Evaluation of Simulated Snow and Snowmelt Timing in the Community Land Model Using Satellite-based Products and Streamflow Observations

Ally M. Toure¹, Kari Luojus², Matthew Rodell³, Hiroko Beaudoin³, ⁴, Augusto Getirana ³, ⁴

¹Wifrid Laurier University, Department of Geography and Environmental Sciences, Waterloo, Ontario, Canada.

²Finnish Meteorological Institute, Erik Palmenin Aukio 1, FI-00560, Helsinki, Finland

³NASA Goddard Space Flight Center, Hydrological Sciences Laboratory, Code 617, 8800 Greenbelt Road, Greenbelt, MD 20771, USA.

⁴Earth System Science Interdisciplinary Center (ESSIC), University of Maryland, College Park, MD, USA.

Corresponding author: Ally M. TOURE
Email: allytoure@gmail.com
Abstract

The purpose of this study was to evaluate snow and snowmelt simulated by version 4 of the Community Land Model (CLM4). We performed uncoupled CLM4 simulations, forced by Modern-Era Retrospective Analysis for Research and Applications Land-only (MERRA-Land) meteorological fields. GlobSnow snow cover fraction (SCF), snow water equivalent (SWE) and satellite-based passive microwave (PMW) snowmelt-off day of year (MoD) data were used to evaluate SCF, SWE, and snowmelt simulations. Simulated runoff was then fed into a river routing scheme and evaluation was performed at 408 snow-dominated catchments using gauge observations. CLM4 and GlobSnow snow cover extent showed a strong agreement, especially during the peak snow cover months. Overall there was a good correlation between simulated and observed SWE (correlation coefficient, R = 0.6). Simulated and observed SWE were similar over areas with relatively flat terrain and moderate forest density. The simulated MoD agreed (MoD differences (CLM4-PMW) = +/-7 days) with observations over 39.4% of the study domain. Snowmelt-off occurred earlier in the model compared to the observations over 39.5% of the domain and later over 21.1% of the domain. Large differences of MoD were seen in the areas with complex terrain and dense forest cover. We also found that, although streamflow seasonal phase was accurately modeled (R=0.9), the peaks controlled by snowmelt were underestimated. Routed CLM4 streamflow tended to occur early (by 10 days on average).

Keywords: snow model; snowmelt, snow cover extent; snow water equivalent; streamflow timing; CLM; simulation.
1 Introduction

Seasonal snow cover and snowmelt strongly influence the global climate system and the hydrological cycle (Kukla, 1981). In particular, snow cover regulates the energy balance of the land surface and tends to amplify perturbations in the global atmospheric circulation (Hare and Thomas, 1979; Walsh et al., 1982; Walsh et al., 1985). Additionally, there are possible feedbacks between inter-annual changes of snow cover extent and global climate fluctuation (Bojariu and Gimeno, 2003). Water from snowmelt is an important component of the annual water budget in many parts of the world. For example, over 50% of the water supply in the western United States originates from mountain snowmelt, and in 73% of these basins, snowmelt onset has advanced by one to three weeks due to global warming (Cayan et al., 2001; Regonda et al., 2005; Stewart et al., 2005; Lundquist et al., 2009). Arctic snow cover declined at a rate of about 18% per decade in June over the 1979-2014 time period, largely due to global-warming induced delay of onset of snow accumulation and earlier spring melt (Derksen et al., 2016).

Monitoring snow cover extent, water equivalent, and melting onset is essential for hydrological forecasting in cold regions. Predicting snowmelt runoff is of great importance to hydrologists and decision-makers, as it has consequences for hydroelectric power generation, fresh water supply, agricultural production, and flood control. Because of the strong sensitivity of the climate system to snow cover, largely through its control of surface albedo, accurate parameterization of snowmelt onset is crucial in global climates models. Better simulation of the onset of snowmelt would improve climate simulations, and snowmelt onset is also a useful climate indicator (Drobot and Anderson, 2001).
Version 4 of the Community Land Model (CLM4; Lawrence et al., 2011; Oleson et al., 2013) is a state-of-the-art land surface model (LSM) developed primarily as the land component of the Community Earth System Model. The CLM community is large and the model has benefitted greatly from the contributions of that community, including evaluation studies such as this one. CLM4 represents snow as a granular multi-layer medium and simulates snow cover fraction, snow depth, and snow water equivalent at any prescribed time step. Recently a newer version of CLM (4.5) has been released, however, the snow parametrization of this latest version is not fundamentally different from that of CLM4 as far as the subgrid snow representation (topography, land cover and land use) is concerned. Toure et al. (2016) performed an evaluation of CLM4, using in situ observations including snow telemetry (SNOTEL, Serreze et al. 1999) and cooperative station (COOP, Durre et al. 2010) observations, a satellite product (Moderate Resolution Imaging Spectroradiometer (MODIS) Hall et al., 2006a, 2006b), a satellite-based product (Interactive Multisensor Snow and Ice Mapping System (IMS, Ramsay, 1998, Helfrich et al. 2007), and blended model-ground-based observations (Canadian Meteorological Centre (CMC), Brasnett, 1999; Brown and Brasnett, 2010). The results showed good agreement between simulated snow cover fraction and snow depth in most areas of the globe during their snow accumulation seasons, but they also suggested a model tendency toward premature snowmelt. This study complements the work of Toure et al. (2016) by evaluating the model snowmelt and its impact on simulated streamflow.

Previous studies have evaluated snow schemes in LSMs. Reichle et al., (2011; 2017) and Toure et al., (2018) evaluated the snow model in the NASA Catchment LSM (CLSM, Koster et al.,
2000; Ducharme et al., 2000), the land model component used for the Modern-Era Retrospective Analysis for Research and Applications (MERRA, Rienecker et al., 2011) and its latest version MERRA-2 (Gelaro et al., 2017). Results showed a reasonably good agreement with in situ snow depth observations and against CMC data. Better agreement was found between the MERRA-2 and CMC snow depth, especially in areas with good ground-based observations used to constrain the CMC snow depth product. Minder et al. (2016) evaluated snow albedo estimates from the Noah LSM (Ek et al., 2003) and Noah-MP (Multi-parameterization, Niu, G.-Y., et al., 2011) using MODIS observations. While the two models were able to reproduce the general observed seasonal and spatial variability of snow cover, they also exhibited significant biases. Wang et al. (2013) evaluated snow related outputs including albedo from a variant of the ORCHIDEE (Organizing Carbon and Hydrology in Dynamic Ecosystems, Krinner et al., 2005) LSM. The results showed good skill compared with in situ snow depth and snow water equivalent observations at Col de Porte station (France) and sites in northern Eurasia. Uncertainty in snow albedo was shown to be the main reason for discrepancies between simulations and observations. The Snow Model Intercomparison Project (SnowMIP, Etchevers et al 2004; Brown et al., 2006), a comparison of 23 LSMs showed that the Canadian Land Surface Scheme (CLASS, Verseghy, 2010) with a single-layer snow model was among the better performing models along with more sophisticated multi-layer snowpack models such as CROCUS (Brun et al., 1989, 1992). A subsequent intercomparison of 33 snow models at forest sites (SNOWMIP2, Essery et al., 2009) showed that all models generally predicted the SWE seasonal cycle with some broad differences in the magnitude of snow accumulation. A study by Dutra (2012) investigated three different levels of complexity in snow schemes implemented in the Hydrology Tiled ECMWF Scheme of Surface Exchanges over Land (HTESSEL; Balsamo et al., 2009) LSM. The results showed that
stand-alone multi-layer model runs were comparable to a single-layer snow model run in shallow snow areas. The results also showed that the multi-layer model outperformed the single-layer model in open, deep snowpack conditions. Parajka et al. (2010) evaluated snow depth and snow cover from the Joint UK Land Environment Simulator (JULES, Best et al., 2011; Clark et al., 2011) LSM. The results indicated good agreement between the JULES simulations and observations of snow cover at climate stations. However, poor model performance was observed in areas with complex and heterogeneous terrains.

Other studies have also evaluated CLM4 streamflow over a mountainous catchment in the U.S. and globally (Li et al., 2011; 2015). The analyses of Li et al. (2011) were limited to CLM4 estimates of streamflow at five stations in the Western U.S., while Li et al. (2015) used CLM4 coupled to the Scale Adaptive River Transport (MOSART) model to estimate global annual, seasonal, and daily average runoff. Here, we focus on the high latitudes where the quality of snowmelt estimates has a direct impact on the accuracy of simulated streamflow.

The objectives of this study are threefold: 1) Evaluate the performance of the CLM4 snow model using the latest version of the gridded satellite-based snow cover extent (SCE) and snow water equivalent (SWE) produced by the European Space Agency (ESA) GlobSnow project (Takala et al., 2011, Metsämäki et al. 2015); 2) Use a unique satellite-based snowmelt-off day product (Takala et al., 2009) to evaluate the modeled snowmelt; and 3) Evaluate the impact of simulated snowmelt driven runoff using a coupled river routing scheme. Snowmelt-off day is the day of the year when the last of the snow melts, indicating the end of the snow season. We qualitatively discuss inter-annual variability in snowmelt-off day by examining annual snowmelt-off day of
year (denoted as MoD here onwards) and its spatial distributions, and the mean, range, and time
series of snowmelt-off as a function of topography and forest density. This study also attempts
to relate modeled snow cover and snow mass to the timing of snowmelt and its impact on
streamflow timing.

The paper is organized as follows: Section 2.1 and 2.2 describe the LSM and the river routing
scheme, section 2.3 describes the validation datasets, and sections 2.4 and 2.5 present the
experimental design and evaluation metrics, respectively. Results and discussion are presented
in sections 3 and 4. A summary and conclusions are presented in section 5.

2 Method and Data

2.1 Model Description

CLM4 is the land surface parameterization that provides the lower boundary conditions for the
atmosphere in the coupled Community Earth System Model (CESM1.0.4; Gent et al., 2011;
Vertenstein et al., 2012). CLM4 includes a state-of-the-art five-layer snow scheme, 10-layer soil
scheme, and single-layer vegetation scheme. The one-dimensional vertical layer snow model
(Dai et al., 2003; Kluzek, 2012) is based primarily on parameterizations developed by Anderson
forcing, CLM4 simulates processes such as snow accumulation, depletion, densification,
metamorphism, percolation and refreezing of water (Lawrence and Slater, 2009). A
sophisticated snow densification scheme is employed. The computed density is then used to
calculate snow depth based on the SWE while taking into consideration the snow static sublimation. The snow cover fraction (SCF) is calculated as a function of snow density and snow depth (Niu and Yang, 2007).

Snow effective grain size is parameterized as a function of temperature, layer temperature gradient, and density (see Flanner and Zender, 2006). Snow albedo, a factor that controls snowmelt, is calculated as a function of snow metamorphism stage, effective grain size, solar angle, and pollution (Warren and Wiscombe, 1980; Flanner and Zender, 2005; Flanner and Zender, 2006; Flanner et al., 2007). The influence of impurities, such as black carbon and mineral dust on the albedo, is also parameterized in the model. A snow layer is initialized when the snow depth is greater than or equal to 0.01 m and the surface temperature is below freezing. Depending on whether there is accumulation or depletion of snow, snow layers are combined or subdivided. Excess liquid water from rain or snowmelt within a snow layer percolates to the underlying layer whenever the layer’s water holding capacity is reached. 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). The energy available for snowmelt is calculated as follows (Oleson et al., 2004):

\[(1 - \alpha_s)SW_{down} - (LW_{up} - LW_{down}) - SH - LH - G = 0 \quad (1)\]

where \(\alpha_s\) is the snow/soil surface albedo, \(SW_{down}\) is the downward solar radiation (W.m\(^{-2}\), positive down), \((1 - \alpha_s)SW_{down}\) is the absorbed solar radiation (W.m\(^{-2}\), positive down), \(LW_{up}\) is the upward longwave radiation from the surface (W.m\(^{-2}\), positive up), \(LW_{down}\) is the downward
atmospheric longwave radiation (W.m\(^{-2}\)), LW\(_{\text{up}}\) - LW\(_{\text{down}}\) is the net longwave radiation (W.m\(^{-2}\), positive up); SH is sensible heat flux (W.m\(^{-2}\), positive up), LH is latent heat flux (W.m\(^{-2}\), positive up), and G is ground heat flux (W.m\(^{-2}\) includes snowmelt, positive into the soil/snow surface).

2.2 The Hydrological Modeling and Analysis Platform (HyMAP) River Routing Scheme

The Hydrological Modeling and Analysis Platform (HyMAP; Getirana et al., 2012) is a global scale river routing scheme capable of simulating flow dynamics in both rivers and floodplains. The model is composed of modules representing advance physical processes such as 1) the surface runoff and baseflow time delays; 2) the flow routing in river channels and floodplains, and 3) the evaporation from open water surfaces. The runoff and baseflow generated by a LSM pass through the surface water and groundwater linear reservoirs, respectively, and are then routed using a kinematic wave formulation through a prescribed river network to oceans or inland seas. The river network is represented by a river channel reservoir and a floodplain reservoir in each grid cell. Both are treated as continuous reservoirs where the water spilling from the river channel is stored in the floodplain.

The model simulates water level, discharge, and storage in rivers and floodplains at a daily time step. The internal forward integration time step varies from a few minutes to several hours. At each time step, the inflow water is redistributed between the river channel and floodplain reservoirs following stage–volume relationships derived from the topography of each grid cell. HyMAP has been extensively evaluated on a regional scale over the Amazon basin (Getirana et al., 2012; 2014; 2017a), East Africa region (Jung et al., 2015), and the Middle East (Getirana et al., 2015); on a continental scale over North America (Kumar et al., 2015); and on a global scale (Getirana et al.,
In this study, HyMAP is utilized with a local inertia formulation, which enables a more stable and computationally efficient representation of backwater effects. The Courant–Freidrichs–Levy condition is used in order to determine HyMAP’s optimal time steps for numerical stability (Bates et al., 2010; Getirana et al., 2017b). More details on the model parameterization can be found in Getirana et al. (2012, 2013, 2017b).

2.3 Validation Datasets

Table 1 summarizes the models and datasets used in this study including 1) GlobSnow snow cover fraction and snow water equivalent, 2) a snowmelt-off day product derived from passive microwave data, and 3) Global Runoff Data Centre discharge gauging station data.

2.3.1 GlobSnow Snow Cover Fraction and Snow Water Equivalent Products

The European Space Agency GlobSnow project has produced two new hemispheric records of snow parameters intended for climate research purposes. The datasets contain satellite-retrieved information on SCF and SWE, extending 17 and 37 years, respectively. The dataset on SCF (Metsämäki et al., 2015) is based on optical data of Envisat Advanced Along-Track Scanning Radiometer (AATSR) and European Remote-Sensing Satellite (ERS)-2 the Along Track Scanning Radiometers (ATSR)-2 sensors covering Northern Hemisphere (NH) between years 1995 to 2012. The SWE record (Takala et al. 2011) is produced using a combination of passive microwave radiometer data and ground-based weather station data for the period 1979 to present.
a) GlobSnow Snow Cover Fraction Product

The GlobSnow SCF processing system applies optical measurements in the visual-to-thermal part of the electromagnetic spectrum acquired by the ERS-2 sensor (ATSR-2), and the Envisat sensor (AATSR). The snow cover information is retrieved using the semi-empirical SCAmod-algorithm (Metsämäki et al., 2015). Clouds are masked out using a cloud-cover retrieval algorithm developed for GlobSnow (Metsämäki et al., 2015). Large water bodies (ocean and lakes) are also masked out. The resulting product is provided on a latitude-longitude grid with a 0.01° spatial resolution (approximately 1 km) covering the Northern Hemisphere from 25°N to 84°N (excluding Greenland). This corresponds to the seasonally snow-covered land areas of the Northern Hemisphere.

The SCAmod method was originally developed for Northern European boreal forest and sub-arctic regions to support the Finnish Environment Institute (SYKE)’s hydrological forecasting. The semi-empirical reflectance model-based method SCAmod originates from radiative transfer theory. It describes the scene reflectance as a mixture of three major constituents (opaque forest canopy, snow, and snow-free ground) all of which are interconnected through the apparent forest transmissivity and the snow cover fraction. The daily product provides the SCF as a percentage (%) in each grid cell for all satellite overpasses in a given day. If there are multiple snow observations within a single day, the satellite observations applied are those acquired under the highest solar elevation. The SCF is provided only for observations at sun zenith angle < 73°.
GlobSnow SCF product retrieval accuracy has been assessed using ground-based reference data and satellite-based snow products (Metsämäki et al., 2005; Metsämäki et al., 2012). Results showed mean differences on the order of about 10% SCF, with mean bias values between -3.5% and 3.5%, and mean standard deviation of about 25%. In general, partial discrepancies between the GlobSnow SCF estimates and other snow products/observations were found in forested and mountainous areas, while agreement was typically good in non-forested plain areas (Metsämäki et al., 2015). A bi-weekly aggregation from daily SCF was used in this study.

b) GlobSnow Snow Water Equivalent Product

GlobSnow SWE data production is based on a data assimilation-based approach that combines space-borne passive radiometer data (SMMR, SSM/I and SSMIS) with data from ground-based synoptic weather stations (Takala et al., 2011). The satellite sensors utilized provide data at K and Ka bands (19 GHz and 37 GHz) at a spatial resolution of ~25km. SWE fields are produced daily, weekly, and monthly bases. SWE information is provided for Northern Hemisphere land excluding mountainous regions, glaciers, and Greenland. A semi-empirical snow emission model is used for interpreting the passive microwave radiometer observations through model inversion to the corresponding SWE estimates (Pulliainen, 2006). This SWE retrieval methodology is complemented by a time series melt-detection algorithm (Takala et al., 2009), which is used to determine the snow extent for the final SWE product.

The GlobSnow system has been calibrated against a vast pan-Arctic dataset covering most of the land areas of northern Eurasia for the period 1979 to 2001. The SWE estimation and snow line
detection algorithms are combined to produce SWE maps at a spatial resolution of 25 km × 25 km. The SWE estimates are accompanied by uncertainty information on a grid cell level.

Takala et al. (2011) evaluated the GlobSnow SWE record against distributed snow transect data collected in the former Soviet Union during 1979 to 2009. The evaluation showed that the bias and RMSE were +1.1 mm and 43.4 mm, respectively. The GlobSnow SWE product has been shown to be a robust and reliable product on a continental to hemispherical scale, having matched well with the ground-based reference data as well as the more advanced model-based gridded SWE products (Mudryk et al. 2015).

2.3.2 Snowmelt-off Day Product Derived from Passive Microwave (MoD) Data

The snowmelt-off day product is derived from passive microwave (PMW) data using an algorithm developed at the Finnish Meteorological Institute (Takala et al., 2009). The algorithm utilizes the time series of radiometer data to estimate the yearly snowmelt timing. Satellite-based brightness temperature difference information from frequencies at 18/19 and 36/37 GHz and a sequence of data from consecutive days are used to determine the day of year when the snow has melted from the ground.

The melt algorithm is applied to the Scanning Multichannel Microwave Radiometer (SMMR), the Special Sensor Microwave/Imager (SSM/I), and the Special Sensor Microwave Imager/Sounder (SSMIS) instruments onboard Nimbus and the Defense Meteorological Satellite Program (DMSP) series satellites. A yearly snowmelt dataset has been produced for the
available satellite time series from 1979 to the present day. The algorithm and the resulting dataset have been validated against point-scale ground-based data from the quality controlled INTASS SCCONE data, which are derived from weather station observations (Kitaev et al., 2002), and against optical-based snow-covered area data (Metsämäki et al., 2012). In addition, the algorithm has been assessed with optically derived snow-covered area data. Assessments showing good estimation accuracy are presented in Takala et al. (2009). The snowmelt data was recently assessed for coarse resolution snowmelt timing (Metsämäki et al. 2018) across Europe. These assessments confirmed good applicability as an independent reference for evaluating modeled snow output.

### 2.3.3 Global Runoff Data Centre (GRDC) Discharge Gauging Station Dataset

Daily streamflow data were obtained from the Global Runoff Data Centre (GRDC). The GRDC is a worldwide archive of river discharge data. It operates under the auspices of the World Meteorological Organization (WMO) and is hosted by the Germany Federal Institute of Hydrology. The discharge dataset is publicly available at the GRDC’s website (http://grdc.bafg.de). GRDC has continuously updated data for many of its stations (Grabs, 1997). Daily discharge data from 408 gauging stations located in the northern hemisphere poleward of 30° latitude were used in this study and the analysis was performed for the period from 2000 to 2012. The streamflow gauging stations were screened to meet the requirement that stations drain areas larger than 10,000 km².
2.4 Experimental Design

The meteorological forcing fields used to drive CLM4 were those of the Modern-Era Retrospective Analysis for Research and Applications (MERRA; Rienecker et al. 2011) Land dataset (Reichle et al., 2011; 2012). The MERRA standard precipitation forcing field was bias-corrected with a gauge-based product from the National Oceanic and Atmospheric Administrations’ (NOAA) Climate Prediction Center to obtain the MERRA-Land meteorological forcing dataset. MERRA-land was recently discontinued in favor of MERRA-2 (Gelaro et al., 2017). A study by Reichle et al. (2017) showed that in high latitudes and in an off-line setting (no coupling with an atmospheric model), there were no significant difference between MERRA-Land and MERRA-2 surface meteorological fields. To create the initial conditions, we ran CLM4 for 600 years driven with a repeating 25-year (1979–2004) period of MERRA-Land meteorological fields, satellite-based plant phenology, land use, aerosol and nitrogen deposition, and the CO2 level for the year 2000, as a reference case created at NCAR. The resulting restart file was then used in the subsequent 1 January 2000 to 1 September 2012 simulation using the same initial atmospheric and satellite-derived phenological composition set as for the spin-up process (Toure et al. 2016). CLM4-estimated SWE and SCF were then evaluated against the corresponding GlobSnow SWE and SCF data for the time period 2002-2011.

CLM4 melt off day (MoD) was calculated by assessing the day of the year when snow completely melted from a given grid cell. That day is interpreted as the first day with SWE=0 after winter seasons’ maximum SWE or, alternatively, the first day with SWE=0 with no subsequent snow events until the next snow accumulation season begins in the following autumn. The latter condition may be met substantially later than the former, since ephemeral
snow events may well happen after the original snowmelt-off day (Räisänen et al., 2014). We used the latter definition to define MoD in CLM4 and to determine the satellite-based MoD. Simulated MoDs were then compared against the satellite-derived MoDs over the period from 1 January 2002 to 1 September 2011 for the Northern Hemisphere.

The daily output generated by CLM4, including runoff, baseflow, and evapotranspiration, was fed into the HyMAP river routing scheme to simulate surface water dynamics over the period 1 January 2000 to 1 September 2012. The simulated streamflow was then compared with ground-based observations at 408 stations.

2.5 Evaluation Metrics

We quantitatively evaluated the inter-annual variability in snowmelt by examining snowmelt time series, annual snowmelt distributions, as well as the mean, range, and spatial distribution of MoD. Standard metrics such as mean bias, root mean square error (RMSE), Pearson correlation coefficient (R), and Nash-Sutcliffe model efficiency (NSCE) were used to evaluate CLM4 SCE against GlobSnow SCE. To assess CLM4 SWE, we computed the Pearson correlation coefficient and climatological seasonal cycle over the 9-year period (2002-2011). We also computed anomaly time series by subtracting the climatological seasonal cycle from the original time series. We then computed anomaly correlation R between the model anomaly time series and corresponding anomaly of the observations. A summer mask (excluding June–July–August–September) and minimum data requirement (<5% of Non-Not-a-Number (NaN) values across
the nine-year daily series) were applied to GlobSnow data prior to the computation of bias, RMSE, R, and NSCE.

Daily time series of the SWE bias, RMSE, and R were computed. A mask was applied to the GlobSnow and CLM4 daily gridded SWE fields to exclude locations for which GlobSnow indicated snow-free conditions. Areas with no data in both data sets were also excluded (i.e., Glaciers, mountains & Ice Sheets). For the computation of the domain SCF, the snow cover extent (in km²) of each grid cell was first computed. Snow cover extent of all grid cells were then summed up to obtain the total snow cover extent of the domain on that given day. The domain SCF (in percentage) was then computed by dividing the total snow cover extent of the domain by the total area of the study domain as defined in Table 1.

Snowmelt sensitivity to the presence of forest cover and complex terrain was also evaluated. We used the MODIS-derived global forest-cover density (fd in g / cm³) data mask (Hansen et al., 2003) to evaluate the impact of the forest density for fd<20, 20<fd<40, 40<fd<60 and fd>60. We also used the global forest cover fraction (ff, ranging from 0 to 1) data derived from observations from the same MODIS sensors as the mask, and computed relative bias, RMSE, and correlation coefficient R between CLM4 and GlobSnow SCE for areas with different ff. Topographic information for the masking of complex terrain was derived from ETOPO5 (National Geophysical Data, 1988) data, by calculating topography variances within each product grid cell. For the GlobSnow SWE product, a grid cell was considered as mountainous if the standard deviation of the elevation within that grid cell was above 200 m.
To evaluate the performance of the daily streamflow estimates, we used the delay index (DI in days) and the ratio of standard deviations (\( \gamma = \frac{\sigma_s}{\sigma_o} \)) of the simulated (\( \sigma_s \)) over the observed (\( \sigma_o \)) streamflow for the time period from 1 January 2000 to 1 September 2012. The delay index measures errors related to the time delay between simulated and observed hydrographs (Getirana et al., 2014). Positive DI values indicate delayed simulated hydrographs while negative DI values indicate advanced simulated hydrographs relative to the observations. The optimum value of the ratio of standard deviations is 1. Values lower than 1 indicate simulated streamflow that is less variable than the observations and values greater than 1 indicate that the model overestimates streamflow variability.

3 Results

SWE - the amount of water contained in a snowpack - is one of the most important snow model outputs. SWE simulation is primarily affected by the quality of precipitation and temperature forcings and the model parameterizations of snowmelt and sublimation as well as land cover and the topography of the terrain.

First, SCF and SWE were evaluated against daily satellite-based GlobSnow observations to provide insights into the skill of snow cover fraction and snow mass from CLM4. The snowmelt and streamflow timing were then evaluated using daily satellite-derived PMW MoD and ground-based streamflow observations, respectively.
3.1 Comparison against GlobSnow Snow Cover Extent and SWE Product

3.1.1 Comparison against GlobSnow Snow Cover Extent

The time series of monthly CLM4 and GlobSnow SCE for the Northern Hemisphere poleward of 30°N during 1 September 2002 – 1 September 2011, are presented in Figure 1. Table 2 presents the summary of statistics between CLM4 and GlobSnow SCE. There is a strong agreement between the simulated SCE and the GlobSnow product with a correlation coefficient of 0.99, accompanied by Nash-Sutcliffe model efficiency (NSCE) of 0.98 and low RMSE of 1.2 x 10^6 km^2. The peaks of maximum SCE occur in February. The slightly positive bias of 0.99 x 10^6 km^2 suggests that overall the model tends to overestimate SCE. The overestimation mainly occurs during the accumulation phase from mid-September to mid-March as shown by the climatology of SCE (Figure 2). The maximum difference occurs in early October (Figure 2b). SCE approaches zero in July and August (recall that Greenland was excluded from the analysis). From mid-March to August, the model tends to underestimate SCE with the peak difference occurring in May (Figure 2b). The model simulated snowmelt-off day is assessed in section 3.2.

The time series of the spatial distribution of the difference between the climatology (2002-2011) of CLM4 and GlobSnow (Figure 3) confirms a strong agreement. The model slightly overestimates SCE during the accumulation phase (October-November-December) in northern Canada, in the Arctic, Alaska and Eurasia. From March to May, the model underestimates SCE near the 49th parallel North (the US-Canadian border) and in southern Russia, Western Europe and the Tibetan plateau. Toure et al. (2016) observed the same pattern of CLM4 model underperformance compared to MODIS SCE, which was attributed to a parameterization deficiency in CLM4 SCF model (Swenson and Lawrence 2012).
Snow falling in forested areas can either be intercepted by the vegetation, penetrate through the canopy and reach the soil surface, or accumulate onto the leeward side of trees due to turbulent airflow above and within the canopy (Liston and Hiemstra, 2011). During the snow accumulation phase, intercepted loss through sublimation and throughfall increases. Both interception and snowdrift around trees created by the turbulent fluxes form a heterogeneous distribution of the snow in the forest. The distribution patterns of snow in the forest are related to the vegetation type, vegetation density, and presence of nearby clearings (Pomeroy and Gray, 1995). Model SCE metrics (taking GlobSnow as truth) as a function of the forest cover fraction showed that areas with little forest cover (ff $<$ 0.25), which are known to have a high GlobSnow snow detection coefficient, also show a strong agreement with CLM4 with a bias, RMSE and R of $0.13 \times 10^6$ km$^2$, $0.37 \times 10^6$ km$^2$ and 0.99, respectively. For moderate forest density ($0.25 < ff < 0.5$) the bias does not significantly change but the RMSE doubles (Table 3). At high forest cover ($ff > 0.75$), the bias, RMSE and R increase by 75%, 4% and 10%, respectively, compared to moderate forest cover. Dense forest is a known challenge for optical satellite-based retrieval methods. Even though the GlobSnow retrieval approach compensates for forest effects and reduces the underestimation of SCE in dense forests, uncertainty remains higher than in non-forested regions, which may in part explain the discrepancies with the model.

3.1.2 Comparison against GlobSnow SWE Product

Simulated and observed SWE time series in the Northern Hemisphere from 30°N and statistics (bias, RMSE, and correlation coefficient) are shown in Figure 4. For each year, bias is low
during the accumulation season and increases rapidly from March to the end of the ablation season. Average bias for the study period is 8.6 mm. The RMSE shows a similar pattern with an average over the 9-year study period of 48.2 mm. There is also strong correlation throughout the time series. The Northern Hemisphere average correlation coefficient for the study period is 0.6. For each seasonal cycle, the correlation starts low at the beginning of the season (R = ~0.2) and increases significantly with the accumulation of snow to reach values between 0.6 and 0.8. The skill starts to deteriorate from the onset of the snowmelt to the end of the ablation.

We investigated the correlation coefficient for different regions of the Northern Hemisphere in Figure 5. GlobSnow and CLM4 have the strongest correlation (R = 0.8) in Siberia. There is no significant difference in model performance between the East and West Siberia. Alaska and the Tibetan Plateau, two regions with complex terrains, have the lowest skill. Mountainous areas are characterized by heterogeneous snow distribution due to wind-induced erosion and deposition of snow on the lee sides of mountains. This redistribution of snow often occurs when wind speeds exceed 5 m.s\(^{-1}\) (Liston et al., 2007). The turbulent flux that is responsible for snow erosion also leads to large snow losses to the atmosphere through sublimation (Male, 1980; Pomeroy, 1988). CLM4 lacks the capability to simulate those processes. However, CLM4 performed well in Europe (R = 0.6). Model performance in Western Europe (R = 0.60) is less than that of the Eastern Europe, where the model performance is close to that of Siberia (R = 0.7). Western Europe is dominated by thin and wet snow cover, which is a known challenge for passive microwave retrievals (i.e. GlobSnow), which may degrade correlations. Across Canada, the model performance is modest (R = 0.5), due to the diversity of land covers and snow types. Tundra regions have a lower correlation coefficient (R = 0.4). There are very little ground-based
observations used to drive GlobSnow SWE retrieval, and also the precipitation forcing data used to run CLM4 are more commonly erroneous in those areas due to gauge undercatch. These regions are also subject to snow erosion and sublimation losses, processes which the model does not simulate. Regions with mixed-forest and boreal forest in Canada have the same performance (R = 0.5). The retrieval of SWE using passive remote sensing data is challenged in areas with dense forest canopy or deep snow packs as discussed in previous section. Imperfectly modeled interception losses also add uncertainty to the SWE output. Other regions with lower performance in Canada are the Prairies, with R = 0.48. In these regions, sublimation is the main process that controls snow cover depletion. Wind-induced sublimation removes up to 74% of annual snowfall in some areas of the Prairies (Pomeroy and Gray, 1995). Earlier model snowmelt onset discussed in section 3.1.1 may also explain this relatively poor performance. Contiguous United States (CONUS) model performance is fairly good overall (R = 0.6). Better performance is seen in the eastern and mid-western regions with correlation coefficients of 0.6 and 0.7, respectively. Low correlations are seen in the western U.S. and Rocky Mountains (R = 0.4), where the simulation of snow in the complex terrain is not accurately represented in the model.

In order to evaluate the model’s ability to simulate temporal variability of SWE, which is often more important than its absolute magnitude, we calculated the departure (anomaly) of CLM4 and GlobSnow daily SWE from their long-term (9 years) averages for the Northern Hemisphere (poleward of 30°N). We then computed the anomaly correlation. The spatial representation of the anomalies (Figure 6) shows good agreement between the model and observation in areas including parts of Canada and Siberia for all seasons. During the build-up and maximum snow
accumulation phases (OND and JFM), the mean anomaly correlation coefficients equal 0.5, and they degrade to 0.4 during the melting season (AMJ). This degradation of skill may in part be explained by the model tendency toward premature snowmelt.

3.2 Evaluation of Model Snowmelt

In this section, we use a satellite-derived snowmelt product (PMW MoD) to evaluate the CLM4 modeled snowmelt. First, we assess the mean difference between the model and PMW MoD over the North Hemisphere during the period 2002-2011 (Figure 7). There is good agreement (MoD differences (CLM4-PMW) = +/-7 days) between the model and PMW over 39.4% of the study domain. Premature model snowmelt-off (MoD differences < -7. days) occurs over 39.5 % of the study domain. Late model snowmelt-off compared to PWM observations occurs in 21.1% of the domain. In 12.3% of the domain, simulated snowmelt-off occurs two weeks earlier than observed. In 8.6 % of the domain, melt-off occurs between 15 days and three weeks too early. In 5.9 % of the domain, early melt-off occurs between 22 days and 4 weeks too early. Finally, in 12.8 % of the domain, the melt-off occurs more than a month too early. Snowmelt-off occurs more than a month too early in areas that include southern Russia, the Tibetan Plateau, the Rocky Mountains, and the Canadian Prairies. Modeled snowmelt-off occurs in average 10 days late (areas in yellow) in a large part of Eurasia and small areas of Alaska and the Eastern U.S. Snowmelt-off occurs 20 days to a month late in Central and Eastern Europe and Northern Siberia (areas in orange and red). Time series of the frequency distribution of CLM4 and PMW MoDs from 2002 to 2011 (Figure 8) show a similarity between the datasets. For every year from 2002 to 2011, model frequency distributions are equal to or narrower than PMW observations.
Snowmelt is controlled not only by energy availability but also by forest cover and topography. We investigated the impact of the land features on the model simulation of snowmelt-off. Figure 9 represents the errors of CLM4 with respect to PMW observed snowmelt-off as a function of forest density (fd in g / cm³). For most years, the snowmelt-off occurred earlier in the model than in PMW for all evaluated forest density classes. The errors generally increase as forest density increases from moderate (20 <fd<60 g / cm³) to high (fd>60 g / cm³). However, the average error over the study period for fd < 20 g / cm³ is -11.3 days, which is the same as for areas with high forest density (fd> 60 g / cm³). For 20<fd<40 and 40<fd<60 g / cm³ average errors of MoD were -3 days and -6 days, respectively. Areas with low forest density are mostly in the Canadian Prairies, U.S. Rocky Mountain, most of the Midwestern U.S., southern Russia, and Western Europe. In these areas, the model typically simulated an early melt-off. CLM4 is also known to underestimate SWE in those areas (Toure et al. 2016).

Figure 10 shows time series of average differences between CLM4 and PMW MoD for the Northern Hemisphere. Simulated snowmelt-off timing is improved when areas of high elevation are excluded from the analysis. Average errors remain constant from 2002 to 2008 and decrease during the last 3 years of the study. The simulation resolution (0.5°) is too coarse to represent terrain elevation precisely. Finer terrain resolution would help to improve the representation of snow accumulation, redistribution, and melt (Jin and Wen, 2012).
3.3 Streamflow Evaluation

The delay index (DI) is presented in Figure 11. Overall, it indicates good model performance in the Eastern and Central United States as well as in Western Europe with DI=±5 days in ~48% of the stations. Conversely, at 35% of the stations negative DI values indicate advanced simulated hydrographs compared to the observations. These stations are mostly located in the Western U.S., Canada, and Northern Europe, where the flows arrive a week to more than 4 weeks too early. Early peaks in high latitude basins simulated by CLM4 were also reported by Li et al. (2015). Simulated streamflow was late (positive DI >5 days) at 15% of the stations. A clear regional pattern emerges between eastern and western North America. CLM4 is better at reproducing the snowmelt timing in the Eastern U.S. and tends to melt snow prematurely in the Western U.S. The amplitude of streamflow was underestimated at 69% of the stations (values less than 1, in blue and green in Figure 12), which is consistent with previous findings in snowmelt-dominated basins when other LSMs were evaluated using a river routing model (Yang et al., 2015). The seasonal cycle of simulated and observed discharge showed increases from March to mid-June in response to snowmelt and decreases through August (not shown). The model correctly reproduced the spatial variability of the observed discharge (R=0.9). However, low bias is apparent in Figure 12, resulting in average of -99 m.s⁻¹ (RMSE of 100 m.s⁻¹). The seasonal variation follows characteristics of high latitude snowmelt-dominated basins (Clow, 2010). The discrepancies between simulated and observed discharge reflect uncertainties in the simulation of SWE and the timing of snowmelt as shown in Figure 11.
The analyses showed that CLM4 simulated snow strongly agrees with GlobSnow observations (in terms of both snow cover extent and snow water equivalent). Seasonal variations in snow cover are well reproduced by the model. However, the results also showed overall a slight overestimation of SCE by the model during the accumulation phase and an underestimation during the ablation phase. The seasonal and spatial patterns of SCF were similar to those of SWE. Regional variations in the results were also observed. The best model performance was seen in areas with relatively flat terrain and low to moderate forest density. Poor model performance was found in the Tundra, areas with complex terrain, deep snow cover, or dense vegetation cover. The analyses also revealed a model tendency toward premature snowmelt, which impacts streamflow timing. The comparisons of simulated SCE with the GlobSnow data produced outcomes similar to those reported by Toure et al. (2016), who used MODIS and IMS data for evaluation of CLM4 SCE.

Most of the discrepancies between CLM4 modeled and observed SCE and SWE can be attributed to the following: 1) simplifications in the CLM4 snow model parameterization; 2) uncertainties in the forcing data; 3) uncertainties related to the retrieval of snow cover information from satellite measurements; and 4) coarse representation of topography and land cover. Parametrization of processes such as snow erosion and its redistribution in forest clearings is lacking in CLM4. In areas such as the Rocky Mountains and the Boreal forest, the lack of representation of sublimation and turbulent transport of snow caused by interactions between snow, wind, and topography (Liston et al., 2007), prevent the model from reproducing
the effects of snow erosion, transport, loss, and redistribution. Errors in precipitation and the other forcing fields also degrade the model output, through no fault of the model. These errors can be blamed on the scarcity of ground-based observations (especially in high latitudes) for use in constructing, calibrating, and validating forcing data fields, particularly precipitation (Reichle et al. 2017). The measurement instruments that are available in those areas do not always capture the true snowfall rate due to wind-induced precipitation undercatch (Goodison, 1978, Groisman et al. 1999). The aforementioned modelling issues combined with the scarcity of ground-based snow observations used in GlobSnow SWE retrieval make it difficult to ascertain the closeness of the simulated and GlobSnow snow data to reality.

Overall, this evaluation effort lends confidence in the current CLM4 snow model while also suggesting where additional development effort should be directed in the future. The coarseness of the topographic information (0.5° × 0.67°) is one thing that could be improved upon fairly easily. The coarse topography limits the model’s ability to simulate snow cover in alpine regions because it smooths out terrain features such as altitude, slope, and curvature. This has many implications: first, it can reduce the environmental lapse rate, leading to an underestimate of the fraction of precipitation that is modeled as snowfall. Secondly, uncertainty in subgrid representation of gentle rolling slopes (slope < 12 degrees), characteristic of the Arctic tundra environment, can lead to errors in the spatial distribution of snow. Inability to represent snow drifts can cause a model to melt-off snow too early.

GlobSnow provides what is among the most robust and reliable SWE products available on a continental to hemispherical scale (Mudryk et al. 2015), making it a valuable resource for model
validation. Nevertheless it is not perfect, and uncertainty in GlobSnow SCE and SWE fields complicate model evaluation. The GlobSnow SWE product is derived from satellite-based radiometer measurements (SMMR, SSM/I, and SSMIS) combined with ground-based weather station observations. It has been evaluated extensively using ground truth data gathered from Canada, Scandinavia, Russia, and the Alps (Takala et al., 2015). However, there are not enough snow stations in the Canadian Arctic for a satisfactory validation of the product in that region; the best validated areas are Northern Eurasia and Northern Europe. Furthermore, passive microwave SWE retrieval algorithms tend to systematically underestimate SWE in deep snow areas (Pulliainen and Hallikainen, 2001; Clifford, 2010). This is due to the saturation of the microwave signature (channel difference), which starts to degrade the retrievals when SWE exceeds 150 mm. Moreover, passive microwave retrieval is not possible for wet snowpack. Thus, for regions of deep or wet snow GlobSnow applies a kriging interpolation scheme to estimate SWE, based on the weather station observations. As a result, the accuracy of the product diminishes during the spring melt period. Another source of uncertainty in GlobSnow SWE is that effective snow grain size from ground-based snow observations is used as a fitting parameter in the SWE retrieval algorithm, yet spatial variability of the snow grain size is rarely well captured because of the sparsity of ground-based observation stations. Finally, errors related to the challenges of retrieving SWE in forested regions and mountains also add to uncertainty in the GlobSnow product. As a result of all this, uncertainty in the GlobSnow SWE product may increase the discrepancies between simulated and GlobSnow SWE, especially in Northern Canada and Eastern Eurasia.
In addition to evaluating SCE and SWE themselves we paid particular attention to the simulation of snowmelt in CLM4. In general, the model tends to melt snow early, particularly in dense vegetation. The term LWdown (Eq. (1)), which represents the downward longwave radiation from the atmosphere and the bottom of the forest canopy to the surface, seems to be overestimated in the model. When snow and vegetation coexist in a model grid cell, incoming shortwave and longwave radiation are partitioned into vegetation interception and surface radiation. Canopy density underestimation leads to an underestimation of the part of energy intercepted by the vegetation, which leads to more energy available for snow melt. Surface albedo, used to calculate net solar radiation, could also be a source of uncertainty. Grid cell albedo (Eq. (1)) is calculated as an area-weighted average of the albedos of vegetation, snow surface, and other coexisting surface types. Therefore, inaccuracies in the input land cover map would lead to erroneous albedo.

A realistic representation of the terrain is critical for accurate simulation of snow accumulation and melt, as coarse model resolution causes underestimation of SWE in areas with complex topography (Jin and Wen, 2012). Increasing resolution would enable better representation of high elevations and steep slopes, which would allow more snow deposition at higher altitudes, where the temperature is lower. Moreover, the orientation of the slope is a determining factor in the rate of melt, as snowmelt is accelerated in south facing slopes due to the more direct incidence of sunlight and reduced shading (Ellis et al, 2013).

There are several energy fluxes involved in the snowmelt process. The key drivers are direct solar radiation (shortwave radiation, as described in Eq. (1)) and energy convergence.
Ground heat flux is the energy delivered to the snowpack from the soil below by conduction at a more or less constant rate. Latent heat flux, which is the energy removed from or supplied to the snowpack through sublimation, evaporation, or condensation, is relatively small. During the time period of study, the snowmelt season did not trend significantly. However, inter-annual variability was observed due to variations in downwelling longwave radiation, which is controlled by cloud cover and synoptic-scale conditions that generate warm air advection (Mioduszewski et al., 2014), as well as heat and water vapour transport in the atmosphere (Stone et al., 2002). Further understanding the differences in MoD between the model and the observations requires a better understanding of the energy balance terms at the local scale, which emphasizes the need for in-situ observations at high latitudes.

Interannual variability of snowmelt timing could also be related to rapid change of land cover and land use in many snow-dominated basins (Barnett et al., 2008). Barnett et al. (2008) estimated that as much as 60% of perturbation in river flow in the Western U.S. from 1950 to 1999 are human-induced. Moreover, land cover and land use change affect the rate of direct solar radiation. Forest thinning, for instance, may significantly change the insolation of snow cover in forested areas and therefore affect the timing of snowmelt. Reduction of forest cover also results in reduced infiltration and increased runoff.

Finally, it is necessary to note that the river routing scheme could also be a source of uncertainty in simulated streamflow. Many studies (Getirana et al., 2012; 2014; 2015; 2017a,b; Jung et al., 2015; Kumar et al., 2015) conducted using HyMAP have concluded that the primary sources of
uncertainty the meteorological forcing and the vertical water balance computed by the LSM.

However, coarse parameterization of river geometry and roughness and errors in the digital
elevation model used to derive HyMAP parameters also degrade accuracy in the routed
streamflow. Although lakes and reservoirs have an important role in water storage dynamics in
some regions and at the global scale, they are currently not incorporated in HyMAP. As a
consequence, they have been neglected in this study. This means that the amplitude and timing
of simulated water storage components in regions affected by water management might differ
from reality.

Summary and Conclusions

In this study, we used observation-based datasets to evaluate snow mass, snowmelt, and
streamflow timing simulated by the CLM4 LSM. Our findings are consistent with previous
studies showing that the model has a tendency towards early snowmelt and streamflow timing.
We found that snow cover fraction estimates from CLM4 agree well with GlobSnow
observations. We also found that the modeled snow water equivalent estimates agree
reasonably well with GlobSnow SWE estimates, with large spatial variability across the globe.
Seasonal variation of streamflow is well reproduced by the coupled CLM4-river routing model.
However, peak flow amplitudes are greatly underestimated. The results are not uniform across
the globe due to the presence of complex terrain and differences in land cover and land use.
Studies using ground-based observation have shown that the Western United States is experiencing a shift toward earlier snowmelt and snowmelt runoff (Hodgkins et al. 2003; Clow, 2010). CLM4 has the potential to be a useful tool for predicting these changes, as it has one of the most sophisticated snow models currently implemented in a climate model. Its snow metamorphism and snowmelt modules are the key features that distinguish it from many other snow models. However, a few upgrades to the model would be helpful for improving snowmelt and streamflow timing, including: 1) better subgrid representation of land cover and land use; 2) higher resolution topographical information; and 3) representation of the wind-induced snow erosion (including sublimation), transport, and redistribution, which are key factors in many cold regions. Additionally, better meteorological forcing data would greatly benefit CLM4 as well as other models, especially better precipitation data.

6 Acknowledgments

This study was funded by NASA’s Science of Terra and Aqua program. Computing was supported by the NASA High End Computing Program. 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. MERRA data was obtained from NASA’s Modeling and Assimilation Data and Information Services Center (MDISC), managed by the NASA Goddard Earth Sciences (GES) Data and Information Services Center (DISC) (Dataset DOIs: 10.5067/RKPH18KC1Y1T, 10.5067/OQ6B1RHOHB18).
References


Hare, F.K.; Thomas, M.K. Climate Canada, 2nd ed.; John Wiley and Sons Canada Limited: Toronto, ON, Canada, 1979; p. 230.


Kumar, S.V., Peters-Lidard, C.D., Arsenault, K., Getirana, A.C.V., Mocko, D., & Liu, Y., (2015), Quantifying the added value of snow cover area observations in passive microwave snow
depth data assimilation. Journal of Hydrometeorology, 16, 1736-1741. DOI: 10.1175/JHM-D-15-0021.1


Coupled with the Community Land Model. Journal of Hydrometeorology, 16, 948-971.


Oleson, K.W., Dai, Y., Bonan, G., Bosilovich, M., Dickinson, R., Dirmeyer, P., Hoffman, F.,
Technical Note NCAR/TN-461+STR. National Center for Atmospheric Research,

Parajka, J., S. Dadson, T. Lafon, and R. Essery (2010), Evaluation of snow cover and depth
simulated by a land surface model using detailed regional snow observations from

National Hydrology Research Institute Science Report No 7. Environment Canada,
Saskatoon.

University of Saskatchewan, Saskatoon, Sask., 226 pp.

Pulliainen, J., & Hallikainen, M. (2001). Retrieval of regional snow water equivalent from space

zones by assimilating space-borne microwave radiometer data and ground-based

12, 1537-1546.

Räisänen, P., Luomaranta, A., Järvinen, H., Takala, M., Jylhä, K., Bulygina, O. N., Luojus, K.,


List of Figures

1. Time series of monthly GlobSnow and model estimates of monthly maximum SCE (x10^6 km^2) for the Northern Hemisphere northward of 30°N (Greenland excluded). a) The blue solid line is the GlobSnow and the red line is the model estimate. b) The black line represents the difference between CLM4 and GlobSnow SCE (x10^6 km^2); the horizontal red solid line denotes a perfect match between GlobSnow and the model.

2. Average climatology (1 September 2002-1 September 2011) of the maximum snow cover extent from CLM4 and GlobSnow. The red and the blue shading represents the +/- standard deviation of CLM4 and GlobSnow respectively. The red solid line in b) is the line for CLM-GlobSnow equals zero.

3. Time series of spatial distribution of the difference between climatologies (Sept 2002-Sept 2011) of CLM4 and GlobSnow maximum SCF (in percentage). White represents water, ice, snow-free areas or areas with no data (due to cloud cover for example).

4. Time series of statistics (bias, RMSE and correlation coefficient (R)) of simulated SWE relative to the GlobSnow SWE product for the Northern Hemisphere (poleward of 30°N). Black lines represent the domain snow cover fraction. In the top panel, the red line represents bias, in the middle panel, the red line represents RMSE and in the bottom panel, the red line represents correlation coefficient (R).
Figure 5: a) Correlation coefficient between CLM4 and GlobSnow SWE in the Northern Hemisphere and the regional break down for the study period (1 Sept 2002- 1 Sept 2011). The locations of the regions are shown in b). The Northern Hemisphere includes the land area poleward of 30°N excluding Greenland.

Figure 6: Seasonal maps of the anomaly correlation between CLM4 and GlobSnow SWE for the study period (1 Sept 2002- 1 Sept 2011).

Figure 7: Mean different (CLM4-PMW) snowmelt-off Day of Year (DOY) for the years 2002-2011 for North Hemisphere. The DOY values for snowmelt-off were identified by determining the transition from snow-cover (SWE>0) to no-snow (SWE =0) for individual CLM4 pixels, throughout the annual melt period (from 1 January to 30 June).

Figure 8: Distribution of CLM4 and PMW snowmelt day of year (DOY) for Northern Hemisphere. The DOY values for snowmelt-off were identified by determining the transition from snow-cover (SWE>0) to no-snow (SWE = 0) for individual CLM4 pixels, throughout the annual melt period (from 1 January to 30 June).

Figure 9: Errors between CLM4 and PMW snowmelt-off day of year for different forest density (fd in g / cm³) for fd<20, 20<fd<40, 40<fd<60 and fd>60
Figure 10: Errors between CLM4 and PMW snowmelt-off day of year for Northern Hemisphere (NH). Black bars represent errors for North including terrain with high topography and grey bars represent errors calculated when high topography is masked out.

Figure 11: Delay index between simulated and observed streamflow for the period from 1 January 2000 to 1 September 2012.

Figure 12: Ratio of standard deviation between simulated and observed streamflows for the period from 1 January 2000 to 1 September 2012.
Table Captions

Table 1: Summary of the dataset and land surface model used in the study

Table 2: Summary of the statistics of the comparison of GlobSnow and CLM4 Snow cover extent for Northern Hemisphere.

Table 3: Comparison of CLM4 GlobSnow and snow cover extent for various forest fractions over the entire Northern Hemisphere (ff is the MODIS-based forest fraction).
<table>
<thead>
<tr>
<th>Dataset</th>
<th>Description</th>
<th>Resolution</th>
<th>Domain</th>
<th>Parameters and study time periods</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLM4</td>
<td>Land surface model with a Snow model with up to 5 layers</td>
<td>$0.5^\circ \times 0.67^\circ$ Converted into $0.25^\circ \times 0.25^\circ$</td>
<td>Northern Hemisphere poleward of 30°N. (Glaciers, mountains &amp; Ice Sheets masked out)</td>
<td>SCF, SWE Model run period: 1 January 2000 – 1 September 2012.</td>
</tr>
<tr>
<td>GlobSnow SCE</td>
<td>Based on data from ERS-2/ATSR-2 (1995–2003) and Envisat/AATSR (2002–2012)</td>
<td>$0.01^\circ \times 0.01^\circ$ geographical grid Converted into (EASE-grid, $0.25^\circ \times 0.25^\circ$)</td>
<td>Northern Hemisphere poleward of 30°N. (Glaciers, mountains &amp; Ice Sheets masked out)</td>
<td>SCF (Range from 0-100%) 1 September 2002 – 1 September 2011</td>
</tr>
<tr>
<td>GlobSnow SWE</td>
<td>Passive microwave radiometer data+ ground-based synoptic snow observations - Variational data-assimilation</td>
<td>25 x 25 km resolution (EASE-grid, $0.25^\circ \times 0.25^\circ$)</td>
<td>Northern Hemisphere poleward of 30°N. (Glaciers, mountains &amp; Ice Sheets masked out)</td>
<td>SWE 1 September 2002 – 1 September 2011</td>
</tr>
<tr>
<td>PMW snowmelt-off day</td>
<td>Snowmelt-off day product derived from Satellite-derived Passive Microwave data</td>
<td>25 x 25 km resolution (EASE-grid, $0.25^\circ \times 0.25^\circ$)</td>
<td>Northern Hemisphere poleward of 30°N. (Glaciers, mountains &amp; Ice Sheets masked out)</td>
<td>Snowmelt-off day of year (MoD) 1 January 2000 to 1 September 2011.</td>
</tr>
<tr>
<td>GRDC daily discharge</td>
<td>Daily discharge from 408 gauge stations</td>
<td>Point-scale at stations</td>
<td>Stations located above 30° N latitude and draining areas larger than 10,000 km²</td>
<td>Daily data 1 January 2000 to 1 September 2012.</td>
</tr>
</tbody>
</table>
Table 2: Summary of the statistics of the comparison of GlobSnow and CLM4 Snow cover extent for Northern Hemisphere for a time period from 1 September 2002 – 1 September 2011.

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Bias</th>
<th>RMSE</th>
<th>NSCE</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(10^6 km^2)</td>
<td>(10^6 km^2)</td>
<td>[-]</td>
<td>[-]</td>
</tr>
<tr>
<td>Northern Hemisphere</td>
<td>0.99</td>
<td>1.2</td>
<td>0.98</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Table 3: Comparison of CLM4 GlobSnow and snow cover extent for various forest fractions over the entire Northern Hemisphere (ff is the MODIS-based forest fraction [0-1]).

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Bias</th>
<th>RMSE</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[10^6 km^2]</td>
<td>[10^6 km^2]</td>
<td>[-]</td>
</tr>
<tr>
<td>Forest Fraction</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[0-1]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ff &lt; .25</td>
<td>0.13</td>
<td>0.37</td>
<td>0.99</td>
</tr>
<tr>
<td>0.25 &gt; ff &lt; 0.75</td>
<td>0.11</td>
<td>0.7</td>
<td>0.96</td>
</tr>
<tr>
<td>ff &gt; 0.75</td>
<td>0.45</td>
<td>0.63</td>
<td>0.86</td>
</tr>
</tbody>
</table>
Figure 1: Time series of monthly GlobSnow and model estimates of monthly maximum SCE (x10^6 km^2) for the Northern Hemisphere northward of 30°N (Greenland excluded). a) The blue solid line is the GlobSnow and the red line is the model estimate. b) The black line represents the difference between CLM4 and GlobSnow SCE (x10^6 km^2); the horizontal red solid line denotes a perfect match between GlobSnow and the model.
**Figure 2:** Average climatology (1 September 2002–1 September 2011) of the maximum snow cover extent from CLM4 and GlobSnow. The red and the blue shading represents the +/- standard deviation of CLM4 and GlobSnow respectively. The red solid line in b) is the line for CLM-GlobSnow equals zero.
Figure 3: Time series of spatial distribution of the difference between climatologies (Sept 2002-Sept 2011) of CLM4 and GlobSnow maximum SCF (in percentage). White color represents water, ice, snow-free areas or areas with no data (due to cloud cover for example).
Figure 4: Time series of statistics (bias, RMSE and correlation coefficient (R)) of simulated SWE relative to GlobSnow SWE product for the Northern Hemisphere (poleward of 30°N). Black lines represent the domain snow cover fraction. In top panel, red line represents bias, in the middle panel, the red line represents RMSE and in the bottom panel, the red line represents correlation coefficient (R).
Figure 5: a) Correlation coefficient between CLM4 and GlobSnow SWE in the Northern Hemisphere and the regional break down for the study period (1 Sept 2002- 1 Sept 2011). The locations of the regions are shown in b). The Northern Hemisphere includes the land area poleward of 30°N excluding Greenland.
Figure 6: Seasonal maps of the anomaly correlation between CLM4 and GlobSnow SWE for the study period (1 Sept 2002 - 1 Sept 2011).
Figure 7: Mean difference (CLM4 - PMW) of snowmelt-off Day of Year (DOY) for the years 2002-2011 for the North Hemisphere. The snowmelt-off days were identified by determining the transition from snow-cover (SWE>0) to no-snow (SWE =0) for individual CLM4 pixels, throughout the annual melt period (from 1 January to 30 June).
Figure 8: Distribution of CLM4 and PMW snowmelt-day of year (DOY) for the Northern Hemisphere. The DOY values for snowmelt-off were identified by determining the transition from snow-cover (SWE>0) to no-snow (SWE = 0) for individual CLM4 pixels, throughout the annual melt period (from 1 January to 30 June).
Figure 9: Errors between CLM4 and PMW snowmelt-off day of year for different forest density (fd in g / cm$^3$) for fd<20, 20<fd<40, 40<fd<60 and fd>60
Figure 10: Errors between CLM4 and PMW snowmelt-off day of year for the Northern Hemisphere. Black bars represent errors for all terrain and grey bars represent errors calculated when high topography is masked out.
Figure 11: Delay index of simulated CLM4/HyMAP2 streamflow against streamflow observations based on daily data from 1 January 2000 to 1 September 2012.
Figure 12: Standard deviation ratio between simulated CLM4 and observed streamflow based on daily data from 1 January 2000 to 1 September 2012.