieglitz, and P. Kumar, 2000: A catchment-based approach to modeling land surface processes in a general circulation model: 2. Parameter estimation and model demonstration. *J. Geophys. Res.*, **105**, 24 823–24 838, doi:10.1029/2000JD900328.

Durre, I., M. J. Menne, B. E. Gleason, T. G. Houston, and R. S. Vose, 2010: Comprehensive automated quality assurance of daily surface observations. *J. Appl. Meteor. Climatol.*, **49**, 1615–1633, doi:10.1175/2010JAMC2375.1.

Flanner, M. G., and C. S. Zender, 2005: Snowpack radiative heating: Influence on Tibetan Plateau climate. *Geophys. Res. Lett.*, **32**, L06501, doi:10.1029/2004GL022076.

——, and ——, 2006: Linking snowpack microphysics and albedo evolution. *J. Geophys. Res.*, **111**, D12208, doi:10.1029/2005JD006834.

——, ——, J. T. Randerson, and P. J. Rasch, 2007: Present day climate forcing and response from black carbon in snow. *J. Geophys. Res.*, **112**, D11202, doi:10.1029/2006JD008003.

Forman, B. A., R. H. Reichle, and M. Rodell, 2012: Assimilation of terrestrial water storage from GRACE in a snow-dominated basin. *Water Resour. Res.*, **48**, W01507, doi:10.1029/2011WR011239.

Frei, A., and S. Lee, 2010: A comparison of optical-band based snow extent products during spring over North America. *Remote Sens. Environ.*, **114**, 1940–1948, doi:10.1016/j.rse.2010.03.015.

Gent, P. R., and Coauthors, 2011: The Community Climate System Model version 4. *J. Climate*, **24**, 4973–4991, doi:10.1175/2011JCLI4083.1.

Groisman, P. Y., and D. R. Legates, 1994: The accuracy of United States precipitation data. *Bull. Amer. Meteor. Soc.*, **75**, 215–227, doi:10.1175/1520-0477(1994)075<0215:TAOUSP>2.0.CO;2.

Hall, D. K., and G. A. Riggs, 2007: Accuracy assessment of the MODIS snow-cover products. *Hydrol. Processes*, **21**, 1534–1547, doi:10.1002/hyp.6715.

——, ——, V. V. Salomonson, N. E. DiGirolamo, and K. J. Bayr, 2002: MODIS snow-cover products. *Remote Sens. Environ.*, **83**, 181–194, doi:10.1016/S0034-4257(02)00095-0.

——, V. V. Salomonson, and G. A. Riggs, 2006a: MODIS/*Aqua* snow cover daily L3 global 0.05deg CMG, version 5. National Snow and Ice Data Center, Boulder, CO, digital media, accessed 1 December 2011. [Available online at http://nsidc.org/data/myd10c1.]

——, ——, and ——, 2006b: MODIS/Terra snow cover daily L3 global 0.05deg CMG, version 5. National Snow and Ice Data Center, Boulder, CO, digital media, accessed 1 December 2011.

Helfrich, S. R., D. McNamara, B. H. Ramsay, T. Baldwin, and T. Kasheta, 2007: Enhancements to, and forthcoming developments in the Interactive Multisensor Snow and Ice Mapping System (IMS). *Hydrol. Processes*, **21**, 1576–1586, doi:10.1002/hyp.6720.

Jordan, R., 1991: A one-dimensional temperature model for a snow cover: Technical documentation for SNTERERM.89. Special Rep. 91-16, Cold Region Research and Engineers Laboratory, U.S. Army Corps of Engineers, Hanover, NH, 61 pp.

Karl, T. R., P. Y. Groisman, R. W. Knight, and R. R. Heim, 1993: Recent variations of snow cover and snowfall in North America and their relation to precipitation and temperature variations. *J. Climate*, **6**, 1327–1344, doi:10.1175/1520-0442(1993)006<1327:RVOSCA>2.0.CO;2.

Klein, A. G., and A. C. Barnett, 2003: Validation of daily MODIS snow cover maps of the Upper Rio Grande river basin for the 2000–2001 snow year. *Remote Sens. Environ.*, **86**, 162–176, doi:10.1016/S0034-4257(03)00097-X.

Kluzek, E., 2012: CESM Research Tools: CLM4 in CESM1.0.4 User's Guide Documentation. Tech. Doc., 164 pp. [Available online at http://www.cesm.ucar.edu/models/cesm1.0/clm/index.shtml.]

Koster, R. D., M. J. Suarez, A. Ducharne, M. Stieglitz, and P. Kumar, 2000: A catchment-based approach to modeling land surface processes in a general circulation model: 1. Model structure. *J. Geophys. Res.*, **105**, 24 809–24 822, doi:10.1029/2000JD900327.

Kukla, G. J., 1981: Snow covers and climate. Glaciological data: Snow watch 1980, Rep. GD-11, G. Kukla, A. Hecht, and D. Wiesnet, Eds., World Data Center A for Glaciology, Boulder, CO, 27–39. [Available online at http://nsidc.org/pubs/documents/gd/GD-11_web.pdf]

Lawrence, D. M., and A. Slater, 2010: The contribution of snow condition trends to future ground, climate. *Climate Dyn.*, **34**, 969–981, doi:10.1007/s00382-009-0537-4.

——, and Coauthors, 2011: Parameterization improvements and functional and structural advances in version 4 of the Community Land Model. *J. Adv. Model. Earth Syst.*, 3, M03001, doi:10.1029/2011MS000045.

Li, H., M. Huang, M. S. Wigmosta, Y. Ke, A. M. Coleman, L. R. Leung, A. Wang, and D. M. Ricciuto, 2011: Evaluating runoff simulations from the Community Land Model 4.0 using observations from flux towers and a mountainous watershed. *J. Geophys. Res.*, **116**, D24120, doi:10.1029/2011JD016276.

Liston, G. E., 2004: Representing subgrid snow cover heterogeneities in regional and global models. *J. Climate*, **17**, 1381–1397, doi:10.1175/1520-0442(2004)017<1381:RSSCHI>2.0.CO;2.

Maurer, E. P., J. D. Rhoads, R. O. Dubayah, and D. P. Lettenmaier, 2003: Evaluation of the snow-covered area data product from MODIS. *Hydrol. Processes*, **17**, 59–71, doi:10.1002/hyp.1193.

Moser, C. L., G. A. Tootle, A. A. Oubeidillah, and V. Lakshmi, 2011: A comparison of SNOTEL and AMSR-E snow water equivalent data sets in western US watersheds. *Int. J. Remote Sens.*, **32**, 6611–6629, doi:10.1080/01431161.2010.512936.

Niu, G.-Y., and Z.-L. Yang, 2006: Effects of frozen soil on snowmelt runoff and soil water storage at a continental scale. *J. Hydrometeor.*, **7**, 937–952, doi:10.1175/JHM538.1.

——, and ——, 2007: An observation-based formulation of snow cover fraction and its evaluation over large North American river basins. *J. Geophys. Res.*, **112**, D21101, doi:10.1029/2007JD008674.

——, ——, R. E. Dickinson, and L. E. Gulden, 2005: A simple TOPMODEL-based runoff parameterization (SIMTOP) for use in global climate models. *J. Geophys. Res.*, **110**, D21106, doi:10.1029/2005JD006111.

Oleson, K., and Coauthors, 2008: Improvements to the Community Land Model and their impact on the hydrological cycle. *J. Geophys. Res.*, **113**, G01021, doi:10.1029/2007JG000563.

Pan, M., and Coauthors, 2003: Snow process modeling in the North American Land Data Assimilation System (NLDAS): 2. Evaluation of model simulated snow water equivalent. *J. Geophys. Res.*, **108**, 8850, doi:10.1029/2003JD003994.

Peterson, D. H., R. E. Smith, M. D. Dettinger, D. R. Cayan, and L. Riddle, 2000: An organized signal in snowmelt runoff over the western United States. *J. Amer. Water Resour. Assoc.*, **36**, 421–432, doi:10.1111/j.1752-1688.2000.tb04278.x.

Ramsay, B., 1998: The interactive multisensor snow and ice mapping system. *Hydrol. Processes*, **12**, 1537–1546, doi:10.1002/(SICI)1099-1085(199808/09)12:10/11<1537::AID-HYP679>3.0.CO;2-A.

Rango, A., E. G. Landesa, M. Bleiweiss, K. Havstad, and K. Tanksley, 2003: Improved satellite snow mapping, snowmelt runoff forecasting, and climate change simulations in the Upper Rio Grande basin. *World Resour. Rev.*, **15**, 26–41.

Reichle, R. H., 2012: The MERRA-Land Data Product (version 1.2). GMAO Office Note 3, 38 pp. [Available online at http://gmao.gsfc.nasa.gov/pubs/docs/Reichle541.pdf.]

——, R. D. Koster, G. J. M. De Lannoy, B. A. Forman, Q. Liu, S. P. P. Mahanama, and A. M. Toure, 2011: Assessment and enhancement of MERRA land surface hydrology estimates. *J. Climate*, **24**, 6322–6338, doi:10.1175/JCLI-D-10-05033.1.

Rienecker, M., and Coauthors, 2008: The GEOS-5 Data Assimilation System Documentation of Versions 5.0.1 and 5.1.0, and 5.2.0. NASA Tech. Rep. NASA/TM-2008-104606, Vol. 27, 92 pp. [Available online at http://gmao.gsfc.nasa.gov/pubs/docs/tm27.pdf.]

——, and Coauthors, 2011: MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. *J. Climate*, **24**, 3624–3648, doi:10.1175/JCLI-D-11-00015.1.

Riggs, G. A., and D. K. Hall, 2012: Improved snow mapping accuracy with revised MODIS snow algorithm. *Proc. 69th Annual Eastern Snow Conf.*, Claryville, NY, Eastern Snow Conference, 209–214. [Available online at http://www.easternsnow.org/proceedings/2012/riggs_and_hall.pdf.]

Serreze, M. C., M. P. Clark, R. L. Armstrong, D. A. McGuiness, and R. S. Pulwarty, 1999: Characteristics of the western United States snowpack from Snowpack Telemetry (SNOTEL) data. *Water Resour. Res.*, **35**, 2145–2160, doi:10.1029/1999WR900090.

Simmons, A. J., S. M. Uppala, D. Dee, and S. Kobayashi, 2007: ERA-Interim: New ECMWF 759 reanalysis product from 1989 onwards. *ECMWF Newsletter*, No. 110, ECMWF, Reading, United Kingdom, 25–35.

Simpson, J. J., J. R. Stitt, and M. Sienko, 1998: Improved estimates of the areal extent of snow cover from AVHRR data. *J. Hydrol.*, **204**, 1–23, doi:10.1016/S0022-1694(97)00087-5.

Stieglitz, M., A. Ducharne, R. D. Koster, and M. J. Suarez, 2001: The impact of detailed snow physics on the simulation of snow cover and subsurface thermodynamics at continental scales. *J. Hydrometeor.*, **2**, 228–242, doi:10.1175/1525-7541(2001)002<0228:TIODSP>2.0.CO;2.

Sturm, M., J. Holmgren, and G. E. Liston, 1995: A seasonal snow cover classification system for local to global applications. *J. Climate*, **8**, 1261–1283, doi:10.1175/1520-0442(1995)008<1261:ASSCCS>2.0.CO;2.

——, B. Taras, G. E. Liston, C. Derksen, T. Jonas, and J. Lea, 2010: Estimating snow water equivalent using snow depth data and climate classes. *J. Hydrometeor.*, **11**, 1380–1394, doi:10.1175/2010JHM1202.1.

Su, H., Z.-L. Yang, R. E. Dickinson, C. R. Wilson, and G.-Y. Niu, 2010: Multisensor snow data assimilation at the continental scale: The value of Gravity Recovery and Climate Experiment terrestrial water storage information. *J. Geophys. Res.*, **115**, D10104, doi:10.1029/2009JD013035.

Swenson, S. C., 2010: Assessing high-latitude winter precipitation from global precipitation analyses using GRACE. *J. Hydrometeor.*, **11**, 405–420, doi:10.1175/2009JHM1194.1.

——, and M. D. Lawrence, 2012: A new fractional snow-covered area parameterization for the Community Land Model and its effect on the surface energy balance. *J. Geophys. Res.*, **117**, D21107, doi:10.1029/2012JD018178.

Thackeray, C. W., C. G. Fletcher, and C. Derksen, 2014: The influence of canopy snow parameterizations on snow albedo feedback in boreal forest regions. *J. Geophys. Res. Atmos.*, **119**, 9810–9821, doi:10.1002/2014JD021858.

Turner, A. G., and J. M. Slingo, 2011: Using idealized snow forcing to test teleconnections with the Indian summer monsoon in the Hadley Centre GCM. *Climate Dyn.*, **36**, 1717–1735, doi:10.1007/s00382-010-0805-3.

Vertenstein, M., T. Craig, A. Middleton, D. Feddema, and C. Fisher, 2012: CESM1.0.4 user's guide. Tech. Doc., 146 pp. [Available online at http://www.cesm.ucar.edu/models/cesm1.0/cesm/cesm_doc_1_0_4/ug.pdf.]

Warren, S. G., and W. J. Wiscombe, 1980: A model for the spectral albedo of snow. II: Snow containing atmospheric aerosols. *J. Atmos. Sci.*, **37**, 2734–2745, doi:10.1175/1520-0469(1980)037<2734:AMFTSA>2.0.CO;2.

Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences.* 2nd ed. International Geophysics Series, Vol. 91, Academic Press, 627 pp.