Assessments of the Northern Hemisphere snow cover response to 1.5 °C and 2.0 °C warming

Aihui Wang, Lianlian Xu, and Xianghui Kong

Nansen-Zhu International Research Center, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, 100029, China

Correspondence to: Aihui Wang (wangaihui@mail.iap.ac.cn)

Abstract The 2015 Paris Agreement set a goal to pursue a global mean temperature below 1.5 °C and well below 2 °C above pre-industrial levels. Although it is an important surface hydrology variable, the response of snow under different warming levels has not been well investigated. This study provides a comprehensive assessment of the snow cover fraction (SCF) and snow area extent (SAE), as well as the associated Land Surface Air Temperature (LSAT) over the Northern Hemisphere (NH) based on the Community Earth System Model Large Ensemble project (CESM-LE), CESM 1.5 °C and 2 °C projects, as well as the CMIP5 historical RCP2.6 and RCP4.5 products. The results show that the spatiotemporal variations of those modeled products are grossly consistent with observations. The projected SAE magnitude change in RCP2.6 is comparable to that in 1.5 °C, but lower than that in 2 °C. The snow cover differences between 1.5 °C and 2 °C are prominent during the second half of the 21st century. The Signal-Noise-Ratios (SNRs) of both SAE and LSAT over the majority of land areas are greater than one, and for the long-term period, the dependences of SAE on LSAT changes are comparable for different ensemble products. The contribution of an increase in LSAT on the reduction of snow cover differs across seasons, with the greatest occurring in boreal autumn (49-55%) and the lowest occurring in boreal summer (10-16%). The snow cover uncertainties induced by the ensemble variability are invariant over time across CESM members but show an increase in the warming signal between the CMIP5 models. This feature reveals that the physical parameterization of the model plays the predominant role in long-term snow simulations, while they are less affected by internal climate variability.

1. Introduction

Snow mass on the ground is one of the most important surface hydrology elements. Due to its unique physical

properties, such as high albedo, emissivity and absorptivity, low thermal conductivity, and roughness length, snow strongly affects the energy and water exchange between land and atmosphere over cold regions (Zhang, 2005). The snowpack is a moisture reservoir, and it stores rainfall (or snowfall) in winter and recharges the surface runoff and ground water in spring (Zakharova et al., 2011; Belmecheri et al., 2016). It is also an insulator for heat and radiation, which blocks the solar radiation arriving at the soil surface, and it prevents the loss of ground heat to the atmosphere in the winter. At high-altitude areas with snow cover, the ground temperature is usually higher than the air temperature (Stieglitz et al., 2003). Furthermore, ground snow influences the rainfall in remote regions through large-scale atmospheric circulations (e.g., Liu and Yanai, 2002), and it has been extensively used in data assimilation to improve climate predictions (e.g., Dawson et al. 2016).

Snow ablation and accumulation are affected by many factors, such as the land surface air temperature (LSAT) and surface radiation. In general, increases in LSAT enhance the ratio of rainfall to total precipitation over land and speed up the snow melting. As a result, the snow retention time on the ground will be shortened (Smith et al., 2004). During the past three decades, evidence through observations has shown that the annual snow area extent over the Northern Hemisphere (NH) has reduced substantially (e.g., Dye, 2002), and such terrestrial changes are partially attributed to the increase in air temperature (Mccabe and Wolock, 2010). Based on the 5th Coupled Model Intercomparison Project (CMIP5) (Taylor et al., 2012), researchers have found that surface warming can lead to earlier snowmelt (Oki and Kanae, 2006), with a lower rate compared to historical periods due to the reduction in snow cover areas in the projected warmer 21st century (Musselman et al., 2017). The relationship between snow cover and LSAT has been discussed in many studies (Cohen and Entekhabi, 1999; Brown and Robinson, 2011; Brutel-Vuilmet et al., 2013; Mudryk et al., 2017). For example, Brown and Robinson (2011) reported that LSAT explained approximately 50% of the change in spring NH mid-latitude snow area during 1920-2010. Brutel-Vuilmet et al. (2013) also found that the spring LSAT is closely linearly correlated with snow cover in boreal spring, and they further indicated that this relationship would persist from historical to future periods. However, comprehensive assessments of the snow cover response to different warming levels (e.g., 1.5 °C and 2.0 °C above pre-industrial levels, hereafter referred as to 1.5 °C and 2.0 °C for short) have not been extensively performed.

The impacts of global warming on terrestrial variables have been investigated in various studies, most of which have focused on the risks associated with 2 °C warming (Meinshausen et al., 2009; Schaeffer and Hare, 2012). Recently, science

communities have argued that a warming of 1.5 °C would significantly reduce climate risks compared to a 2 °C warming, and the 2015 Paris agreement set a goal to reach a Global Mean Air Temperature (GMAT) below 1.5 °C and well below 2 °C above pre-industrial levels (UNFCCC, 2015). The academic community has shown a great interest in this initiative (e.g., Hulme, 2016; Peters, 2016; Schleussner et al., 2016; Mitchell et al., 2017). The Intergovernmental Panel on Climate Change (IPCC) also plans to propose a special report on the impacts of 1.5 °C warming in 2018 (http://www.ipcc.ch/report/sr15/pdf/information_note_expert_review.pdf). However, present comparison studies regarding the differences between 2 °C and 1.5 °C have all been conducted by analyzing the CMIP5 outputs under the Representative Concentration Pathway (RCP) scenarios (Vuuren et al., 2011; Schaeffer and Hare, 2012; Schleussner et al., 2016). For example, based on the CMIP5 model outputs, Schleussner et al. (2016) assessed the impacts of 1.5 °C and 2 °C warming levels on extreme weather events, water availability, agricultural yields, sea-level rise and risk of coral reef loss and concluded that there are substantial risk reductions with 1.5 °C warming compared to 2 °C warming, thus further demonstrating the regional differentiation in both climate risks and vulnerabilities. Indeed, a 1.5 °C warming is a relatively low warming target to achieve compared to the projections in RCPs. Jiang et al. (2016) showed that the probability of 2 °C warming before 2100 would be 26, 86, and 100% for the RCP2.6, RCP4.5 and RCP8.5 scenarios, respectively, crossing all available CMIP5 model outputs. Based on this premise, there should be a much higher probability of the occurrence of 1.5 °C warming. In fact, the RCPs are not specifically designed for targeting the climate impacts and risks of different warming levels, such as 1.5 °C and 2 °C. In RCPs, the projected rise in surface air temperatures and the greenhouse gas emissions exhibit a near-linear relationship (IPCC, 2014). However, other variables in the climate system do not always change linearly with the surface air temperature, and thus, it is difficult to quantify the changes of some quantities (e. g., snow cover) under specific warming levels using the transient RCP simulations. Regarding the IPCC 1.5 °C special report, Peters (2016) addressed seven existing knowledge gaps in the current literature, for which he suggested "clearly specifying methods for temporal and spatial averaging of temperatures and the desired likelihood to stay below given temperature levels". Therefore, it is necessary to design scenarios under the specified GMAT rising goals.

To achieve 1.5 °C and 2.0 °C goals in line with the IPCC special report, the Community Earth System Model (CESM) research group at the National Center for Atmospheric Research (NCAR) has performed a set of ensemble modeling

experiments under the emulated concentration pathway leading to the stable 1.5 °C and 2 °C warming targets by 2100 (Sanderson et al., 2017). These experiments are the first Earth system model simulation projects targeting both 1.5 °C and 2 °C warming goals. Together with the CESM Large Ensemble (CESM-LE), the above simulations provide the best available datasets to assess the potential impacts and risks on both climatic and environmental elements under 1.5 °C and 2 °C

5   warming levels.

Based on the previously mentioned CESM simulations, CMIP5 model outputs, and the observed snow cover fraction datasets, this study extensively investigates the spatiotemporal change in snow cover over the NH land area for both historical (1920-2005) and future (2006-2100) periods under 1.5 °C and 2.0 °C warming levels, as well as under RCP2.6 and RCP4.5 scenarios. The snow cover reproductions of CESM are evaluated with both in situ and satellite data. The

10   contribution of LSAT change in the snow cover is also quantified. Furthermore, a prominent advantage is that the CESM ensemble simulations provide insight into the impacts of internal climate variability on those surface variables, which is also addressed in this study.

2.   Models, Scenarios, and Data

15   ### 2.1 The CESM and snow cover

The CESM consists of the Community Atmosphere Model version 5, the land surface model version 4.0 (CLM4.0), parallel ocean program version 2, and the Los Alamos sea ice model version 4 (Hurrell et al., 2013). The fully coupled CESM has been used in many studies and adopted in the CMIP5 project (Taylor et al., 2012). The CESM and its performance have been extensively assessed in a special issue of the Journal of Climate

20   (http://journals.ametsoc.org/topic/ccsm4-cesm1). The snow process in the CESM is described in the land component of CLM4, in which the snowpack on the ground is divided into five layers based on snow depth. The life cycle of snow, such as aging, compaction, thawing, and sublimation, are parameterized, and the effects of black carbon, organic carbon, and mineral dust on snow are also considered in CLM4.0 (Oleson et al., 2010).

The SCF is defined as the fraction area of a land grid cell covered by snow. In the CESM, the SCF (f_{sno}) is described as

25   (Niu and Yang 2007; Oleson et al., 2010)

$$f_{sno} = \tanh\left\{\frac{z_{sno}}{2.5Z_0\left[\min(\rho_{sno},800)/\rho_{new}\right]^m}\right\} \qquad (1)$$

where Z_{sno} is the snow depth over the ground, $m = 1$ is a parameter representing the snow melting rate that can be calibrated with observations, $Z_0 = 0.01$ m is the momentum roughness length for soil, $\rho_{new} = 100$ kg m^{-3} is the density of new snow, and ρ_{sno} is the density of current snow, computed as the ratio of snow water equivalent and Z_{sno}. Equation (1) is modified based on satellite and in situ observation data (Niu and Yang, 2007). In the CLM4.0, the SCF directly affects the surface hydrology and heat processes (Oleson et al., 2010). The snow products in the offline CLM4.0 simulation have been well evaluated by both satellite and in suit observations, and the general conclusion is that the model simulations have overall captured the temporal variations in both SCF and snow water equivalence, whereas the model expresses a fast, but too early, snow ablation process (Swenson and Lawrence, 2012; Toure et al., 2016; Wang et al., 2016).

2.2 The CESM-LE project

The CESM-LE is a 40-member ensemble project that uses the fully coupled CESM version 1.1. Under the CMIP5 design protocol, all ensemble simulations have the same specified historical external forcing for 1920-2005, and RCP8.5 is used for the future scenario (2006-2100). The ensemble member No. 1 was run continuously from 1850 to 2100, while the ensemble members No. 2 to 40 were restarted on January 1920 using the ensemble No. 1-generated initial condition with slight perturbations in air temperature (Kay et al., 2015). The horizontal resolution of the CESM-LE products is 0.9° × 1.25°. Those products have been used in various studies, such as investigating the impacts of internal climate variability on global air temperature variations (Dai et al., 2015) and meteorological drought in China (Wang and Zeng, 2017). In this study, the monthly SCF and LSAT from the CESM-LE for 1920-2005 are treated as the historical simulations, and the simulations from member No. 1 for 1850-1919 are regarded as the pre-industrial periods.

2.3 CESM 1.5 °C and 2.0 °C projects

The CESM 1.5 °C and 2.0 °C projects are specifically designed to achieve the goal of the 2015 Paris Agreement (Sanderson et al., 2017). The scenarios employ an emulator to simulate both the GMAT and emission concentration

evolution in Earth systems, and the parameters in the emulator were calibrated using the CESM simulations (Sanderson et al., 2017). Based on the methodology established in Sanderson et al. (2016), three idealized emission pathways, including 1.5 °C never-exceed (hereafter referred to as 1.5 °C), 1.5 °C overshoot (1.5 degOS), and 2.0 °C never-exceed (hereafter referred to as 2.0 °C), were defined to limit the GMAT to increasing within 1.5 °C and 2.0 °C by 2100 (Sanderson et al., 2017). In these pathways, before 2017, the carbon emissions follow RCP8.5; then, the combined fossil fuel and land carbon emissions rapidly decline to net-zero. Finally, the emission fluxes are reduced to negative to ensure that the GMAT achieves 1.5 °C and 2.0 °C warming targets by 2100. The difference between 1.5 degOS and 1.5 °C is that, after 2017, the declines in carbon and combined fossil fuel emissions show different rates. In 1.5 degOS, the rise in GMAT can reach over 1.5 °C before 2100, and emissions decline slightly less than those in 1.5 °C after 2017. For example, the emissions decrease to zero in 2046 for 1.5 degOS and in 2038 for 1.5° scenario. Details regarding the emulator establishment and scenario design are described in Sanderson et al. (2017).

To "provide comprehensive resources for studying climate change in the presence of internal climate variability", a set of multi-member projected simulations has been produced under three new scenarios, branching from the corresponding historical simulations of CESM-LE in 2006 (Kay et al., 2015; Sanderson et al., 2017). There are 11 simulations (visited in May 2017) available for both the 1.5 °C and 2.0 °C scenarios, and the products can be downloaded from the Earth system grid website (https://www.earthsystemgrid.org/dataset/ucar.cgd.ccsm4.lowwarming/). In this study, the monthly SCF and LSAT data from the above ensemble simulations under 1.5 °C and 2.0 °C scenarios are analyzed.

2.4 CMIP5 data

The monthly SCF and LSAT data from 12 models in CMIP5 for both the historical simulations (1850-2005) and future projections (2006-2100) under RCP2.6 and RCP4.5 are used in this study (Taylor et al., 2012). The selection of models is performed according to data availability and the spatial resolution of each product, and only the first ensemble run (i.e., r1i1p1) in each model is used. The models used in this study are BCC-CSM1.1, BNU-ESM, CanESM2, CCSM4, CNRM-CM5, FGOALS-g2, FIO-ESM, GISS-E2-H, MIROC-ESM, MPI-ESM-MR, MRI-CGCM3, and NorESM1-M. General

information about those models is summarized in Table S1. SCFs of the selected models were evaluated using satellite observations, and the results indicated that, overall, the model products were not only able to capture spatial patterns, seasonal change, and annual variations but also showed the apparent disparities between different models. The simulations from both the RCP2.6 and RCP4.5 scenarios are chosen because the surface warming rates in these two scenarios are, to some extent, comparable to the 2.0 °C warming target (IPCC, 2013; Jiang et al., 2016). To facilitate this comparison, the monthly SCFs from 12 models are rescaled to 0.9° × 1.25° to match the resolution of the CESM outputs.

2.5 Validation data

To validate the simulated SCF, daily snow cover products of the 0.05° MODIS Climate-Modeling Grid (CMG) version 6 are used (Hall and Riggs, 2016). It is well known that the satellite-based SCF has obvious biases when clouds are present. To reduce the impacts of cloud cover, the daily confidence index (CI), defined as the percentage of clear sky within a grid cell from the CMG product, is used to filter the CMG SCF products. Similar to the method used in Toure et al. (2016), we begin by filtering out the daily SCF data with CI values of less than 20%. Then, the daily filtered CMG SCF data are averaged per month, and, finally they are aggregated in line with the CSME-LE resolution (i.e., 0.9° × 1.25°).

Excluding the MODIS SCF product, the monthly snow area extent (SAE) time series from the National Oceanic and Atmospheric Administration Climate Data Record (NOAA-CDR) is also adopted to compare with the modeled products (Robinson et al., 2012). The NOAA-CDR SAE is computed from the gridded monthly snow cover databases, which are derived from the NOAA weekly snow charts for 1966-1999 (Robinson, 1993) and the Interactive Multi-Sensor (IMS) daily snow products for 1999 and afterwards (Ramsay, 1998;Helfrich et al., 2007). The NOAA CDR SAE monthly time series averaged over the NH are obtained from Rutgers University (http://climate.rutgers.edu/snowcover/).

3.   Methods

In this study, the pre-industrial period is taken as 1850-1919 in each modeled product, which is consistent with that used in Sanderson et al. (2017). The SAE for each modeled product is computed as the SCF multiplied by the land area of each grid cell. The monthly SAE and LSAT, averaged over the NH land area for 1920-2100, are then derived from all products.

The annual anomaly of each variable, with its corresponding 1850-1919 mean, denotes the change with respect to the pre-industrial period. The linear trend is derived from the least-square-fit method. The period of 1971-2000 is used as the common baseline period. To assess this change, the mean value of each product for 2071-2100 is compared with those in the baseline period. The standard deviation (STD) across CESM ensemble members, or CMIP5 multi-models, represents the spread of simulations due to ensemble variability. To address the contribution of change in SCF due to LSAT warming, both the pattern correlation coefficient and the coefficient of determination between the two for different seasons and different products are also computed. The linear regression method is used to analyze the dependence of SAE on the LSAT anomaly for different periods and different products in four seasons, annually. The four seasons represent the boreal winter (December-February, DJF), spring (March-May, MAM), summer (June-August, JJA), and autumn (September-November, SON).

4. Validation of modeled SCF

Figure 1 shows the mean (2001-2005) SCF from MODIS, the CESM-LE ensemble mean, and the CMIP5 ensemble mean. The SCF biases of two ensemble means with regarding to MODIS and STDs of their biases, are also plotted. Overall, the spatial patterns from these three products are similar, with the greatest SFCs over the high latitudes. The annual mean SCF values, averaged over the entire NH land area from 2001-2005, are 17.97% for MODIS, 22.3 ± 0.26% (STD) for CESM-LE, and 16.24 ± 7.87% (STD) for CMIP5. Compared with the MODIS mean, the CESM-LE ensemble mean overestimates the SCF over most of the land area, with the exception of a small portion in western Eurasia (Fig. 1d), while the CMIP5 ensemble mean is comparable to that of MODIS, with a slight underestimation over the Eurasian continent, North America, and Greenland (Fig. 1e). Toure et al. (2016) evaluated the MODIS SCF with offline CLM4.0 simulations using the observation-based atmospheric forcing data set, and they found that, overall, the model underestimated the mean SCF average. They attributed the modeled SCF biases to the snow process parameterization, the sub-grid effect in CLM4.0, and the forest-coverage- and cloud-cover-induced uncertainties in MODIS. Those issues still exist in the CESM-LE. In contrast to the underestimation by offline CLM4.0 simulations, the overestimation of SCF in CESM-LE is partially attributed to the biases in surface atmospheric forcing variables (e.g., precipitation, air temperature, humidity), which are

8

produced by the atmospheric model in CESM-LE (Wang and Zeng 2017). The biases due to rainfall and snowfall separation are also responsible for the above SCF biases in CESM-LE (Wang et al., 2016). For example, during 1979-2005, the CESM-LE ensemble mean, averaged over the NH land area, has an annual precipitation of 2.13 mm/day, while the Global Precipitation Climatology Project (GPCP) product has a smaller value of 2.08 mm/day (Huffman et al., 2009). The GPCP

5   product has been used to bias-correct precipitation in the atmospheric forcing dataset by both Toure et al. (2016) and Wang et al. (2016). The STDs of SCF biases from CESM-LE are generally less than 5%, with the greatest values located in the western United States, the mid-latitude portion of the Eurasian continent, and the Tibetan Plateau (Fig. 1f). In contrast, the STDs from CMIP5 are above 10% over the majority of snow-covered regions (Fig. 1g), which are much larger than the magnitude of their ensemble mean differences (Fig. 1e). In fact, the spread from CESM-LE is induced by the internal climate

10   variability due to the interaction of intrinsic dynamical processes within the Earth system, in which a slight perturbation of the initial condition in the CESM-LE experiment will lead to different climate variabilities (Kay et al., 2015). The STD from CMIP5 is derived from the inter-model variability, which is mainly caused by the model structure and physical parameterization, namely, the representation of the snow process in different models because all models used the same external forcing (Taylor et al., 2012). Therefore, these results indicate that the SCF heavily relies on the representation of the

15   physical process in the model, while the internal climate variability might play a relatively minor role.

   Figure 2 shows the 12-month moving averaged SAE anomalies over the NH from the NOAA-CDR, CMIP5, and CESM-LE ensemble mean during 1967-2005. The full spread of the 12 CMIP5 models and CESM-LE 40 ensemble members is also shown. The SAE from NOAA-CDR exhibits apparently annual variations, with the anomaly varying within $\pm 2\times10^6$ km^2, while SAEs from both the CMIP5 and CESM-LE ensemble means show less temporal variation. The spreads

20   from both products are remarkable, and their ranges cover most NOAA-CDR curves, implying that the SAEs from both modeled products are reasonable.

   To further investigate the SAE temporal variations, we compute the R values between the modeled products and NOAA-CDR, along with the linear trends of the three products for the period of 1976-2005 (Table 1). For CESM-LE, R varies between -0.41 and 0.55 with a mean of 0.18 ± 0.17 (STD), and 35 of the 40 members have a positive R. However, for

25   CMIP5, R varies from 0.10 to 0.50 with a mean of 0.24 ± 0.12 (STD). The linear trends of SAE from all three products

exhibit a reduction, with values of -3.98×10^4 km^2/year, -2.36 ± 0.76 (STD) $\times 10^4$ km^2/year, and $-2.62 \pm 1.33 \times 10^4$ km^2/ year for NOAA-CDR, CSEM-LE, and CMIP5, respectively. The trend ranges from -4.35×10^4 km^2/year to -0.22×10^4 km^2/year across CESM-LE ensemble members and from -5.22×10^4 km^2/year to -1.02×10^4 km^2/year for CMIP5 models. The ensemble means of both modeled products underestimate the magnitude of SAE reduction. However, accounting for the model spreads in the two products, both modeled SAE reductions are roughly comparable to that of NOAA-CDR. On the other hand, the fact that the majority of members have a positive R, along with the consistency in the reduction of SAE, implies that both CMIP5 and CESM-LE products can be used to represent the temporal evolution of SAE. It should be noted that the deficiencies of climate modeling for snow reproduction are beyond this work, and therefore, they are not discussed in this study.

5. Impacts of the LSAT on snow cover

5.1 Long-term SAE variations

To quantify the long-term SAE variations, Figure 3 shows the annual anomalies of both SAE and LSAT averages over the NH land area for the period of 1920-2100. All anomalies are with respect to the mean value for the pre-industrial period. Both the ensemble mean and the full spread of ensemble members are displayed. There are distinct temporal variations in the long-term evolution and the magnitude of diversity between different products from both SAE (Fig. 3a) and LSAT (Fig. 3b). During the period of 1920-2005, the ensemble SAE anomalies from both CESM-LE and CMIP5 show similar annual variations with a correlation coefficient of 0.86, but the actual values from CESM-LE are consistently larger than those from CMIP5. Before the early 1960s, the temporal variability of SAE was relatively small, but afterwards, it shows an apparent trend of decline. Overall, SAE reduction from the CMIP5 ensemble is much larger than that from the CESM-LE ensemble mean. For example, from 1920 to 2005, the annual SAE ensemble mean decreases by approximately 0.75×10^6 km^2 for CESM-LE, while this value is 1.32×10^6 km^2 for CMIP5. During the future period of 2005-2050, the linear trends of SAEs are all negative, varying between -4.92×10^4 km^2/year (2.0 °C) and -2.37×10^4 km^2/year (RCP 2.6), while after 2050, the trends from both RCP 2.6 (0.32×10^4 km^2/year) and 1.5 °C (0.26×10^4 km^2/year) become positive. Moreover, before 2050, the ensemble mean SAE anomaly from CMIP5 is less than those from CESM-based simulations, but after 2050, they are

comparable to each other from both RCP 2.6 and 1.5 °C. Nevertheless, although the ensemble mean SAE shows an overall trend of decline for the future period, the upper range of the spread from RCP 2.6 gives positive SAE anomalies within a few years, with the maximum value at approximately 1.0×10^6 km^2. This feature implies that the projected SAE under RCP 2.6 in some models would be above pre-industrial levels.

5   In contrast, the LSAT anomaly exhibits an overall trend of increase (Fig. 3b). The linear trends of LSAT from the ensemble mean for 2006-2050 are 0.022, 0.026, 0.034, and 0.043 °C/year for RCP 2.6, 1.5° C, RCP 4.5, and 2.0 °C, respectively. Similar to the SAE, after 2050, the LSAT trends become negative in both RCP 2.6 (-0.03 °C/decade) and 1.5 °C (-0.02 °C/decade), and the magnitudes of the trends from both RCP 4.5 and 2.0 °C also decrease compared with those in the early period. Furthermore, Fig. 3 also shows that the range of the ensemble members displays different variabilities for

10   different products.

To examine the evolution of the SAE anomaly spread between different ensemble members, we have computed the STDs of SAE across all members in each year, and the results are shown in Fig. 4. The STDs from both CESM and CMIP5 show apparent annual variations. For the entire period of 1920-2100, the annual STDs from CESM change slightly with time and vary between 0.3×10^6 km^2 and 0.7×10^6 km^2, while the STDs from CMIP5 increase distinctly with time at a magnitude

15   of up to 1.4×10^6 km^2. Correspondingly, the spread of LSAT was also computed (Fig. S1). The temporal evolutions of the annual STDs of LSAT from different products are similar to those of SAE. The STDs of the LSAT anomaly also represent the spread of the warming rates between different ensemble members. To further investigate the dependence of SAE change on an increase in LSAT, we have conducted a linear regression of the annual SAE anomaly on the LSAT anomaly from each CESM and CMIP5 ensemble member during both the historical and future periods. We then computed the ensemble mean of

20   regression coefficients and their STDs for each product. For the period of 2006-2100, the regression coefficients (unit: 10^6 km^2/°C) are -1.37 ± 0.56 (1.5°C), -1.12 ± 0.07 (2.0°C), -1.18 ± 0.19 (RCP2.6), and -0.97 ± 0.44 (RCP4.5), while for the period of 1920-2005, the magnitude of the regression coefficient decreases, with values of -0.79 ± 0.42 (CESM-LE) and -0.84 ± 0.22 (CMIP5). The results do not clearly show that the dependence of SAE loss on the warming rate in CMIP5 is greater than that from the simulations in CESM. However, based on both Fig. 4 and Fig. S1, we argue that the inter-model diversity

25   of CMIP5 is probably responsible for the increase in the spread of both SAE and LSAT with an increase in time. The above

11

results suggest that the uncertainty induced by internal climate variability is an inherent property in the climate system and is nearly stationary, while the uncertainty (or the inter-model spread) from CMIP5 multi-model simulations increases with warming signals. Therefore, caution should be taken when CMIP5 outputs from the multi-model ensemble are used to address the long-term change of surface variables.

5.2 Future SCF and LSAT changes in both 1.5 °C and 2.0 °C

Figure 5 shows the 30-year annual mean SCF differences between 2071-2100 and 1971-2000 from both the 1.5 °C and 2.0 °C scenarios. Both products show the reduction in SCF for 2071-2100, and the NH average SAE changes are -1.69 $\times 10^6$ km^2 in 1.5 °C and -2.36$\times 10^6$ km^2 in 2.0 °C. The annual mean ensemble changes of SFC are not uniformly distributed. The largest magnitude of change could be above 10%, which occurs at mountain ranges in the middle latitudes, such as the Iran Plateau, northern Canada, western America along the Rocky Mountains, and the western Tibetan Plateau. In comparison to the ensemble mean SCFs between 1.5 °C and 2.0 °C for 2071-2100 (Fig. 5c), the differences are generally below 4% across the majority of snow-covered areas (corresponding to the SAE difference of 0.67×10^6 km^2), with the largest difference appearing at the same locations as the largest SCF reduction with respect to the base period (Figs. 5a and 5b). In contrast, the ensemble mean for LSAT exhibits the largest warming over a polar region during the future period, and the magnitude of warming lessens over the middle and low latitudes (Figs. 5d and 5e). The increase in LSAT for 2071-2100 exceeds 4 °C along the coastline of the Arctic Ocean. Prominent warming over polar regions represents the polar amplification effect, which might be related to sea ice change (Screen and Simmonds, 2010). The inconsistent spatial variations of LSAT and SCF suggest that LSAT is not the only factor in determining the SCF change.

To further examine SAE change in the future, we compute the percentage change in SAE between 2071-2100 and 1971-2000 from 1.5 °C, 2.0 °C, RCP 2.6, and RCP 4.5 scenarios (Table 2). The percentage change is calculated as the mean difference of two periods divided by the mean of 1971-2000, annually and in each season. The STDs are computed from 12 models (RCP 2.6 and RCP 4.5) and 11 CESM simulations (1.5 °C and 2.0 °C). Figure 6 illustrates the Signal-Noise Ratio (referred to as SNR), which is defined as the ratio of the ensemble mean change to the STDs of change across the ensemble members. This metric represents the relative importance of external forcing and internal climate variability on the variable

change (Kay et al., 2015; Wang and Zeng, 2017). Under the 1.5 °C scenario, the SNR of SAE change exceeds 1 across 65% of all snow-covered areas (with respect to the base period), while under the 2.0 °C scenario, it exceeds 1 over 70% of all snow-covered areas in the NH. For the difference in 1.5 °C and 2.0 °C during 2071-2100, the percentage of snow-covered areas with SNR exceeding 1 is approximately 31% (Fig. 6c). For LSAT, the SNR over almost the entire NH land area exceeds 1 under both scenarios. This feature implies that both LSAT and SCF changes are dominantly affected by external (or anthropogenic) forcing and are slightly triggered by internal climate variability. The spatial patterns of SNR for both SCF and LSAT are broadly consistent with each other over snow-covered regions. The SNRs of both SCF and LSAT are relatively small over Eurasian middle-to-high latitudes, compared to other regions, but are they large in the region east of 90 °W in North America as well as along the margin of the Rocky Mountains. Over snow-free regions in low latitudes, the greater magnitude of SNR of LSAT is caused by the smaller STD when compared to high-latitude regions. Moreover, the SNR of LSAT is, overall, larger than that of SCF for a specific scenario. This further reflects that external forcing has a more evident impact on the change in LSAT than that of SCF.

5.3 Contribution of LSAT on snow cover reduction

In the context of climate change, the increase in surface air temperature in recent decades is one of the most prominent features. In CESM, a surface air temperature with a 0 °C-threshold is used to separate rainfall and snowfall. Therefore, an increase in surface air temperature would reduce the chance of snowfall but enhance rainfall. An outstanding question is: to what degree is the increase in local LSAT responsible for snow cover reduction by 2100?

To address the above question, we compute the pattern-correlation (R) values between SCF and LSAT changes for 2071-2100 versus 1971-2100 over the NH from 1.5 °C, 2 °C, RCP 2.6, and RCP 4.5 scenarios (Fig. 7). As previously discussed in section 5.1, the analyses have also shown that, in the long-term, the regression coefficient of SAE anomalies onto LSAT change are all negative in both historical and future periods, and the ensemble mean magnitudes of those coefficients from four scenarios during 2006-2100 are comparable. Therefore, the increase in LSAT will reduce the local snow fraction, and R should be negative. For all four seasons and the annual R, the ensemble mean R is smaller than -0.3, with all passing the significance test at the 95% confidence level. The magnitude of R shows clear seasonal variations, with

the highest occurring in boreal autumn and the lowest occurring in boreal summer. For example, R varies from -0.70 (RCP 2.6, RCP 4.5) to -0.75 (1.5 °C and 2.0 °C) in OSN and from -0.30 (RCP 4.5) to -0.40 (1.5 °C) in JJA. Furthermore, it clearly shows that the ensemble variabilities (denoted as STDs of R in Fig. 7) from CESM-based products are relatively small when compared to the ensemble mean R, as well compared to those from CMIP5. This illustrates that the inter-model differences have greatly influenced the above relationship. To quantify the contribution of LSAT warming on snow reduction, we use an index, the coefficient of determination (CD), which is defined as the percentage of squared pattern correlation (CD = $100\% \bullet R^2$) annually and for four seasons across different scenarios (figure not shown). The CD denotes the percentage of SCF reduction explained by the LSAT increase. Similar to R, the CD shows clear seasonal variations, with the highest in OSN (49% - 55%) and the lowest in JJA (10% -16%), and the STDs of CD are also larger in CMIP5 than those in CESM-based simulations. Although the CDs from the CMIP5 ensemble mean are slightly smaller than those from CESM-based simulations, overall, the two from the specific seasons are comparable. For example, the CDs of the ensemble annual mean difference are approximately 50% for all products. This means that the LSAT change could explain approximately half of annual SCF reduction, and the change in SCF would also be affected by other factors. For example, studies have shown that Arctic sea ice has greatly impacted snow cover over the NH (Kapnick and Hall, 2012), and human activities induce black carbon, thus reducing the snow surface albedo and enhancing the solar radiation absorbed by the snow and, as a result, accelerating the reduction in snow (Flanner et al., 2007).

6. Conclusions

This study investigates the long-term change in SCF and SAE associated with LSAT over the NH during the period of 1920-2100. We have analyzed simulations from the CESM-LE, CESM 1.5 °C and 2.0 °C projects (Sanderson et al., 2017), as well as simulations from historical, RCP 2.6 and RCP 4.5 from CMIP5 12 models (Taylor et al., 2012). The model-simulated, snow cover products are compared with MODIS and NOAA-CDR observations. We emphasize the responses of both SCF and SAE under different warming levels. The reduction of snow cover due to an increase in LSAT is quantified. The relative importance of internal climate variability and external forcing on changes to both SCF and LSAT, and their relationship, are also addressed.

We find that the ensemble annual mean SCF from both CMIP5 and CESM-LE simulations can broadly capture the MODIS spatial pattern, with a slight underestimation in CMIP5 and an overestimation in CESM-LE. The annual SAE from the ensemble means of CMIP5, CESM-LE, and NOAA-CDR all display significant reduction trends for the period of 1967-2005. Compared to pre-industrial periods, the SAE anomalies from the CMIP5 and CESM simulations show gross

5 similarities in their annual variations. Overall, the annual ensemble mean SAE displays a downward trend, but the LSAT exhibits an upward trend for the long-term period of 1920-2100. However, the actual variability differs in different products for different time periods. The trends of the projected SAEs (LSAT) from all products are negative (positive) for the period of 2006-2100, but they become positive (negative) during the second half of the 21st century in both RCP 2.6 and 1.5 °C. The magnitude of the SAE anomaly in RCP 2.6 is comparable to that in 1.5 °C, and it is smaller than that in 2.0 °C. Furthermore,

10 the STDs of SAEs induced by the ensemble member variability show time invariance across CESM ensemble members but increase with warming signals among CMIP5 models. Therefore, caution should be taken when multi-model, projected surface variables are analyzed.

For 30-year mean change between 2071-2100 and 1971-2000, the ensemble mean magnitude change of SAE varies from -14.47% (RCP 4.5) to -8.02% (1.5 °C) across the four scenarios averaged over NH land areas. For the seasonal time

15 scale in a specific scenario, the percentage magnitude of SAE loss is largest in JJA and smallest in DJF. We also find that the spread (STDs) of SAE loss due to ensemble variability is much larger in the two RCPs than those in both 1.5°C and 2.0°C, implying that the inter-model variability will induce larger SAE uncertainty than the internal climate variability. In comparison with the ensemble mean SCF between 1.5 °C and 2 °C for 2071-2100, the SCF differences are less than 4% over most snow grid cells, and the SAE difference is 0.67×10^6 km^2. Moreover, by analyzing the SNR of SAE change, we find

20 that SNRs exceed 1 over a majority of the land area in both 1.5 °C and 2.0 °C, and the percentage of area with SNR exceeding 1 in 2.0 °C is slightly more than that in 1.5 °C. The spatial patterns of SNR for both SCF and LSAT are broadly consistent with each other across snow-covered regions, but the SNR magnitude for SCF is much smaller than that for LSAT. The significant negative R values between the projected LSAT and SCF changes for 2071-2100, versus the baseline period of 1971-2000, suggests that the SCF reduction strongly relies on LSAT warming. For 2006-2100, the regression coefficients of

25 SAE anomalies on the LSAT anomaly are -1.37 ± 0.56×10^6 km^2/°C (1.5 °C), -1.12 ± 0.07×10^6 km^2/°C (2.0 °C), -1.18 ±

0.19×10^6 km^2/°C (RCP 2.6), and $-0.97 \pm 0.44 \times 10^6$ km^2/°C (RCP 4.5). We also find that more than 50% of the OSN and the

annual reduction in projected SCF over the NH is attributed to the increase in LSAT, whereas this value is less than 16% in

JJA. Furthermore, the STDs of CDs are much larger from CMIP5 than in CESM-based simulations. This feature implies that

the SAE uncertainties are mainly derived from the physical structure and the snow process representation in different CMIP5

5   models, while they are less affected by the internal climate variability from the CESM ensemble. From the above results, we

may conclude that external forcing plays the predominant role in future changes in both SFC and LSAT, and with an increase

in warming signals, the effects of external forcing on surface variables will be more evident.

      Finally, we provide a comprehensive analysis of both SCF and SAE from the CESM and CMIP5 simulations for both

historical and future periods in different warming or emissions scenarios. Under different scenarios (e.g., RCP 2.6, RCP 4.5,

10   2 °C, and 1.5 °C warming above pre-industrial levels), the snow cover response to LSAT warming varies with season and

differs in products. In conclusion, surface warming is partially responsible for the surface snow change. Furthermore, it

should be noted that there are some caveats in this study. In the analyses, we only use model-simulated SFC and LSAT to

investigate changes to the two. In the model, SCF largely depends on the parameterization schemes. Many studies have

focused on the validation of the modeled SFC using satellite or in situ observations (e.g., Xia and Wang 2015), and others

15   have tried to improve snow schemes in the model (e.g., Wang and Zeng, 2009). However, it is still difficult to conclude

which model has an overall better snow scheme. Therefore, we suggest examining the relationship between SFC and LSAT

(or other surface meteorology quantities) based on observations and then using this relationship to evaluate the model

simulations. To do this, we can first choose the models with better representation of the relationship and then, based on the

selected model, investigate future changes.

20

Tables

Table 1: Correlation coefficient (R) of snow area extent (SAE) between CESM-LE and NOAA-CDR, between CMIP5 and NOAA-CDR, and annual linear trend of SAE from above three products for the period of 1976-2005. The mean, standard deviation, maximum, and minimum of the corresponding statistics from CESM-LE multi-member and CMIP5 multi-model are also displayed. The value in the last column is the annual linear trend of SAE from NOAA-CDR. The values with superscript stars denote the R values or trends passing the 95% significance level test.

	R (CESM-LE, NOAA-CDR)	R (CMIP5, NOAA-CDR)	Trend CESM-LE 10^4km^2/year	Trend CMIP5 10^4km^2/year	Trend NOAA-CDR 10^4km^2/year
Mean	0.18	0.24	-2.36*	-2.62*	-3.98*
Standard deviation	0.17	0.12	0.76	1.33	---
Maximum	0.55*	0.50*	-0.22	-1.02	---
Minimum	-0.41*	0.1	-4.35*	-5.22*	---

Table 2: Percentage change of snow area extent (SAE) and its standard deviation (STD) between the period of 1971-2000 and 2071-2100 from the 1.5°C, 2.0°C, RCP2.6, and RCP4.5 scenarios. The percentage changes are computed as the difference of two periods divided by the mean of 1971-2000 in each season and annually. The STD is computed from 12 models (RCP2.6 and RCP4.5) and 11 CESM simulations (1.5°C and 2.0°C).

	1.5°C	2.0°C	RCP2.6	RCP4.5
Annual	-8.02±0.78%	-10.92±0.52%	-8.5±5.58%	-14.47±5.71%
DJF	-5.41±0.99%	-7.41±0.61%	-5.62±3.2%	-9.70±3.48%
MAM	-6.74±0.78%	-9.59±0.73%	-9.03±7.1%	-15.77±7.39%
JJA	-19.42±1.19%	-26.38±1.36%	-16.56±13.81%	-25.08±15.53%
OSN	-13.33±1.20%	-17.39±0.79%	-12.85±9.05%	-21.75±8.62%
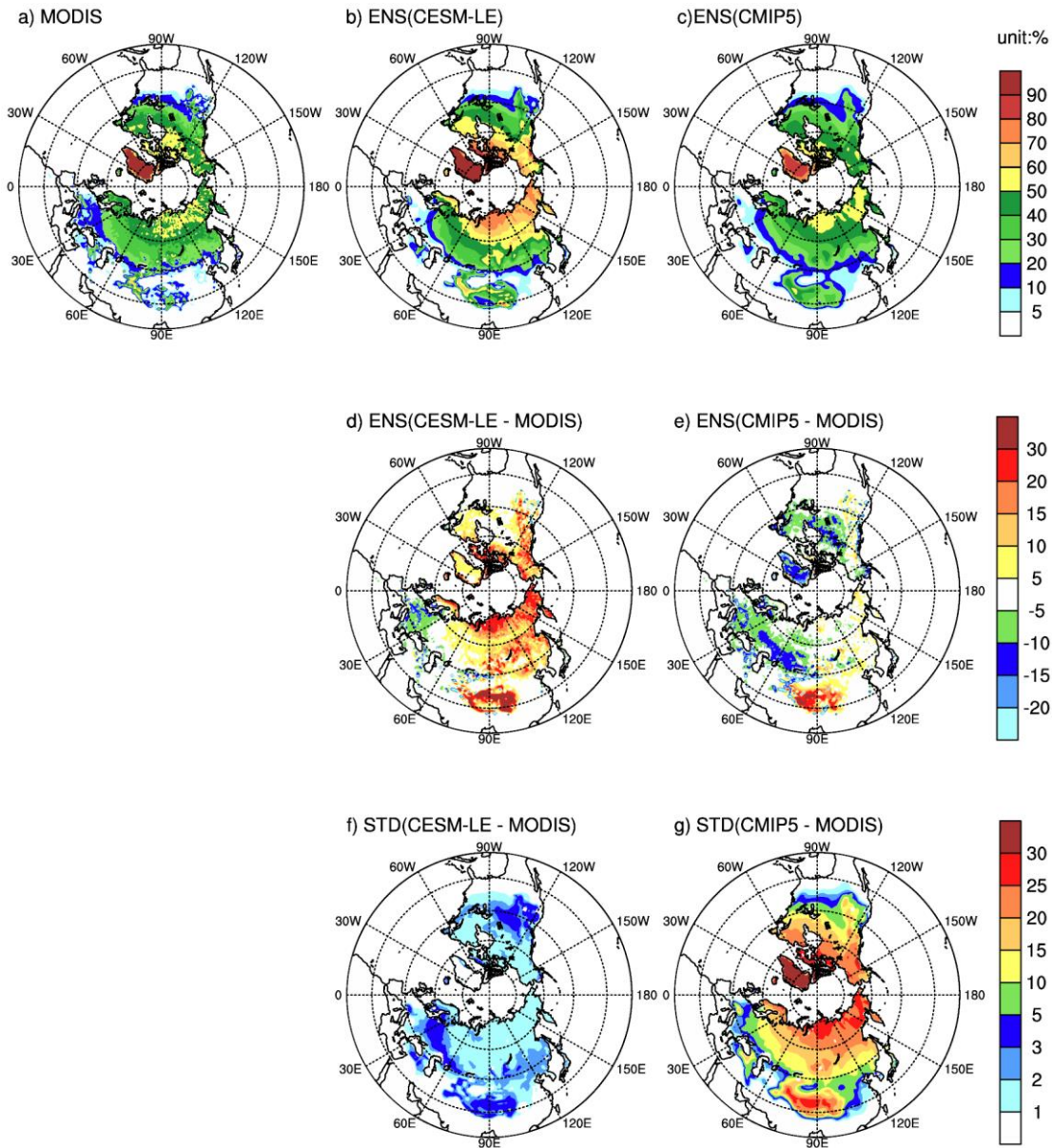
Figures:

Figure 1 Averaged annual snow cover fraction over Northern Hemisphere land area from a) MODIS, b) CESM-LE ensemble, c) CMIP5 ensemble; the difference between d) CESM-LE ensemble and MODIS, e) CMIP5 ensemble and MODIS; f) and g) are the standard deviations of c) and d), respectively. The average was taken for the period of 2001-2005.
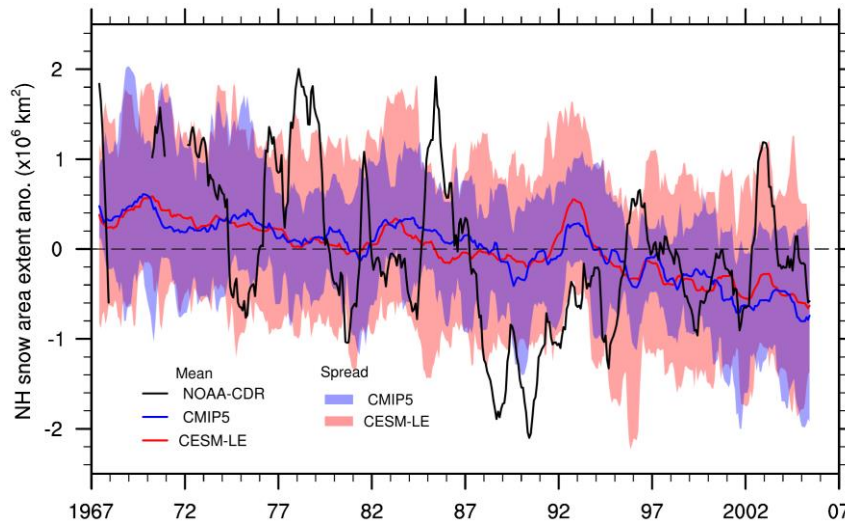
Figure 2 Time series of snow area extent (SAE) anomalies from NOAA-CDR, CMIP5 12 models, and CESM-LE 40 ensemble members over Northern Hemisphere land area for the period of 1967-2005. The spreads of CMIP5 12 models and from CESM-LE 40 ensemble members are shaded.
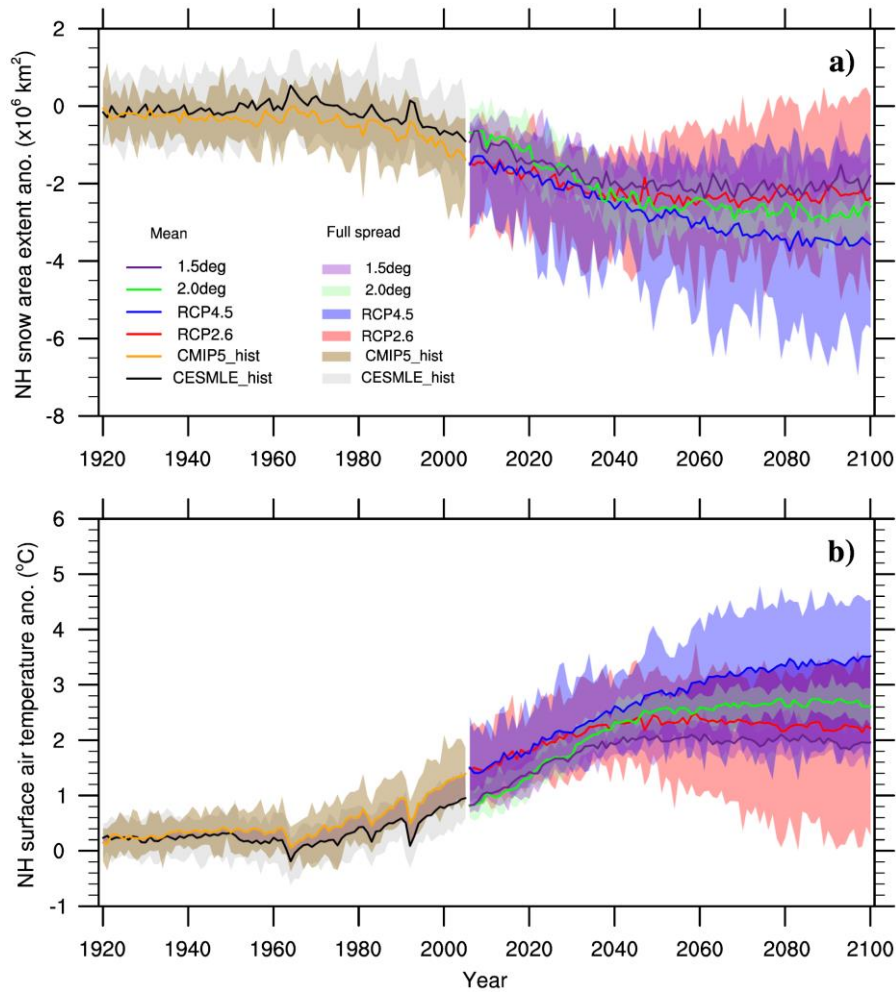
Figure 3 Annual time series of a) snow area extent (SAE) anomalies, and b) land surface air temperature (LSAT) anomalies over Northern Hemisphere for 1920-2100. The different colors represent the simulations from different projects with different scenarios. The shaded areas represent the full spread from simulations in both CMIP5 models and CESM ensemble members. Note that "1.5 deg" and "2.0 deg" represent simulations from 1.5°C and 2.0°C scenarios, respectively.
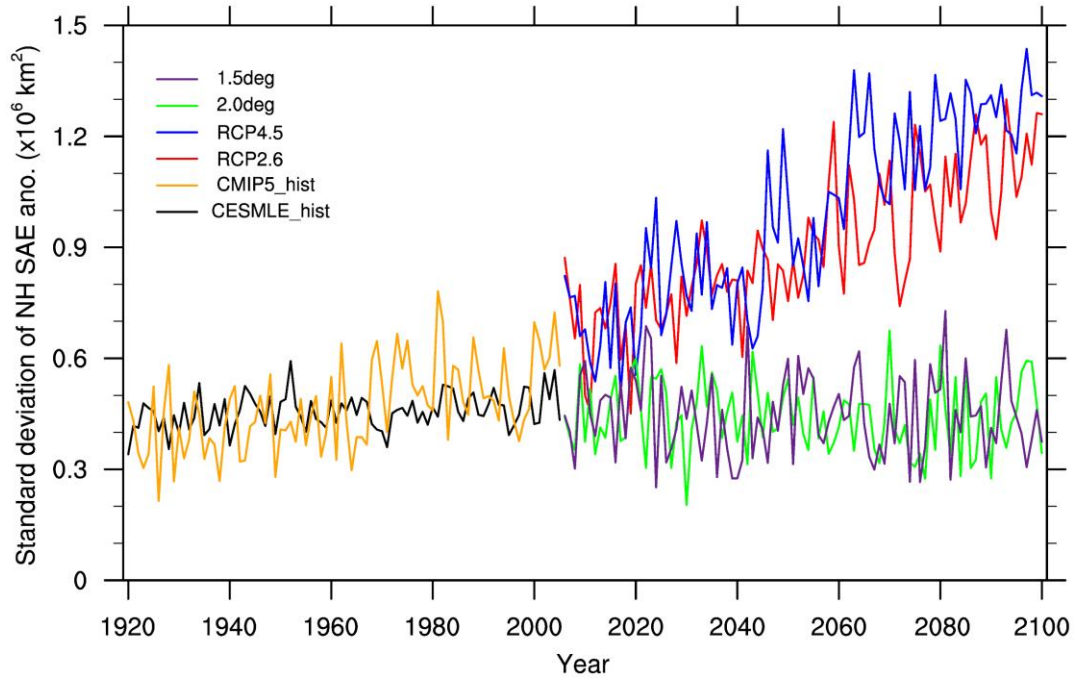
Figure 4 The annual standard deviation of snow area extent anomaly due to the ensemble variability for 1920-2100. Results from CESM-LE, CMIP5 historical, RCP2.6, RCP4.5, 1.5°C and 2.0°C scenarios are shown.
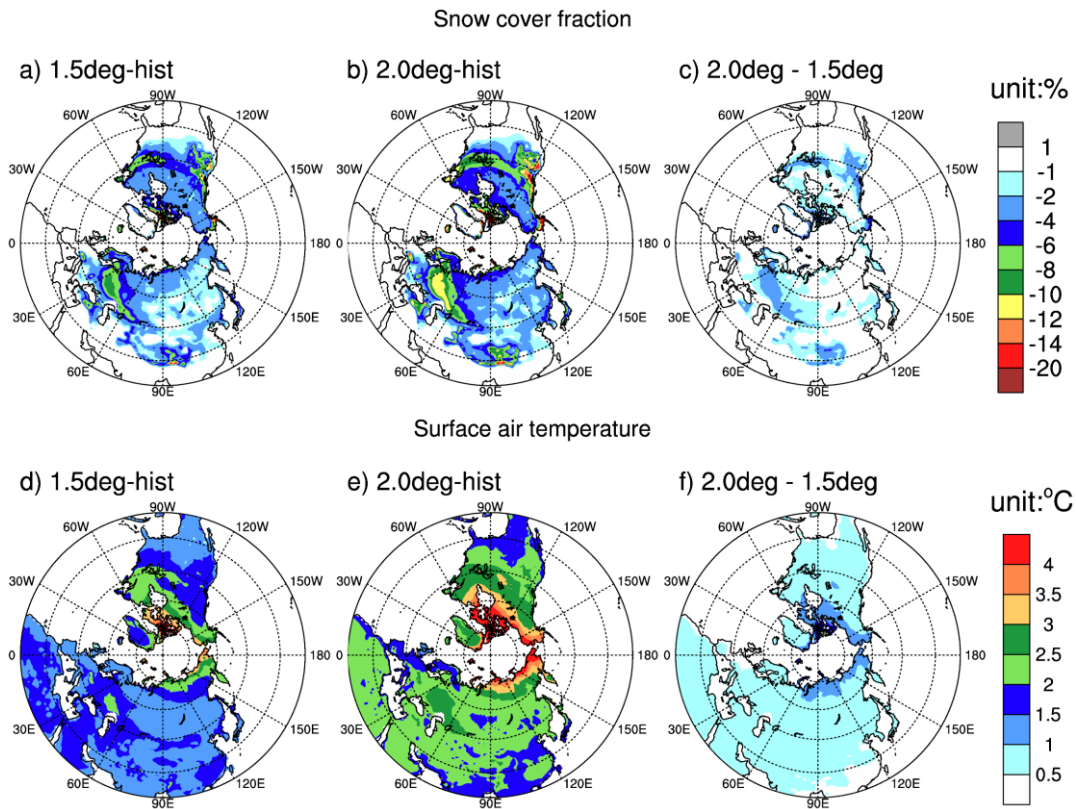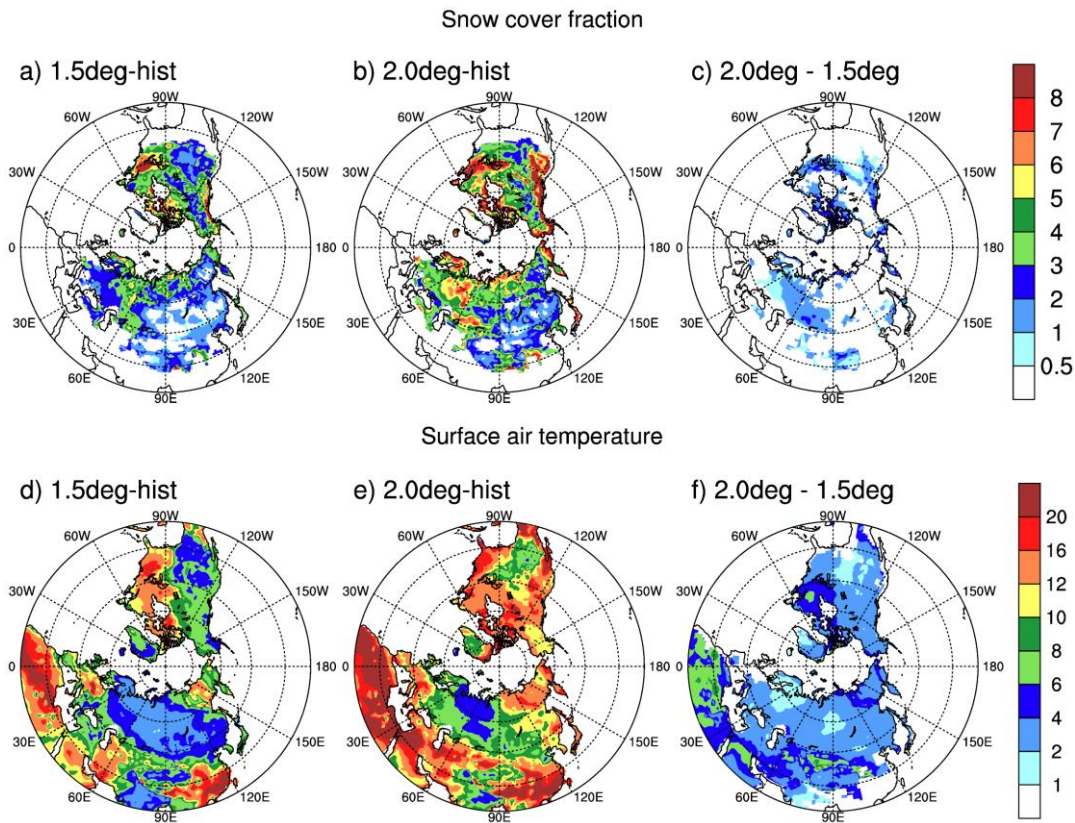
Figure 5 Snow cover fraction (top) and land surface air temperature (bottom) differences between 2071-2100 and 1971-2000 over Northern Hemisphere land area from a) 1.5 deg, b) 2.0 deg, and c) 2.0 deg minus 1.5 deg. d)-f) are the land surface air temperature differences corresponding to a)-c), respectively. "hist" is the ensemble mean for 1971-2000 from CESM-LE.

Figure 6 Similar to Figure 5, but for the signal-to-ratio (SNR) of snow cover fraction (a-c) and land surface air temperature differences (d-f) between 2071-2100 and 1971-2000 over Northern Hemisphere land area. The SNR was computed as the ratio of change in the ensemble mean to the standard deviation due to ensemble variability
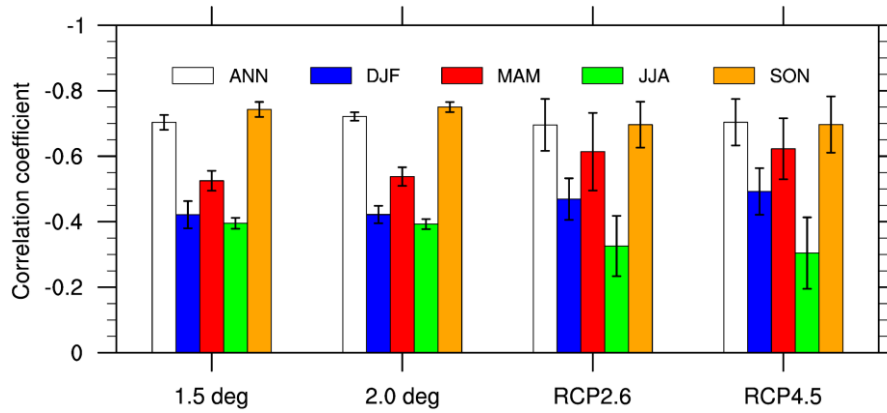
Figure 7 Pattern correlation between surface air temperature change and snow cover fraction change from 1.5°C, 2.0°C, RCP2.6, and RCP4.5 scenarios. The changes are computed as the difference between 2071-2100 and 1971-2000. The bar represents the ensemble mean, and the vertical line is the standard deviation from 12 models (RCP2.6 and RCP4.5) or 11 CESM simulations (1.5°C and 2.0°C).