Evaluation of Northern Hemisphere snow water equivalent in CMIP6 models ~~with satellite-based SnowCCI data~~ during 1982-2014

Kerttu Kouki[1], Petri Räisänen[1], Kari Luojus[1], Anna Luomaranta[1], Aku Riihelä[1]

[1]Finnish Meteorological Institute, Helsinki, P.O. Box 503, 00101, Finland

5   *Correspondence to*: Kerttu Kouki (kerttu.kouki@fmi.fi)

Abstract. Seasonal snow cover of the Northern Hemisphere (NH) is a major factor in the global climate system, which makes snow cover an important variable in climate models. ~~Monitoring snow water equivalent (SWE) at continental scale is only possible from satellites, yet~~ Previously, substantial uncertainties have been reported in NH snow water equivalent (SWE) estimates. A recent bias-correction method significantly reduces the uncertainty of NH SWE estimation, which enables a more

10   reliable analysis of the climate models' ability to describe the snow cover. We have intercompared ~~the~~ NH SWE estimates between CMIP6 (Coupled Model Intercomparison Project Phase 6) models and ~~satellite~~observation-based ~~NH~~ SWE reference data ~~estimates~~ north of 40° N for the period 1982-2014 and analyzed with a regression approach whether model biases in temperature (T) and precipitation (P) could explain the ~~differences~~ model biases in SWE. We analyzed separately SWE in winter and SWE change rate in spring. For SWE reference data, we used bias-corrected SnowCCI data for non-mountainous

15   regions and the mean of Brown, MERRA-2 and Crocus v7 data for the mountainous regions. The SnowCCI SWE data are based on satellite passive microwave radiometer data and in situ snow depth data. The analysis shows that CMIP6 models tend to overestimate SWE, however, large variability exists between models. In winter, P is the dominant factor causing SWE discrepancies especially in the northern and coastal regions. ~~This is in line with the expectation that even too cold temperatures cannot cause too high SWE without precipitation.~~ T contributes to SWE biases mainly in regions, where T is close to 0°C in

20   winter. In spring, the importance of T in explaining the snowmelt rate discrepancies increases. This is to be expected, because the increase in T is the main factor that causes snow to melt as spring progresses. Furthermore, it is obvious from the results that biases in T or P can~~not~~ explain all model biases either in SWE in winter or in the snowmelt rate in spring. Other factors, such as deficiencies in model parameterizations~~-~~ and possibly biases in the observational datasets, also contribute to SWE discrepancies. In particular, linear regression suggests that when the biases in T and P are eliminated, the models generally

25   overestimate the snowmelt rate in spring.

1 Introduction

Seasonal snow cover of the Northern Hemisphere (NH) is an important factor of the global climate system. The seasonal snow cover greatly influences surface albedo and, thus, the Earth's energy balance (Callaghan et al., 2011; Flanner et al., 2011; Qu and Hall, 2005; Trenberth and Fasullo, 2009). This makes snow cover an important variable in climate models (Derksen and

30 Brown, 2012; Loth et al., 1993). Additionally, snow cover significantly affects the hydrological cycle at high latitudes and in mountainous regions (Barnett et al., 2005; Bormann et al., 2018; Callaghan et al., 2011; Douville et al., 2002). In winter, snow cover stores large amounts of fresh water, which limits water availability. In spring and summer, warming temperatures melt the snowpack, releasing water as runoff. In some areas, snow is the largest freshwater storage, and about one sixth of the world's population is dependent on meltwater from snow (Barnett et al., 2005; Hall et al., 2008).

Melting snow is also a major source for hydropower (Callaghan et al., 2011; Magnusson et al., 2020). Due to the global warming, the melting season begins earlier, with the timing of streamflow peaks also becoming earlier (Kundzewicz et al., 2008). In addition, changes in snow cover affect the intensity of spring streamflow, as an increasing proportion of winter precipitation is rain instead of snow (Callaghan et al., 2011; Cohen et al., 2015; Dong et al., 2020; Kundzewicz et al., 2008).

40 Thus, changes in snow cover affecting the hydrological cycle can cause regional water shortages and affect hydropower production.

SWE (snow water equivalent) is the amount of water contained in the snowpack (in units of kg m^{-2}), or equivalently, the height of the water layer (in units of mm) that would result from melting the whole snowpack instantaneously (Fierz et al., 2009).

45 Recent studies show negative trends in global SWE (Bormann et al., 2018; Derksen and Brown, 2012; Essery et al., 2020; Hernández-Henríquez et al., 2015; Mortimer et al., 2020; Mudryk et al., 2017), but significant spatial variability exists: North America shows clear negative trends in observed SWE, while negative trends are less pronounced in Eurasia (Kunkel et al., 2016; Pulliainen et al., 2020). At mid-latitudes, SWE is more sensitive to warming than at high latitudes (Brown and Mote, 2009). Although the overall SWE trends are negative, there are also regions where SWE is observed and projected to increase:

50 SWE will most likely increase in northern Siberia and northern Canada, which largely results from the increased atmospheric moisture holding capacity (Brown and Mote, 2009; Park et al., 2012; Räisänen, 2008). Trends in ~~seasonal~~ snow cover also vary seasonally: the seasonal snow in spring is especially sensitive to warming due to the strong surface albedo feedback and the observed snow cover trends in spring are clearly negative in both Eurasia and North America (Derksen and Brown, 2012; Essery et al., 2020; Hernández-Henríquez et al., 2015). In winter, the observed trends are less pronounced: early winter from

55 October to December shows even slightly positive trends in both Eurasia and North America, while ~~the mid-winter shows~~ in January and February, there are no significant trends (Hernández-Henríquez et al., 2015).

~~Monitoring SWE at continental scale is only possible from satellites, yet substantial uncertainties have been reported in satellite-based NH SWE estimates (Bormann et al., 2018; Mudryk et al., 2015). However, our knowledge of the NH SWE has~~

60 ~~recently improved considerably, with new bias corrections which reduce the uncertainty of the SWE estimate integrated over NH from 33% to 7.4% (Pulliainen et al., 2020). With more accurate and reliable satellite-based SWE estimates, a comparison with the modeled SWE will also provide a more reliable analysis of the models' ability to describe the distribution of seasonal snow.~~

65 ~~Due to the reasons described above, it is crucial that seasonal snow is accurately described in climate models, to properly predict the cryospheric state in future climate. However, previous studies have shown that climate models have had difficulties in correctly reproducing the seasonal snow and its recent trends (Brutel-Vuilmet et al., 2013; Derksen and Brown, 2012; Henderson et al., 2018; Santolaria-Otín and Zolina, 2020; Thackeray et al., 2016). Therefore, it is important to study how the new CMIP6 (Coupled Model Intercomparison Project Phase 6) climate models can describe the seasonal snow, and where the~~
70 ~~uncertainties and discrepancies arise.~~

Observing SWE at continental scale is only possible from satellites, but also model and reanalysis products provide gridded SWE estimates which have been widely used in hydrological and climate research (e.g., Huning and AghaKouchak, 2020; Mortimer et al., 2020; Mudryk et al., 2020). Previously, substantial uncertainties have been reported in the NH SWE estimates
75 (Bormann et al., 2018; Mudryk et al., 2015). However, our knowledge of the NH SWE has recently improved considerably with new bias corrections which reduce the uncertainty of the SWE estimate integrated over NH from 33% to 7.4% (Pulliainen et al., 2020). The bias-correction method, for example, considerably improves SWE estimates in the moderate and deep SWE range (Pulliainen et al., 2020), which has previously caused underestimation in SWE estimates (Cho et al., 2020). However, limitations still exist: the bias-correction method cannot be applied in mountainous regions due to the lack of snow course
80 measurements and the large SWE variability in complex terrain (Pulliainen et al., 2020). Even though the area of mountainous regions is limited, these regions store a considerable portion of seasonal snow (Kim et al., 2021). For observation-based datasets, the bias-correction method mostly increases SWE, and it is therefore likely that without bias-correction or alternative approaches where estimates are corrected using in situ data, SWE in mountainous areas is underestimated (Pulliainen et al., 2020; Wrzesien et al., 2018).

85

Previous studies have shown that climate models have had difficulties in correctly reproducing the seasonal snow and its recent trends (Brutel-Vuilmet et al., 2013; Derksen and Brown, 2012; Henderson et al., 2018; Santolaria-Otín and Zolina, 2020; Thackeray et al., 2016). However, there are clear improvements from CMIP5 to CMIP6. The snow cover extent is better described in CMIP6 than in CMIP5 models (Mudryk et al., 2020; Zhu et al., 2021). The snow cover fraction is underestimated
90 in both CMIP5 and CMIP6 multi-model ensemble means, but the biases are clearly smaller in CMIP6 than in CMIP5 (Zhu et al., 2021). SWE, in turn, is biased high in both CMIP5 and CMIP6, and the peak SWE is also overestimated in almost all models (Mudryk et al., 2020). The model biases also vary seasonally; the snow cover fraction is best described in the period from January to March in CMIP6 models (Zhu et al., 2021), whereas the SWE bias is smallest in autumn and early winter and increases in spring (Mudryk et al., 2020). Both CMIP5 and CMIP6 models are mostly able to reproduce the observed negative
95 trend in snow cover area, but there are uncertainties in the magnitude of the trend (Mudryk et al., 2020; Zhu et al., 2021). The CMIP5 models are able to better capture the trend in snow cover extent than the trend in SWE (Santolaria-Otín and Zolina, 2020).

Due to the reasons described above, it is crucial that seasonal snow is accurately described in climate models, to properly predict the state of the cryosphere in future climate. Therefore, it is important to study how the new CMIP6 climate models can describe the seasonal snow, and where the uncertainties and discrepancies arise. The current paper focuses on the climatological distribution of SWE in CMIP6 models. To our knowledge, only one previous study has ~~compared~~ evaluated NH SWE in CMIP6 models ~~with satellite-based data~~: Mudryk et al. (2020) compared SWE estimates between CMIP6 models and several observational datasets. ~~They found that the models tend to overestimate SWE. The difference increases in spring and is smallest in autumn and early winter.~~ Additionally, they studied the connection between SWE and temperature, but did not consider temperature and precipitation together. However, they stated that a coordinated analysis of temperature and precipitation is needed to determine SWE trend drivers. In general, simulated trends of SWE can be considered more reliable if the current climatological distribution of SWE is simulated accurately. Therefore, in the present study, we consider, for the first time, the role of both temperature and precipitation for SWE ~~differences between~~ biases in CMIP6 climate models ~~and a satellite-based dataset~~. Specifically, the main goals of this study are (1) to intercompare the CMIP6 and ~~satellite~~observation-based SWE estimates and (2) to analyze whether temperature and precipitation biases could explain the ~~differences in~~ SWE biases.

2 Data ~~and Methods~~

2.1 CMIP6 model data

The data of this study consist of CMIP6 climate model data (Table 1) and observational and reanalysis data (Table 2). For CMIP6, we used monthly mean data from those models ~~that had horizontal resolution equal to or higher than 100 km and~~ for which either historical or esm-hist simulations were available for download in ~~April 2020~~ August 2021. ~~A total of ten models fulfilled these criteria.~~ The historical and esm-hist simulations extend from 1850 to 2014. In historical simulations, the CO_2 concentrations are prescribed, whereas in esm-hist simulations, the models calculate the atmospheric CO_2 concentration interactively based on prescribed CO_2 fluxes (Eyring et al., 2016). ~~We only used relatively high-resolution models (50 or 100 km), as coarser resolutions would differ too much from the resolution of the observational datasets, making the comparison more problematic.~~ In this study, we used three variables from CMIP6 models: SWE (variable "snw", unit kg m^{-2}), surface air temperature ("tas", unit Kelvin), and precipitation ("pr", unit kg m^{-2} s^{-1}). Altogether, there are 17 high-resolution models (100 km) and 21 low-resolution models (250-500 km) in Table 1. We have evaluated the SWE sum over the entire study area for every model but performed the subsequent detailed analysis only for the high-resolution (100 km) models. We decided to leave the models with coarser resolution out of the analysis, as coarser resolution would differ too much from the resolution of the observational datasets, making the comparison more problematic. The number of ensemble members available for the chosen models varies between 1 and 16. For simplicity, we only consider the first member of each model ensemble (r1i1p1f1) in this study. A brief analysis showed that the differences between different ensemble members for the same model were generally

130   smaller compared to inter-model differences. Figure S1 in the Supplementary material shows all realizations of three different models (CESM2, MPI-ESM1-2-HR, and EC-Earth3), which were chosen in this figure due to a high number of realizations. The figure illustrates that internal variability of each model is smaller than the intermodel variability.

2.2 Observational and reanalysis data

135   For SWE reference data, we have used ESA CCI-Snow "SnowCCI" (European Space Agency Climate Change Initiative, Snow) data for the non-mountainous regions, and MERRA-2 (The Modern-Era Retrospective analysis for Research and Applications, Version 2; Gelaro et al., 2017; GMAO, 2015a), Brown (Brown et al., 2003) and Crocus v7 (Brun et al., 2013) data for the mountainous regions. ESA CCI-Snow "SnowCCI" (European Space Agency Climate Change Initiative, Snow) SnowCCI data, which is the same product as the GlobSnow v3 SWE product (except provided in geographical

140   latitude/longitude grid for easier comparison with climate model data)data, are based on satellite passive microwave radiometer data and in situ data (Luojus et al., 2021; Pulliainen et al., 2020). The SnowCCI algorithm combines microwave brightness temperature (Tb) data, observed by satellite instruments, with ground-based snow depth measurements from the global network of synoptic weather stations (Luojus et al., 2021). The SWE estimation algorithm is based on the difference in Tb between two frequencies (37 and 19 GHz). The ground beneath the snowpack emits microwaves, which propagate through the snowpack,

145   being partially absorbed during the process. The low-frequency and high-frequency signals attenuate differently as they propagate through the snowpack, which makes the difference in Tb a good indicator for estimating SWE (Cagnati et al., 2004). The attenuation is affected by snow depth, snow grain size, and snow density. The high-frequency signal attenuates more asthan the low-frequency signal when it propagates through a deep, dense, large-grainedthe snowpack, especially for a deep, dense, and large-grained snowpack compared to the low-frequency signal. Thus, a large difference between high and low

150   frequency signals indicates a high SWE (Kelly et al., 2003). The original SnowCCI algorithm approach combines Tb differences with in situ snow depth observations, which considerably improves SWE estimation relative to a satellite-only retrieval (Pulliainen, 2006; Takala et al., 2011).

    A recent bias-correction method combines the original SnowCCI data with extensive ground-based snow course SWE

155   measurements, which significantly reduces the uncertainty of NH SWE estimation (Pulliainen et al., 2020). The method decreases the uncertainty of hemisphere-mean SWE estimation from 33% to 7.4%. The bias-corrected SnowCCI data are mapped to a 25 km EASE-Grid and are available from year 1979. The data cover non-mountainous regions, and glaciers and ice sheets are excluded. Mountainous regions, glaciers, and ice sheets are excluded from the data. The original SnowCCI data are available around the year, while bias-corrected SnowCCI data are only available from February to May. Despite limitations

160   in its temporal coverage, we have used the bias-corrected data in this study. We chose to do this because as the bias-correction method significantly reduced the uncertainty making and made the observational data more accurate, which, in turn, makes the comparison with the models also far more accurate and reliablemeaningful.

Table 1: CMIP6 models used in this study.

Nominal resolution	Institution	Model	Experiment ID	Reference
100 km	Beijing Climate Center (BCC)	BCC-CSM2-MR	historical	Wu et al. (2018a)
			esm-hist	Wu et al. (2018b)
	National Center for Atmospheric Research (NCAR)	CESM2	historical	Danabasoglu (2019a)
		CESM2-WACCM	historical	Danabasoglu (2019b)
	Lawrence Livermore National Laboratory (LLNL)	E3SM-1-0	historical	Bader et al. (2019)
	EC-Earth-Consortium	EC-Earth3	historical	EC-Earth (2019a)
		EC-Earth3-AerChem	historical	EC-Earth (2020a)
		EC-Earth3-CC	esm-hist	EC-Earth (2021)
		EC-Earth3-Veg	historical	EC-Earth (2019b)
	Geophysical Fluid Dynamics Laboratory (NOAA-GFDL)	GFDL-CM4	historical	Guo et al. (2018)
		GFDL-ESM4	historical	Krasting et al. (2018a)
			esm-hist	Krasting et al. (2018b)
	Max Planck Institute for Meteorology (MPI-M)	MPI-ESM1-2-HR	historical	Jungclaus et al. (2019)
	Meteorological Research Institute (MRI)	MRI-ESM2-0	historical	Yukimoto et al. (2019)
	Norwegian Climate Consortium (NCC)	NorESM2-MM	historical	Seland et al. (2019a)
	Seoul National University (SNU)	SAM0-UNICON	historical	Park and Shin (2019)
	Research Center for Environmental Changes, Academia Sinica (AS-RCEC)	TaiESM1	historical	Lee and Liang (2020)
250 km	Commonwealth Scientific and Industrial Research Organisation (CSIRO)	ACCESS-CM2	historical	Dix et a. (2019)
		ACCESS-ESM1-5	historical	Ziehn et al. (2019a)
			esm-hist	Ziehn et al. (2019b)
	Alfred Wegener Institute (AWI)	AWI-ESM-1-1-LR	historical	Danek et al. (2020)
	Beijing Climate Center (BCC)	BCC-ESM1	historical	Zhang et al (2018)
	National Center for Atmospheric Research (NCAR)	CESM2-WACCM-FV2	historical	Danabasoglu (2019c)
	EC-Earth-Consortium	EC-Earth3-Veg-LR	historical	EC-Earth (2020b)
	Goddard Institute for Space Studies	GISS-E2-1-G	historical	NASA/GISS (2018)
		GISS-E2-1-G-CC	esm-hist	NASA/GISS (2019a)
		GISS-E2-1-H	historical	NASA/GISS (2019b)
	Institut Pierre Simon Laplace	IPSL-CM5A2-INCA	historical	Boucher et al. (2020)
		IPSL-CM6A-LR	historical	Boucher et al. (2018)
		IPSL-CM6A-LR-INCA	historical	Boucher et al. (2021)
	Japan Agency for Marine-Earth Science and Technology (JAMSTEC)	MIROC6	historical	Tatebe and Watanabe (2018)
	Max Planck Institute for Meteorology (MPI-M)	MPI-ESM-1-2-HAM	historical	Neubauer et al. (2019)
		MPI-ESM1-2-LR	historical	Wieners et al (2019a)
			esm-hist	Wieners et al (2019b)
	Norwegian Climate Consortium (NCC)	NorESM2-LM	historical	Bentsen et al. (2019)
			esm-hist	Seland et al. (2019b)
500 km	Canadian Centre for Climate Modelling and Analysis, Environment and Climate Change Canada	CanESM5	historical	Swart et al. (2019a)
			esm-hist	Swart et al. (2019b)

165

As the bias-corrected SnowCCI data are only available for non-mountainous grid cells, we have used an average of the MERRA-2, Brown and Crocus v7 SWE products for the mountainous regions. MERRA-2 is a NASA (National Aeronautics

and Space Administration) atmospheric reanalysis, and it is available from year 1980. The spatial resolution of the data is 0.625°×0.5° (Gelaro et al., 2017). From MERRA-2, we have used the SWE product (Gelaro et al., 2017; GMAO, 2015a).

170 Brown SWE product, in turn, uses a simple snow scheme driven by ERA-Interim reanalysis (Brown et al., 2003). For the third SWE dataset, we have used Crocus version 7 product, which is a physical snow model driven by ERA-Interim reanalysis (Brun et al., 2013). Both MERRA-2 and Crocus v7 tend to slightly overestimate SWE under 150 kg m-2 and underestimate SWE over 150 kg m-2 (Mortimer et al., 2020).

175 **Table 2: Observational and reanalysis datasets used in this study.**

Variable, unit	Dataset	Resolution	Reference
Snow water equivalent (SWE), mm	SnowCCI	25 km × 25 km, monthly	Luojus et al. (2021)
	Brown	0.75° × 0.75°, monthly	Brown et al. (2020)
	Crocus v7	0.5° × 0.5°, monthly	Brun et al. (2013)
	MERRA-2	0.5° × 0.625°, monthly	Gelaro et al. (2017) GMAO (2015a)
Precipitation (P), mm month^{-1}	GPCC V2018	0.5° × 0.5°, monthly	Schneider et al. (2018)
2 m air temperature (T), Kelvin	MERRA-2	0.5° × 0.625°, monthly	Gelaro et al. (2017) GMAO (2015b)
Forest cover fraction, %	ESA CCI Land Cover time-series v2.0.7	300 m × 300 m, monthly	ESA (2017)

Additionally, For precipitation (P) and temperature (T) reference data, we used GPCC (Global Precipitation Climatology Centre) Version 2018 precipitation (P) data (Schneider et al., 2018) and MERRA-2 (The Modern Era Retrospective analysis

180 for Research and Applications, Version 2) temperature (T) data (Gelaro et al., 2017; GMAO, 2015b). GPCC is a monthly precipitation product based on data from rain gauge stations, and the data are available on a 0.5-degree global grid from 1891 to the present (Schneider et al., 2018). The product agrees well with other precipitation products (Behrangi et al., 2016; Sun et al., 2018). The unit of CMIP6 precipitation data is kg m^{-2}-s^{-1}, whereas the unit of GPCC data is mm month^{-1}, so we converted the CMIP6 data to monthly values (kg m^{-2} month^{-1}) to make the units of the datasets equivalent to each other. All precipitation

185 data are presented here in units of kg m^{-2} month^{-1} which is equivalent to mm month^{-1}.

MERRA-2 is a NASA (National Aeronautics and Space Administration) atmospheric reanalysis, and it is available from year 1980. The spatial resolution of the data is 0.625°×0.5° (Gelaro et al., 2017). In this study, For T reference data, we have used the monthly mean 2 m air temperature product, which agrees well with observations in the arctic Arctic (Gelaro et al., 2017;

190 Simmons et al., 2017) and the mean values show very small biases (Bosilovich et al., 2015). MERRA-2 daily temperature tends to have a cool daytime bias and a warm nighttime bias (Bosilovich et al., 2015; Draper et al., 2018). However, this is not

a major issue for our study because we use the monthly mean product. In addition, MERRA-2 seems to underestimate global warming trends in the last years of our study period (Gelaro et al., 2017; Simmons et al., 2017). Additionally, we used the ESA CCI Land Cover time-series v2.0.7 (ESA, 2017) to study the effect of forest cover on the results.

195

3 Methods

We used the nearest neighbor method to resample CMIP6, Brown, Crocus v7, MERRA-2 and GPCC data to the 25-km equal-area projection. ~~The SnowCCI data are only available for non-alpine regions, so we filtered out the corresponding grid cells from other datasets as well. The difference between each model and observation was calculated by~~ The fractional forest cover

200 was calculated from the higher-resolution ESA CCI Land Cover time-series as the fraction of forest cover grid cells of all cells within each 25-km × 25-km grid area. The bias-corrected SnowCCI data are available only for non-mountainous regions. Therefore, we used the bias-corrected SnowCCI data for the non-mountainous regions, and for the grid cells that were defined as mountainous in SnowCCI data, we used the mean SWE of the Brown, MERRA-2 and Crocus v7 datasets. The complex terrain causes uncertainties in SWE estimates but averaging over multiple products can improve the accuracy (Mortimer et al.,

205 2020). We calculated the model biases by subtracting the observation value from the model value, i.e. model minus observation. ~~We~~ and compared the biases~~differences~~ grid cell by grid cell. Our study covered land areas north of 40° N and years between 1982 and 2014. In this study, we mainly concentrated on snow-covered areas, i.e. grid cells where SWE > 10 kg m^{-2}. The snow-covered area was computed individually for each model and for each month. We have also filtered out grid cells with modeled SWE above 2~~1~~000 kg m^{-2}, as those values greatly exceed observed SWE (Stuefer et al., 2013).

210

We focus the analysis on two seasons: winter and spring. For the winter season, we consider only the SWE in February, since bias-corrected SnowCCI data are only available from February to May. We studied through linear regression analysis, how the ~~difference in~~ SWE bias in February ~~between the models and the observations~~ depends on the ~~difference in~~ precipitation (P) bias and temperature (T) biases, ~~averaged~~summed over the three preceding months from November to January:

215

$$\Delta SWE = \beta_T \Delta T_{cum} + \beta_p \Delta P_{cum} + C \tag{1}$$

where P_{cum} and T_{cum} are the precipitation and temperature summed over November-January, β_P and β_T are the regression coefficients and C is the constant. Here, as well as in Eq. (2) below, Δ refers to the model bias, i.e. the difference between the

220 modeled value (defined for each year separately) and the observed climatological value (averaged over the whole period considered). We used the climatological average for the observations because climate models cannot be expected to correctly simulate the weather conditions of individual years. Thus, the regression coefficients β_P and β_T depend only on the modeled interannual variations. The equations are presented in more detail in Appendix A.

For the spring season, we considered the monthly changes in SWE from February to March, from March to April, and from April to May. We defined the SWE change rate (SWE_change) as the difference in SWE between each month and the previous month. Positive values indicate an increase in SWE from one month to the next, and negative values a decrease. The model-minus-observation difference ~~in~~ ΔSWE_change (i.e., the model bias) was then regressed against the monthly ~~difference in~~ precipitation and temperature biases:

$$\Delta SWE_{change} = \beta_T \Delta T + \beta_P \Delta P + C \tag{2}$$

For example, when considering SWE_change from February to March, we used P and T for March. We pooled together the values of ΔSWE_{change}, ΔP and ΔT for the whole spring period (February through May) to determine the regression parameters β_P and β_T and C. The equations are presented in more detail in Appendix A.

We included only snow-covered grid cells (SWE > 10 kg m^{-2}) in the analysis and calculated the linear regressions only for grid cells with at least four values available during the study period. We calculated the linear regressions for the whole study period 1982-2014, and separately for three shorter periods: 1982-1991, 1992-2001 and 2002-2014.

By substituting into Eqs. (1) and (2) the mean model biases ~~differences between the models and observations~~, it is possible to split the model SWE biases ~~in SWE~~ into three components: the contribution of P (P_C), the contribution of T (T_C), and the contribution of other factors. For SWE in winter, P_C and T_C are:

$$P_C = \beta_P \Delta P_{cum,mean} \tag{3}$$
$$T_C = \beta_T \Delta T_{cum,mean} \tag{4}$$

Correspondingly, for SWE_change in spring, P_C and T_C are:

$$P_C = \beta_P \Delta P_{mean} \tag{5}$$
$$T_C = \beta_T \Delta T_{mean} \tag{6}$$

The third component in both winter and spring is the residual term, which is the constant from the regression equation (1) or (2). This is the contribution of other factors, including for example, inaccuracies in observational datasets and model parameterizations related to, for example, snow and surface energy budget. The residual (R) gives an estimate for the SWE bias that would remain if P and T were simulated correctly in the model.

4 Results

Figure 1 shows as an example the mean SWE of all CMIP6 high-resolution (100 km) models, the mean of SWE reference data and SnowCCI, and ΔSWE (CMIP6 SnowCCI – SWE reference) in April during 1982-2014. The corresponding figures for precipitation and temperature are in the supplementary material (Fig. S1S2). The SWE distribution has a large spatial variability: the highest values exceed 240 kg m^{-2} in both multi-model mean and SnowCCISWE reference data, and these values are found in northeastern Canada, around the Rocky Mountains, in Scandinavia, and in some parts of Siberia. Although the SWE distribution is similar for the multi-model mean and SnowCCISWE reference data, the models overestimate SWE in several regions, which are mostly located in the northern parts of the study area: in northeastern Canada, northeastern Siberia, and Eurasia around 90° E. In the southern parts of the study area, the multi-model mean mainly underestimates SWE.

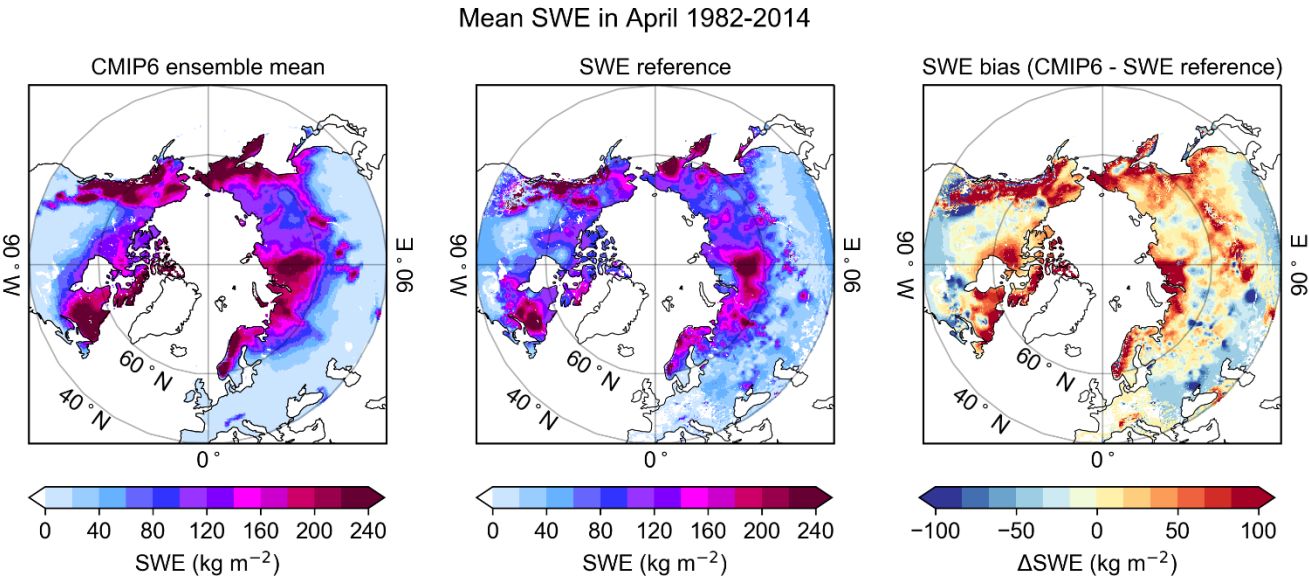



Figure 1. Mean SWE in April for the CMIP6 high-resolution (100 km) multi-model ensemble mean (left), SnowCCI and the SWE reference data (middle) and the difference CMIP6-SnowCCImodel SWE bias (right) for the period 1982-2014.

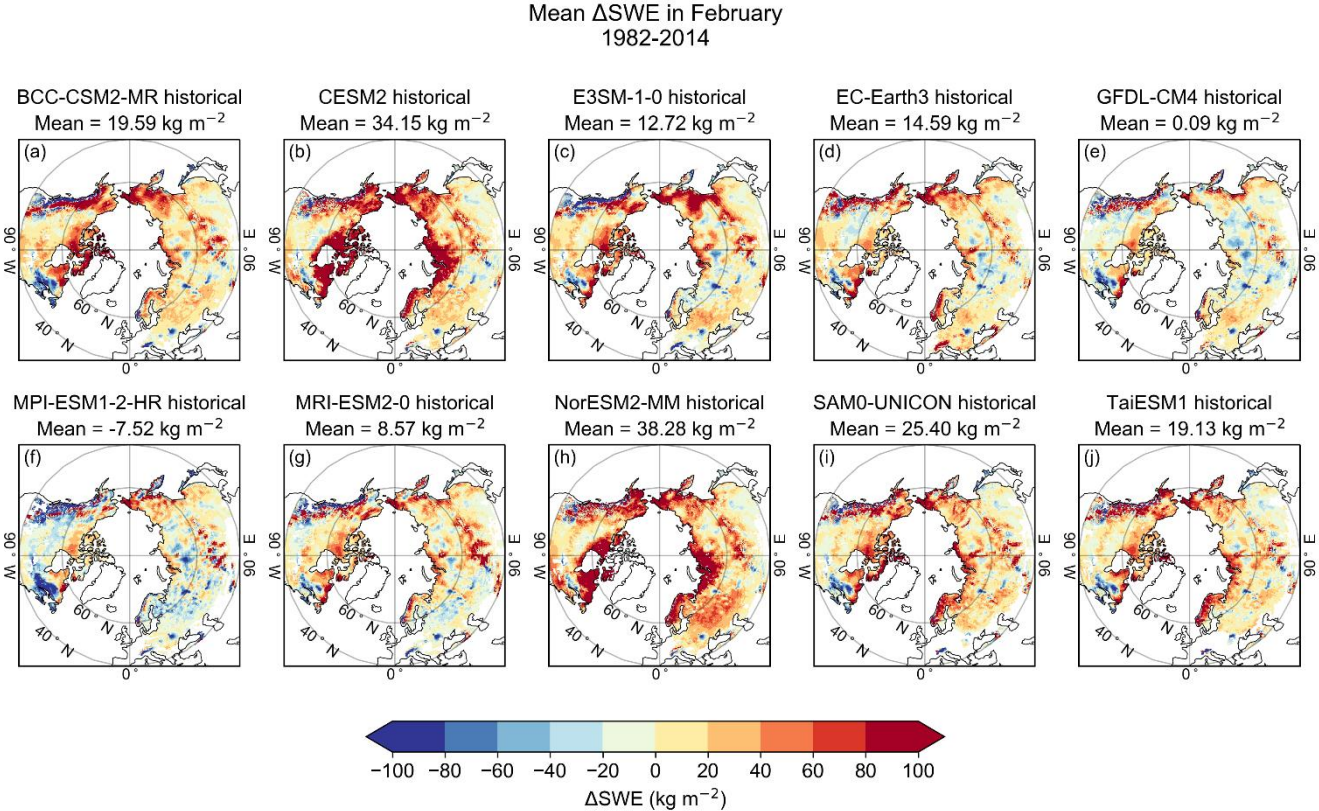
The monthly SWE sum of the whole study area varies considerably between the models. (Fig. 2). In February, March, and April, the modeled SWE (solid and dashed grey lines) vary by a factor of two, and in May, even by a factor of three. The variability between models is notably larger than the uncertainty range of the SWE reference estimate (blue markers). The low-resolution models (thin grey lines) do not significantly differ from the high-resolution models (dashed lines); the SWE sum is mostly in the same range in both resolution groups. Also, the mean values for both high-resolution (red markers) and low-resolution (purple markers) models are very close to each other. However, there are two low-resolution models that show very high SWE sum values in every month, which are clear outliers. These outliers are "GISS-E2-1-G historical" and "GISS-

285

Figure 2. Monthly SWE sum over the entire study area in February, March, April, and May separately for each high-resolution (100 km) CMIP6 model (grey dashed lines), for each low-resolution (> 100 km) CMIP6 model (thin solid lines), for the high-resolution CMIP6 multi-model ensemble mean (red markers), for the low-resolution CMIP6 multi-model ensemble mean (purple markers) and for SnowCCI the SWE reference data (blue markers). The blue shaded area indicates the 7.4% uncertainty range of
290 **the SnowCCI SWE reference data.**

Figure 2 shows the monthly SWE sum of the whole study area separately for each model (grey lines), for multi- model ensemble mean (red markers) and for SnowCCI (blue markers). The blue shaded area illustrates the 7.4% uncertainty range of SnowCCI

295

Mean ΔSWE in February
1982-2014

300 **Figure 3. The M~~m~~ean ~~difference~~ model bias in SWE (ΔSWE, model minus observation) in snow-covered areas ~~between each CMIP6 model and SnowCCI~~ in February 1982-2014.**

4.1 SWE in winter

The mean SWE bias in February shows large spatial variability within each model but also varies much between the models (Figs. 3 and S3). ~~Figure 3 shows the mean difference in SWE in snow-covered areas between each model and SnowCCI in~~
305 ~~February for the entire study period 1982-2014. Large variability exists between the models.~~ The areal-mean ~~difference between the models and SnowCCI~~model bias varies from about -~~7~~8 kg m^{-2} to ~~over~~ 40 kg m^{-2}. However, the largest negative and positive ~~differences~~biases are well concentrated in the same areas in all models. Overall, the models tend to overestimate the SWE in the northern parts of the study area, but also in southern Siberia. The negative biases, in turn, occur mostly in the

south and especially on the coastal areas of North America. In some models, there are negative biases also in the middle parts

310 of Eurasia. ~~Large biases also occur over Rocky Mountains in every model, but the sign of the bias varies.~~ The CESM2 and

NorESM2-MM ~~CESM2 WACCM~~ models show the largest overestimations. For both models, the ~~difference~~ bias is very high

in large regions in northern parts of North America and Eurasia. In these areas the relative bias is typically 150-200%. The

BCC-CSM2-MR, E3SM-1-0 and SAM0-UNICON models also show large positive ~~differences~~biases, which are concentrated

in the same areas, although, the ~~differences~~ biases are clearly smaller than for the CESM2 and NorESM2-MM ~~CESM2

315 WACCM~~ models. In other models, the areal-mean ~~differences~~ biases are closer to 0 kg m^{-2}, however, regional differences

exist. Overall, the GFDL models are the most consistent with the ~~SnowCCI~~ SWE reference data during February.

Mean ΔP_{cum} in Nov-Jan
1982-2014

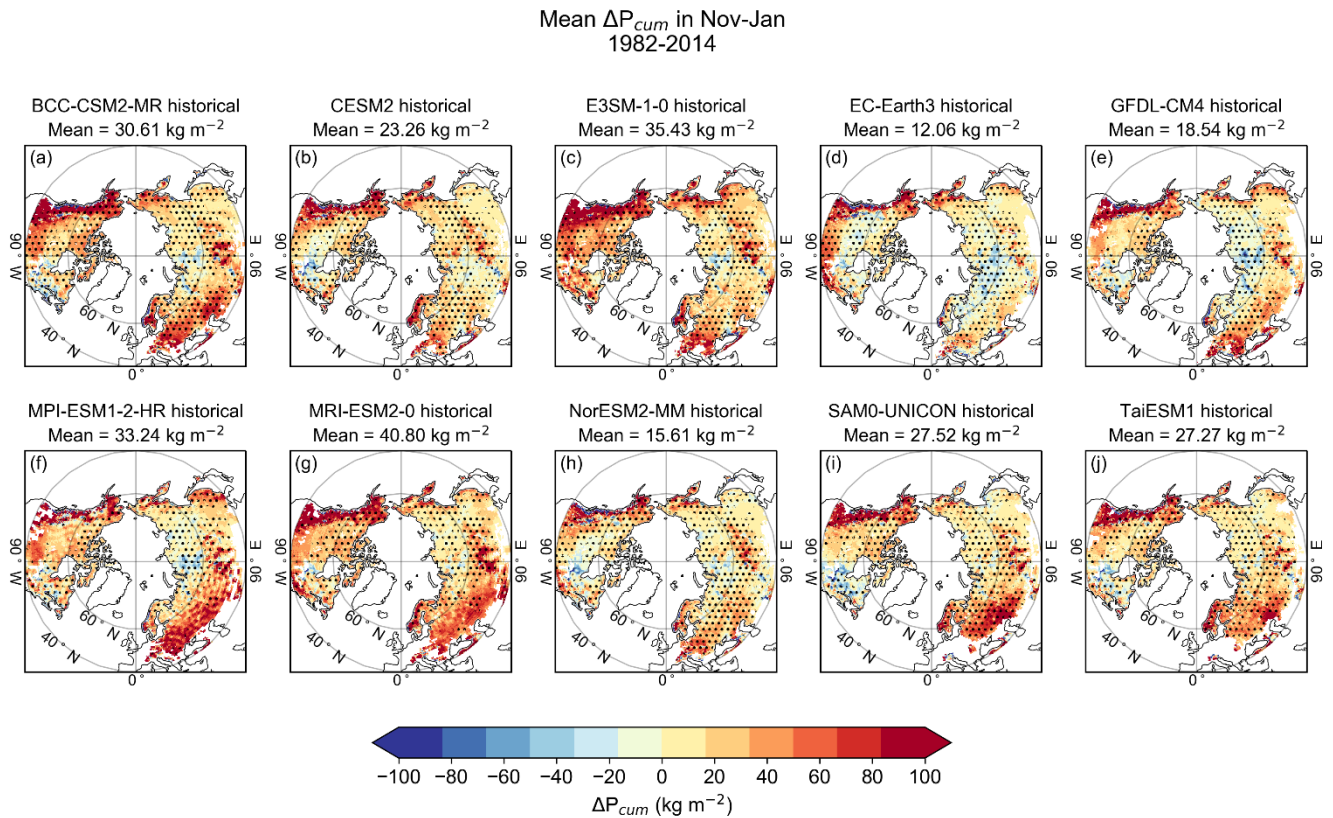



Figure 4. The ~~M~~mean ~~difference in~~ model bias in P$_{cum}$ (ΔP_{cum}, model minus observation) in snow-covered areas ~~between each CMIP6

320 model and GPCC in winter~~from November to January **1982-2014. The dots indicate areas where the models either overestimate both SWE and P or underestimate both SWE and P.**

~~Figure 4 shows the mean difference in monthly P~~$_{cum}$ ~~in snow-covered areas between each model and GPCC in winter for the

entire study period 1982-2014. The dots indicate areas where models either overestimate both SWE and P~~$_{cum}$ ~~or underestimate

325 both SWE and P~~$_{cum}$~~, i.e. the areas where differences in P~~$_{cum}$ ~~could logically explain the discrepancies in SWE. Overall, a~~All

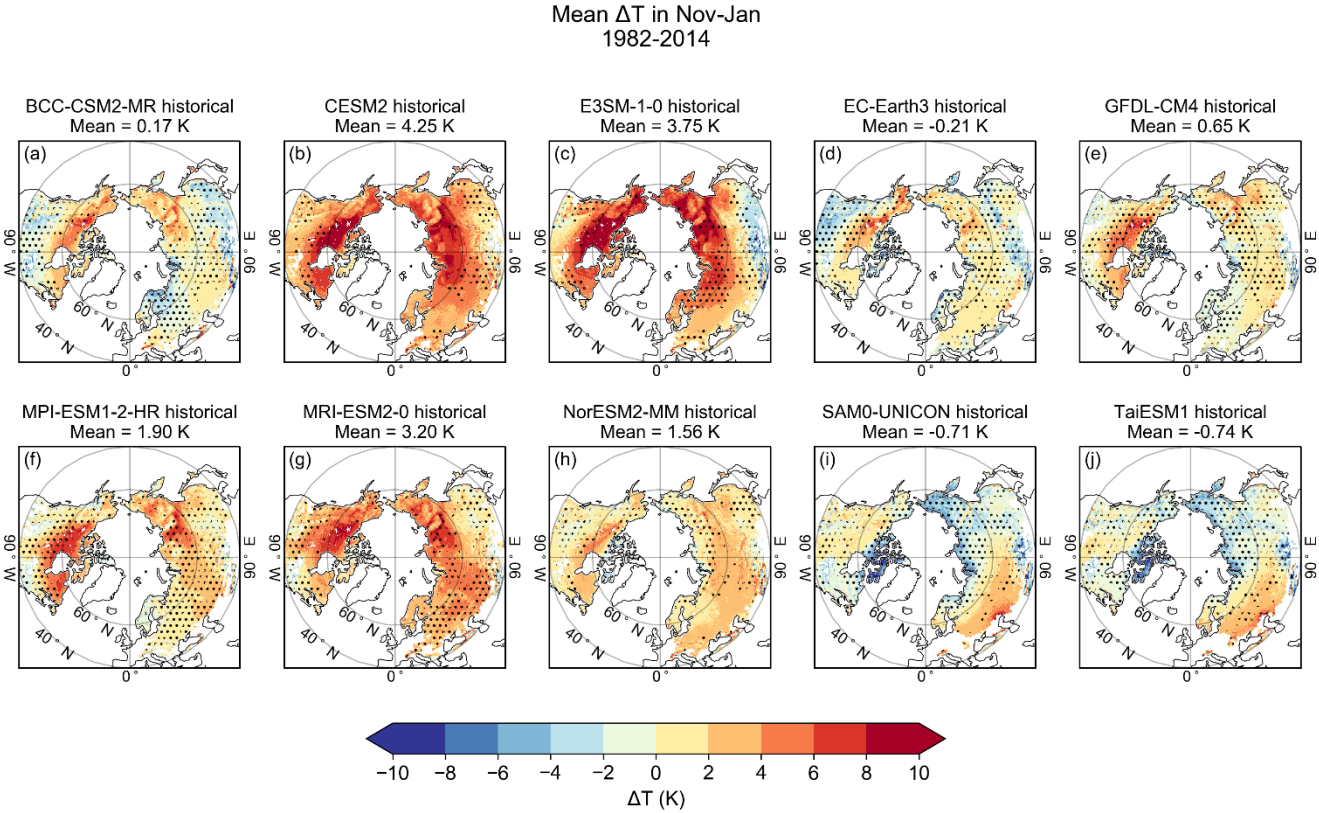
models overestimate NH extratropical precipitation ~~in winter~~from November to January (Figs. 4 and S4). The largest overestimations are mainly in southern regions and in coastal areas. There are small areas where underestimation occurs, especially in Eurasia around 90° E. In every model, there are large ~~dotted~~ regions where P_{cum} bias could logically explain the SWE bias (marked with dots), i.e. areas where models overestimate or underestimate both SWE and P_{cum}. These regions are mostly in the northern parts of the study area, whereas in the south, there are more areas where SWE and P_{cum} discrepancies are more often not consistent with each other.

Mean ΔT in Nov-Jan
1982-2014

Figure 5. The ~~M~~mean ~~difference~~ model bias in T (ΔT, model minus observation) in snow-covered areas ~~between each CMIP6 model and MERRA-2 in winter~~from November to January 1982-2014. We show ΔT (ΔT_{cum} divided by 3) instead of ΔT_{cum} so that the values are more intuitive and easier to interpret. The dots indicate areas with either cold bias and positive SWE bias or warm bias and negative SWE bias, i.e., the areas where T bias could logically explain the SWE bias.~~where models simulate either too high SWE and too low T or too low SWE and too high T.~~

~~Figure 5 shows the mean difference in monthly mean T in snow-covered areas between each model and MERRA-2 in winter for the entire study period 1982-2014. We show ΔT (ΔT_{cum} divided by 3) instead of ΔT_{cum} so that the values are more intuitive and easier to interpret. The differences~~For model T biases, the values are mostly positive; however, regional and inter-model variability exists (Figs. 5 and S5). We show ΔT (ΔT_{cum} divided by 3) instead of ΔT_{cum} so that the values are more intuitive and

easier to interpret. The CESM2 and E3SM-1-0 ~~CESM2-WACCM~~ models generally simulate too warm temperatures, and the largest positive ~~differences~~ biases are in the northern parts of the study area. The GFDL, MPI-ESM1-2-HR and MRI-ESM2-0 models simulate too warm temperatures in the north, while the SAM0-UNICON and TaiESM1 model~~s~~, in turn, simulate~~s~~ too cold temperatures in the north. The BCC-CSM2-MR and~~,~~ EC-Earth3 ~~and EC-Earth3-Veg~~ models are the most consistent with the MERRA-2 data. In every model, there are large ~~dotted~~ areas where the signs of biases for T and SWE are opposite (marked with dots), indicating that biases in T might explain biases in SWE. However, in these areas, the ~~differences~~ biases are mainly quite small.

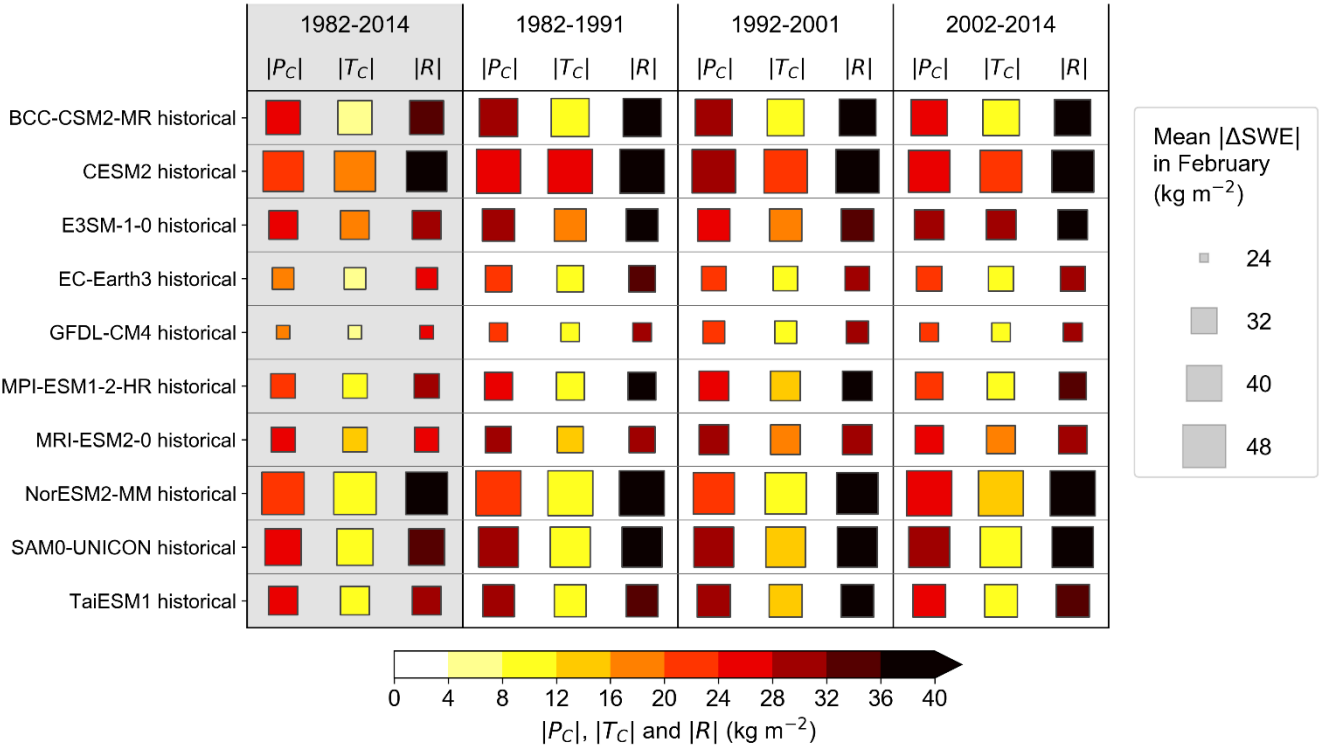



Figure 6. The areal-means of the absolute values of ΔSWE, P_C, T_C, and residual R calculated for the entire study period 1982-2014 (left column, shaded with grey) and for three shorter time periods (1982-1991, 1992-2001, and 2002-2014) for each model in winter. The size of the square indicates the absolute value of ΔSWE of that time period and model, and the color of the square indicates the absolute value of P_C, T_C, and R.

The contributions to ~~differences in~~ SWE biases (ΔSWE) due to precipitation biases (P_C), temperature biases (T_C) and other factors (residual R) are quantified using the regression equations (1), (3) and (4). To summarize their relative importance, Fig~~ure~~s 6 and S6 show~~s~~ the areal-means of the absolute values of ΔSWE, P_C, T_C, and R. The mean ΔSWE varies from under 30 kg m^{-2} in the GFDL-CM4 model to around 50 kg m^{-2} in the CESM2 and NorESM2-MM models. ~~. The CESM2 and CESM2-WACCM models show largest ΔSWE, whereas in other models, ΔSWE is clearly smaller.~~ In all models, the contribution of P

on ΔSWE is clearly larger than the contribution of T. However, the residual is also typically large, indicating that P and T cannot explain the SWE biases alone. This implies that observational uncertainty or model structural factors, such as parameterizations related to the surface energy budget or hydrology, play a considerable part in the observed SWE differencesbiases. We studied the terms' statistical significance by using the t-test and found that all the terms (P_C, T_C, and R) are significantly different from zero at the 99% confidence level. The variability in these parameters between the decadal subperiods and the full three-decade analysis period was slight, indicating consistent behavior across time in both models and observations.

Figures 7 and S7 shows the spatial distribution of the contributions of P and T and the residual for each model for the entire study period 1982-2014. The spatial distribution of regression parameters R^2, $β_P$ and $β_T$ are shown in the supplementary material (Figure Fig. S2S8). Also, the contributions of P and T and the residual calculated for the shorter time periods are in the supplementary material (Figs.Figures S9-S11S3-S5). Figure. 7 shows that overall, the contribution of P is larger than the contribution of T, as Fig. 6 already indicated. P contributes to ΔSWE especially over northern and coastal regions, with fairly similar patterns for all models considered. The regression coefficient $β_P$ also shows large values ($β_P ≈ 1$) especially in the northern and middle parts of both continents (Fig. S2S8), with relatively small inter-model variations. This is consistent with the expectation that in cold regions, an increase of precipitation translates into a similar increase in SWE.

The contribution of T is mostly very weak (Fig. 7); however, for some models, T shows stronger contribution especially over western parts of Eurasia and over northeastern Canada. The regression coefficient $β_T$ is mostly negative or very close to zero (Fig. S2S8). The negative correlation is strongest in Europe and southern parts of North America and Eurasia. In these regions, the temperature is close to 0 °C in winter, which makes temperature an important driver of the SWE. Northern Canada and Siberia, in turn, show areas with positive correlation, meaning that warmer temperatures cause higher SWE. Studies have shown that when the temperature rises due to climate change, the winter precipitation will also increase (Brown and Mote, 2009; Park et al., 2012; Räisänen, 2008). In the south, warming temperatures will shift winter precipitation from snow to rain. In the north, in turn, temperature will stay below 0 °C despite the warming, which will lead to an increase in snowfall in the coldest regions of NH, and therefore, to an increase in SWE. This phenomenon is most likely seen here as well; warmer temperatures in the models will increase winter precipitation, resulting in too high SWE in the models. In fact, this exposes a limitation of the regression Eq. (1): it treats $ΔT_{cum}$ and $ΔP_{cum}$ as independent variables, which is not fully realistic. When these variables are correlated, their contributions to ΔSWE cannot be fully separated. It should be noted that since Eq. (1) treats

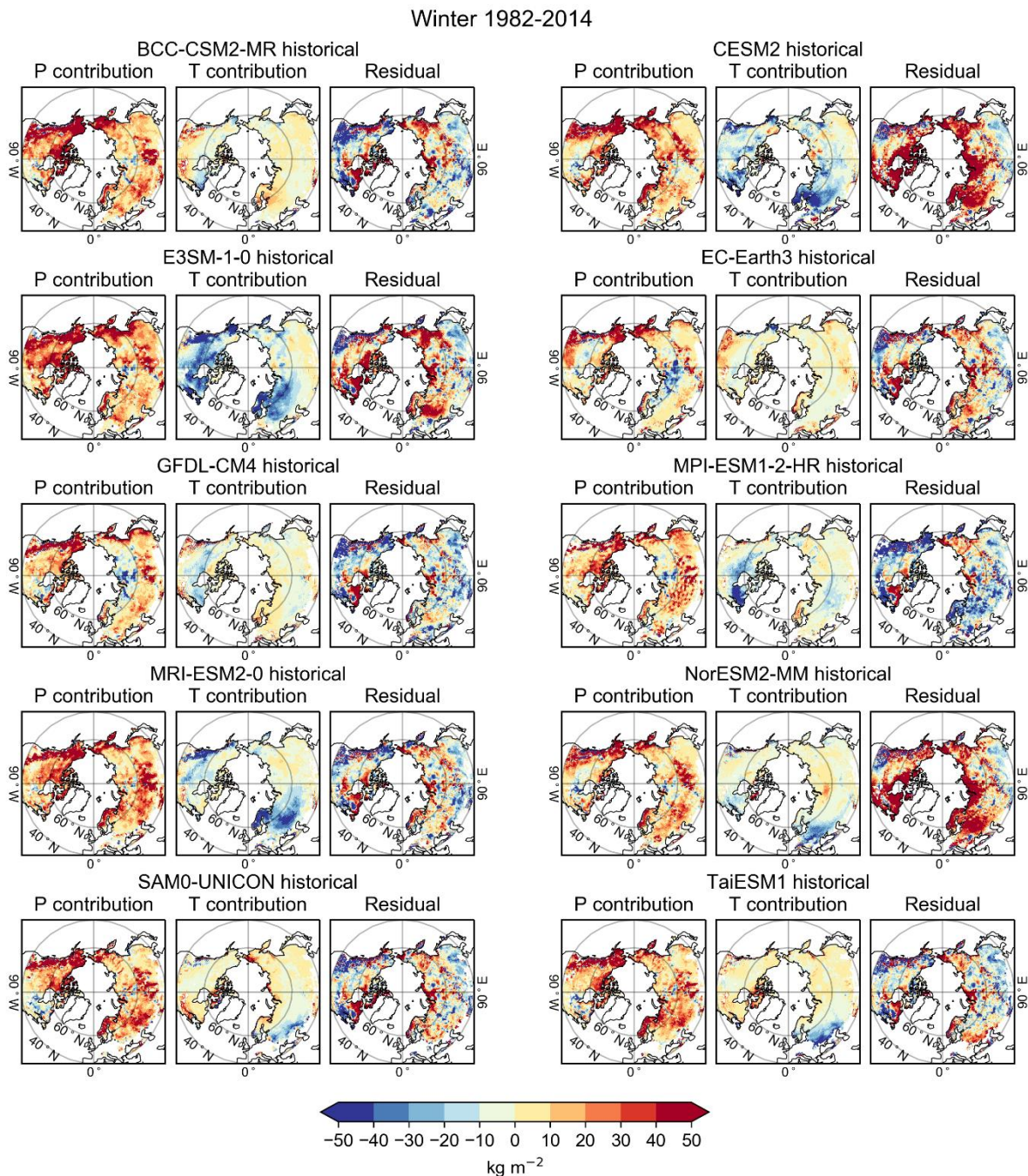



Figure 7. Spatial distribution of the P contribution, the T contribution, and the residual for each model in winter 1982-2014.

The residual shows large spatial and inter-model variability (Fig. 7). Especially for the CESM2 and ~~CESM2-WACCM~~NorESM2-MM models, the residual shows very large positive values. These large positive residuals are mainly concentrated in the same areas where the models clearly overestimate SWE (Fig. 3). This indicates that, for these models, the large SWE ~~differences~~ biases in these areas are mainly caused by some other factors than P or T. For other models, the residual shows both positive and negative values across the study area. The possible factors causing the large residual term are discussed in more detail in Sect. 5.2.

Mean ΔSWE_{change} in spring
1982-2014

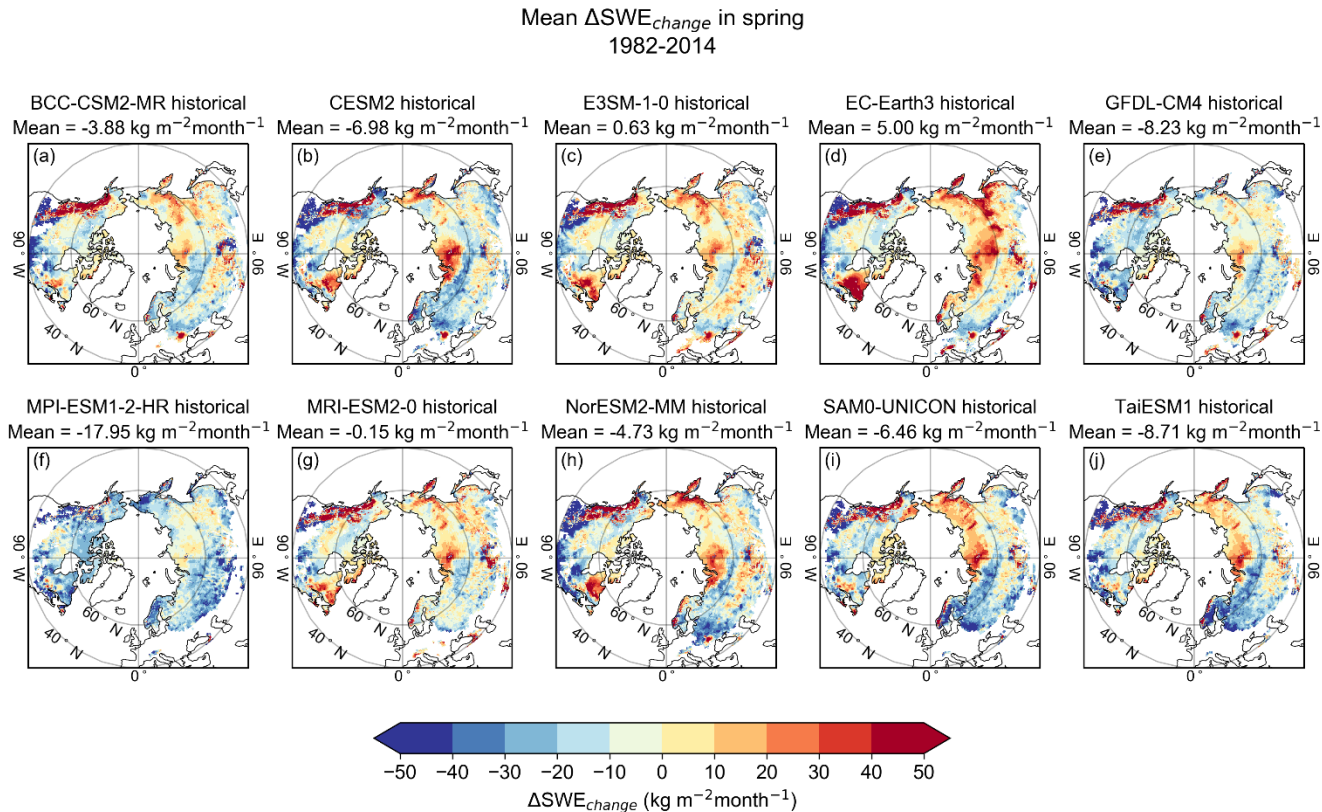



Figure 8. The M~~m~~ean ~~difference~~ model bias in the monthly SWE change (ΔSWE_{change}, model minus observation) in spring ~~between each CMIP6 model and SnowCCI~~ for the period 1982-2014.

~~3~~4.2 Monthly SWE change in spring

Figure~~s~~ 8 and S13 show~~s~~ the mean ~~difference~~ bias in monthly SWE change (ΔSWE_{change}) in spring in the snow-covered areas ~~between each model and SnowCCI~~ for the whole study period 1982-2014. A figure showing the SWE change rate of multi-model mean, SWE reference data and bias in multi-model mean is in the Supplementary material (Fig. S12). Positive ΔSWE_{change} means that there is less snowmelt~~snow melts more slowly~~ in the models compared to ~~SnowCCI~~the SWE reference data, and negative ΔSWE_{change} indicates that snow melts faster in models, respectively. The areal-mean ΔSWE_{change} is mainly

negative in every model, which means that snow melts generally faster in the models compared to ~~SnowCCI~~the SWE reference data. However, there is a large spatial variability in every model and the inter-model variability is also large. In the CESM2

415   and NorESM2-MM~~CESM2-WACCM~~ models, three areas show distinctly positive ΔSWE$_{change}$: northeastern Canada, northern Siberia, and eastern Siberia. In all these areas, the SWE ~~difference~~bias in February (Fig. 3) was already clearly positive, meaning that these models greatly overestimate SWE in these areas also in spring. The EC-Earth3 ~~and EC-Earth3-Veg~~ model~~s~~ shows clear positive ΔSWE$_{change}$ in northeastern Canada but also in Eurasia. The area with positive ~~differences~~biases in Eurasia is very extensive and differs notably from the other models. The EC-Earth3 model has been found to drastically overestimate

420   the snow cover extent as well (Mudryk et al., 2020). The SAM0-UNICON and TaiESM1 model~~s~~ also shows positive values in northern Siberia. The GFDL and MPI-ESM1-2-HR models, in turn, show large areas with negative ΔSWE$_{change}$ in northern Canada and in eastern parts of Siberia, which differs from the other models. Overall, the model ~~minus observation differences~~biases in SWE change rate in spring (Fig. 8) show larger inter-model variations than the corresponding ~~differences~~biases in SWE in February (Fig. 3).

425

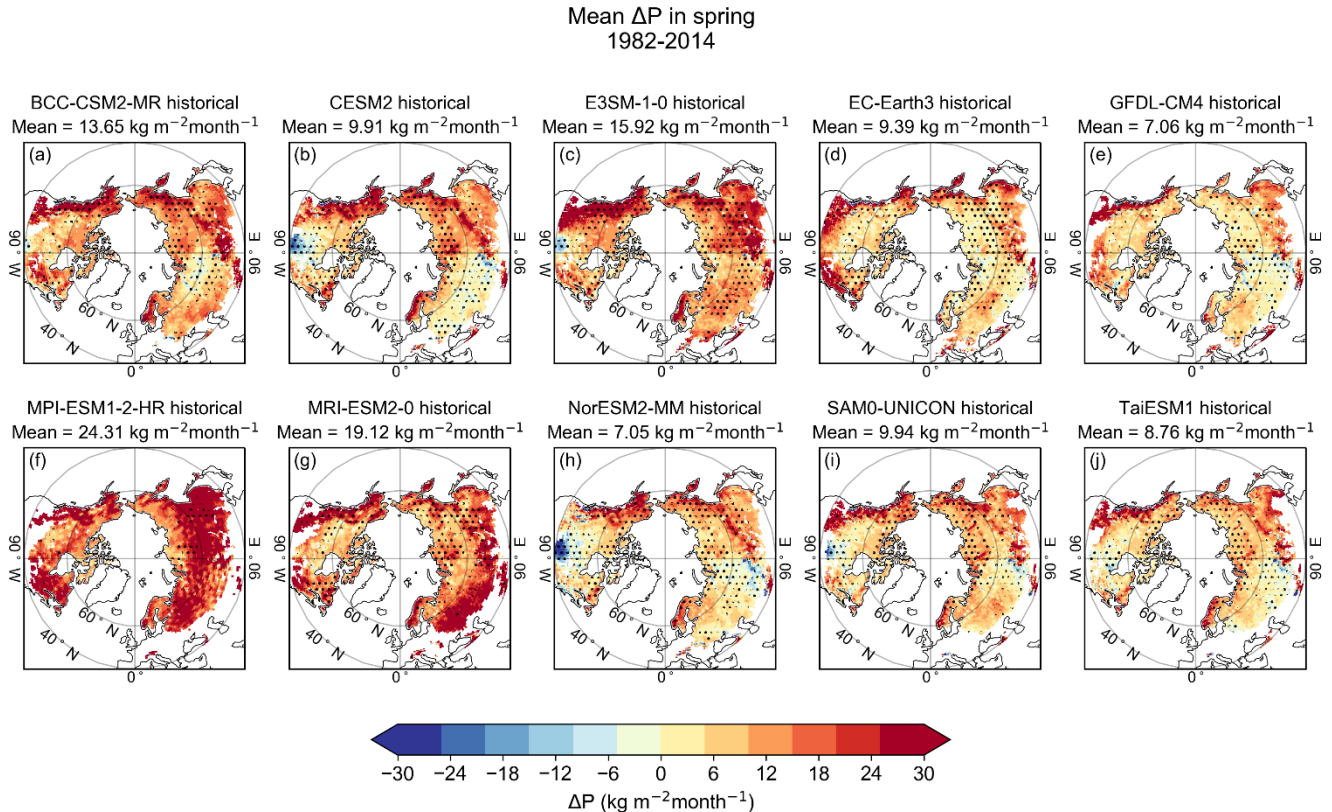



Figure 9. The ~~M~~mean ~~difference~~model bias in P (ΔP, model minus observation) in snow-covered areas ~~between each model and GPCC~~ in spring 1982-2014. The dots indicate areas where models either overestimate both SWE$_{change}$ and P or underestimate both SWE$_{change}$ and P.

430 As in winter (Fig. 4), models on average overestimate precipitation in spring as well (Figs. 9 and S14). The largest overestimations occur mainly in southern regions and in coastal areas.- The regions with mutual~~ly~~ biases in P and SWE_{change} (marked with dots) show large inter-model variability, and they are less extensive than in winter (Fig. 4). This indicates that precipitation is not as important factor in spring ~~than~~ as it is in winter.

Mean ΔT in spring
1982-2014

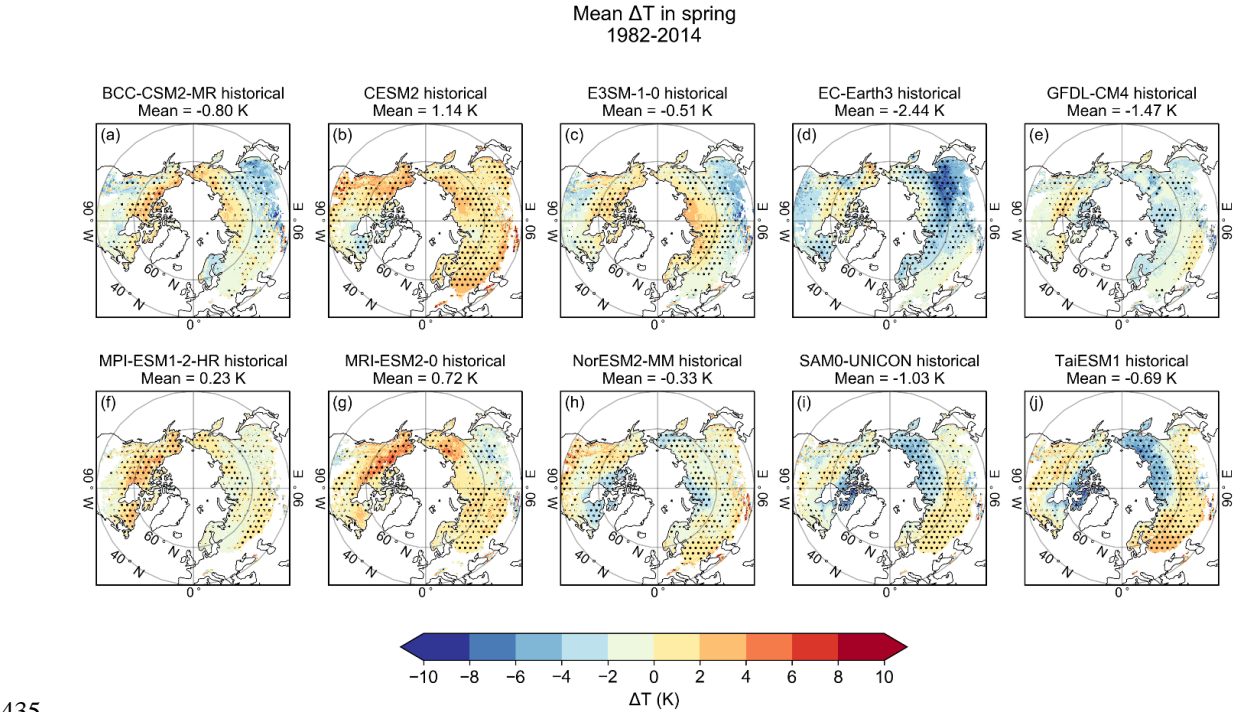



Figure 10. The ~~M~~mean ~~difference~~ model bias in T (ΔT, model minus observation) in snow-covered areas ~~between each CMIP6 model and MERRA-2~~ in spring 1982-2014. The dots indicate areas with either cold bias and positive SWE bias or warm bias and negative SWE bias, i.e., the areas where T bias could logically explain the SWE bias. ~~where models simulate either too positive SWE_change and too low T or too negative SWE_change and too high T.~~

440

~~Figure 10 shows the mean difference in T in snow-covered areas between each model and MERRA-2 in spring for the entire study period 1982-2014. There is~~ The mean T bias in spring shows a large spatial variability within each model but also the variability between models is very clear (Figs. 10 and S15). The BCC-CSM2-MR, E3SM-1-0, ~~GFDL-ESM4,~~ MPI-ESM1-2-HR, and MRI-ESM2-0 models show a warm bias in the northern parts of the study area, whereas the SAM0-UNICON and

445 TaiESM1 models show~~s~~ a cold bias in the same area. The EC-Earth3 ~~and EC-Earth3-Veg~~ models, in turn, ~~have~~ has a cold bias in eastern Eurasia near 60° N, which clearly differs from the other models. These areas also exhibited less snowmelt than the SWE reference data, which indicates that the bias in snow melt rate can be caused by the bias in T. The sizes and locations of the ~~dotted areas (i.e.,~~ areas with mutually consistent biases in T and SWE_{change}~~)~~ (marked with dots) vary greatly between models. Especially in the GFDL-CM4 model, the size of these areas is small, while in most of the models, the dots cover the

450   majority of the study area. Furthermore, in those regions where the biases in T and SWE$_{change}$ are consistent in spring, the cold or warm temperature biases are typically relatively large, when compared with the corresponding biases in winter- (Fig. 5). This indicates that biases in T are a more important driver of biases in SWE in spring than in winter.

   Figures 11 and S16 summarizes the areal-means of the absolute values of ΔSWE$_{change}$, the contribution of P (P$_C$), the
455   contribution of T (T$_C$), and the contribution of other factors (the residual R). ~~In all models, the residual is larger than P$_C$ or T$_C$. This suggests that overall, the biases in snow melt rate in spring are dominated by other factors than T or P.~~ The contributions of P and T are quite similar in magnitude in almost all ~~but varies between~~the models. ~~The EC-Earth3 model stands out from the other models, as T$_C$ is larger than in the other models.~~ None of the variables shows a large dependence on the period considered. The EC-Earth3 model stands out from the other models, as T$_C$ is larger than in the other models. The large
460   contribution of T in ΔSWE$_{change}$ was already seen in Figs. 8 and 10, as the spatial distribution of the biases were similar. This also suggests that the overestimated snow cover extent in the EC-Earth3 model (Mudryk et al., 2020) might be at least partly due to the cold bias. In all models, the residual is larger than P$_C$ or T$_C$. This indicates that overall, the biases in snow melt rate in spring are dominated by other factors than T or P. These factors will be further discussed in Sect. 5.2. We also studied the terms' statistical significance by using the t-test. Even though the contributions of T and P are considerably smaller than the
465   residual term, all the terms (P$_C$, T$_C$, and R) are significantly different from zero at the 99% confidence level.

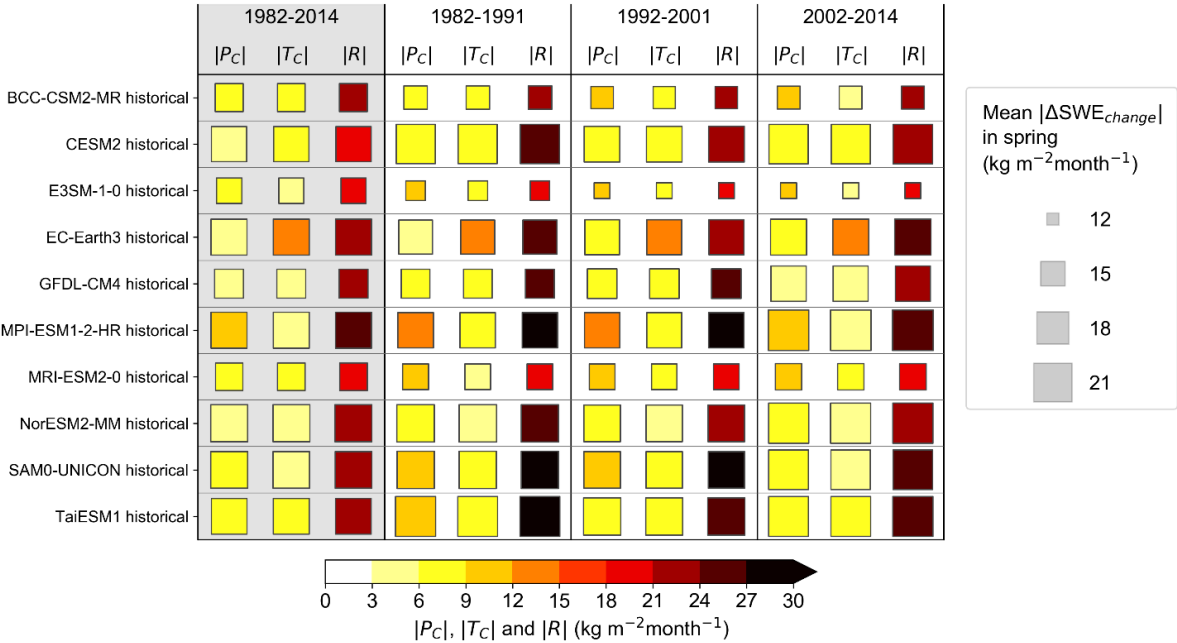



Figure 11. The areal-means of absolute values of ΔSWE$_{change}$, P$_C$, T$_C$, and residual R calculated for the entire study period 1982-2014 (left column, shaded with grey) and for three shorter time periods (1982-1991, 1992-2001, and 2002-2014) for each model in spring. The size of the square indicates the mean absolute value of ΔSWE$_{change}$ of that time period and model, and the color of the square
470   **indicates the mean absolute value of P$_C$, T$_C$, and R.**

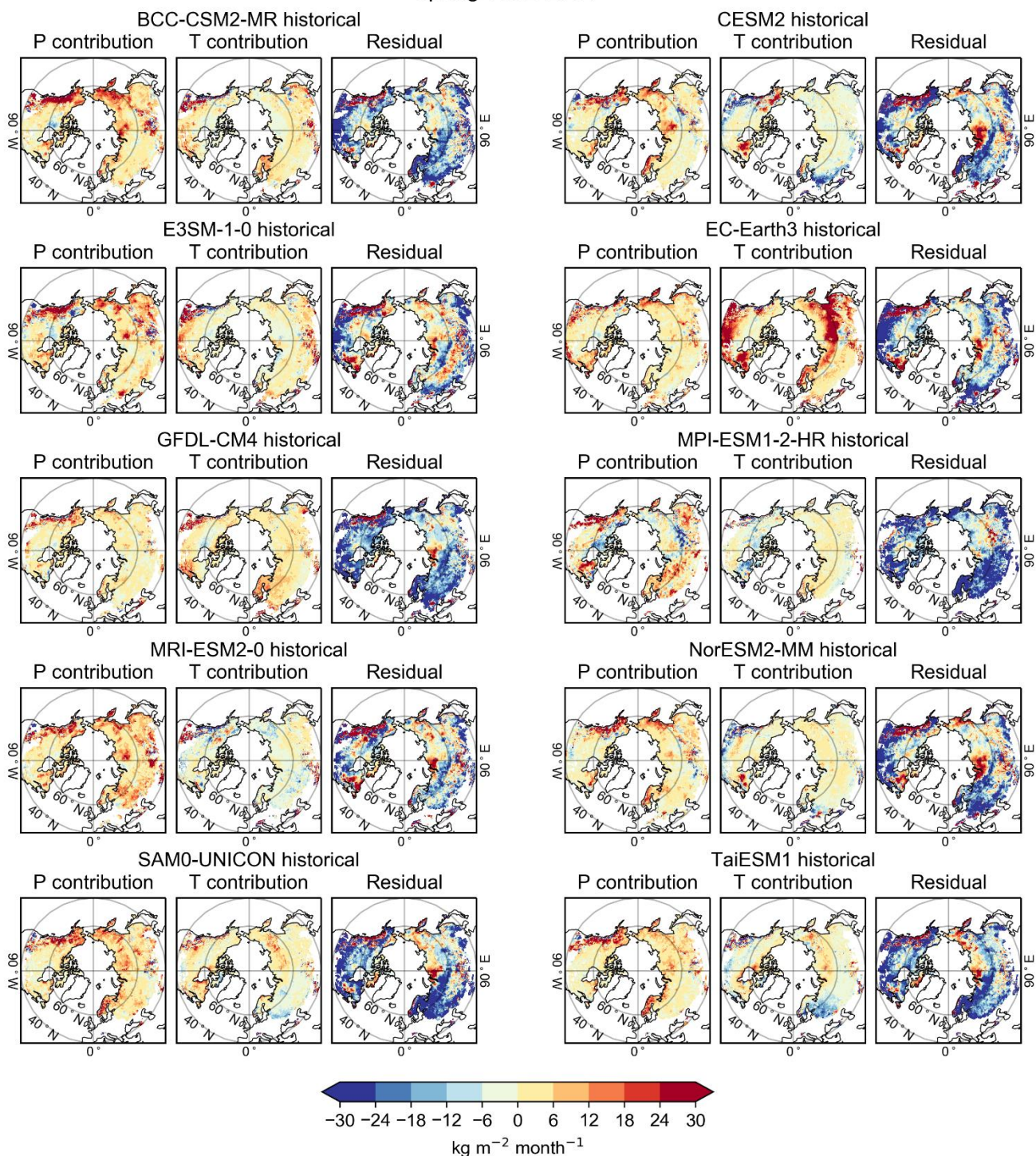



Figure 12. Spatial distribution of the P contribution, the T contribution, and the residual for each model in spring 1982-2014.

Figures 12 and S17 shows the spatial distribution of the contributions of P and T and the residual for each model for the entire study period 1982-2014. The spatial distribution of ~~regression parameters~~ R^2, β_P and β_T are displayed in the supplementary material (~~Figure~~ Fig. ~~S6~~S18). Also, the contributions of P and T and residual calculated for the shorter time periods are in the supplementary material (Figs. S19-S21~~Figures S7-S9~~). P contributes to ΔSWE_{change} mostly in Alaska and northern Siberia, but the effect varies between models. Furthermore, even though T showed clear warm and cold biases in many areas (Fig. 10), the contribution of T is mostly quite weak, because of the small regression coefficient β_T (Fig. S6). However, exceptions exist; especially, the EC-Earth3 and EC-Earth3-Veg models stand out, as in Eurasia, there is a large area where a negative bias in T (Fig. 10) contributes substantially to a positive bias in SWE_{change} (Fig. 8). Also, for the CESM2 and GFDL-ESM4 models, T shows a stronger contribution over northern parts of North America. Overall, however, the contributions of both P and T are small compared to the residual R, which is consistent with Fig. 11. This indicates that other factors than T or P are the dominant drivers for the SWE_{change} discrepancies. The residual is mostly negative in all the models, which means that snow would melt too fast in the models, if T and P were simulated correctly. This issue is discussed further in Sect. 5.24.

5 Discussion

~~We have evaluated NH SWE in CMIP6 models with satellite-based SnowCCI data for the period 1982-2014. While SWE in CMIP6 models has been previously studied by Mudryk et al. (2020), this is to our knowledge the first study to analyze CMIP6 SWE together with temperature and precipitation. In addition, the recent bias correction method has significantly narrowed down the confidence limits of the NH SWE estimate, which makes the comparison more accurate and provides a more reliable analysis of the models' ability to describe the snow cover. The continuous development of climate models is crucial, so that the changing climate can be simulated as reliably as possible.~~

The analysis shows that CMIP6 models overestimate the total SWE mass over the entire study area with a few exceptions (Fig. 2), which is consistent with the previous study (Mudryk et al., 2020). The overestimation of SWE was also observed in CMIP5 models (Mudryk et al., 2020; Santolaria-Otín and Zolina, 2020). The NH SWE sum reaches its peak value in March, but the peak values are overestimated by most of the models. As the spring advances, the variability between models increases, and some of the models clearly overestimate the SWE in May. In some models, in turn, snow melts faster than in the observations and they underestimate SWE in late spring. This is also shown in Figs. 3 and 8: in winter, the ~~differences in~~ SWE biases are mainly positive (Fig. 3), while in spring, there are large differences in snow melt rates between the models (Fig. 8).

~~In this study, we have analyzed whether differences in T and P between models and observations could explain the corresponding differences in SWE. The analysis focused, on one hand, on the SWE in February and, on the other hand, on the SWE change rate (SWE_change) during the spring from February to May. Using linear regression, model biases in SWE (ΔSWE)~~

and SWE~change~ (ΔSWE~change~) were divided into three components: the contribution of P, the contribution of T and the contribution of other factors.

5.1 The contribution of P and T to the SWE bias

In winter, the models mostly overestimate SWE (Fig. 3), but spatial and inter-model variability exists. The overestimations are mostly concentrated in the same areas in all models, but the magnitude varies greatly between the models. The models also overestimate precipitation in winter (Fig. 4), suggesting that overestimated SWE is caused by the overestimated P. The regression coefficient β_P also shows very clear correlation between ΔSWE and ΔP_{cum} (Fig. S2S8). Therefore, P clearly contributes to the ΔSWE bias, whereas the contribution of T is substantially smaller (Fig. 7). This result is consistent with the expectation that precipitation is needed to initiate the snow cover. In other words, even too cold temperatures cannot cause too high SWE without precipitation. Also, studies have shown that both the P bias (Zamani et al., 2020) and the SWE bias (Mudryk et al., 2020) have decreased from CMIP5 to CMIP6. This indicates that there is a connection between these variables, which is consistent with the results from this study.

The link between ΔT_{cum} and ΔSWE is strongest in the warmest regions of the study area (Figs. 7 and S2S8). Especially in the coldest regions, where T is well below 0 °C, variations in T do not significantly affect the amount of snow on the ground. In regions where T is closer to 0 °C in winter, T plays a more significant role and has a greater impact on SWE. This physically reasonable behavior suggests that climate models might be able to simulate SWE trends in the warming climate correctly, even if SWE itself is not reproduced accurately.

In spring, ΔSWE~change~ and ΔT show quite similar patterns in many models (Figs. 8 and 10), which indicates that biases in T affect the biases in SWE~change~. This result is to be expected because the increase in T is the main factor that causes snow to melt in spring. A relationship between temperature and snow cover in spring has also been observed in CMIP5 models (Brutel-Vuilmet et al., 2013; Mudryk et al., 2017). The CMIP5 models have been found to underestimate the observed trend towards a reduced snow cover extent due to an underestimation of the spring warming trend (Brutel-Vuilmet et al., 2013).

Even though ΔSWE~change~ and ΔT are quite consistent with each other, the contribution of T is not very strong (Fig. 12), because the regression coefficient β_T is small (Fig. S6S18). The only model showing clear contribution of T is EC-Earth3, as there is a large area in eastern Eurasia where positive values of ΔSWE~change~ co-occur with a cold bias. The spatial distribution of the bias in SWE~change~ in the EC-Earth3 model differs considerably from other models (Fig. 8). The model has been found to drastically overestimate snow cover extent though not so much the SWE (Mudryk et. al. 2020). This suggests that while the positive bias in SWE~change~ (i.e., underestimated snowmelt rate in spring) is related to the cold bias, the cold bias might be maintained by a too high surface albedo resulting from overestimated snow cover extent.

Several factors can weaken the regression coefficient β_T. The analysis shows that in spring, the relationship between ΔT and ΔSWE_{change} is strongest in western parts of Eurasia and in southern parts of North America (Fig. ~~S6~~S18). These are regions with the earliest snow melt onset. In these areas, T is the dominant factor causing snow to melt throughout the spring season. In the northernmost parts of the study area, the melt season begins later, so that early spring still belongs to the snow accumulation season. As a result, P may still be the dominant factor influencing ΔSWE_{change} in early spring, and T will become a more significant factor later in the spring. This can reduce the correlation and affect R^2, β_P and β_T ~~the linear regression parameters~~ in the northern parts of the study area. Additionally, if SWE in winter is simulated incorrectly, this can affect the melt rates in spring, as there can be too little or too much snow that can melt.

5.2 The residual term

In both winter and spring, the residual term R of the regressions is also significant (Figs. 6, 7, 11 and 12). This means that the model biases in SWE in winter and SWE_{change} in spring cannot be entirely explained by the biases in P and T, but other factors also contribute to these biases. These factors may include inaccuracy in the model parameterizations related to snow and surface energy budget, but also inaccuracy in the observational datasets.

The residual term R is particularly pronounced for the SWE change rate in spring, when it is typically larger than either the contribution of P or T. Interestingly, the residual is mostly negative (Fig. 12). The negative residual means that if P and T were simulated correctly in the models, snow would melt too fast in spring. While understanding the origins of this bias would be worth a separate study, a previous study with ECHAM5 (Räisänen et al. 2014) is of interest here. ECHAM5 is a predecessor of the atmospheric part of MPI-ESM1-2-HR, for which the residual R in Fig. 12 is especially strongly negative. This is consistent with the finding that in ECHAM5 snow melted generally too fast in spring, despite a cold bias in T (Räisänen et al. 2014). A major factor for this was that T was not calculated separately for snow-covered and snow-free parts of the grid cell. Because of that, T was not able to rise above 0 ℃ if there was snow left in the grid cell, and, therefore, a too large fraction of the available energy was used in melting the snow (Räisänen et al., 2014). ~~The parameterization of the surface albedo is another factor that may influence the snowmelt rate in spring. A too low (or high) surface albedo would speed up (or delay) the snowmelt. As a detail, the only region in which MPI-ESM1-2-HR displays a positive residual in Fig. 12 is southeastern Siberia. In this very region, ECHAM5 featured delayed snowmelt, related to overly high albedo in the presence of vegetation over snow (Räisänen et al. 2014). While the specific mechanisms leading to too fast snow melt might differ in different models, the example of ECHAM5 highlights the importance of the treatment of surface energy budget in the presence of snow.~~

The parameterization of the surface albedo is another factor that influences the snowmelt rate. In spring, the surface albedo feedback is strong and if the albedo is not correctly simulated in the models, it can cause large biases in SWE in spring

(Thackeray et al., 2018). CMIP5 models have had challenges in simulating correctly the surface albedo on snow-covered land (Thackeray et al., 2015). The intermodel variability in surface albedo feedback is large in CMIP5 models (Qu and Hall, 2014; Thackeray et al., 2018), and the large variability has been found to persist in CMIP6 models (Thackeray et al., 2021), suggesting that there are still large uncertainties in simulating the surface albedo. Specifically, regarding the discussion on ECHAM5 and MPI-ESM1-2-HR above, the only region in which MPI-ESM1-2-HR displays a positive residual in Fig. 12 is southeastern Siberia. In this very region, ECHAM5 featured delayed snowmelt, related to overly high albedo in the presence of vegetation over snow (Räisänen et al., 2014). While the specific mechanisms leading to too fast snow melt might differ in different models, the example of ECHAM5 highlights the importance of the treatment of surface energy budget in the presence of snow.

In addition to issues related to the models, uncertainties in the observational data can also affect the residual values. ~~All the observational datasets have uncertainties that can also affect the results.~~ With the bias-correction method, SWE data is more accurate than before, but the uncertainty in hemisphere-mean values is still 7.4%. In mountainous regions, the complex terrain poses a challenge in the SWE estimation but averaging over multiple products can improve the accuracy (Mortimer et al., 2020). There are also uncertainties associated with the GPCC and MERRA-2 datasets that can cause errors in ~~the differences between models and observations~~ assessing the model biases; for example, MERRA-2 underestimates global warming trends in the last years of our study period compared to other reanalyses (Gelaro et al., 2017; Simmons et al., 2017). Snow cover in spring is especially sensitive to warming (Hernández-Henríquez et al., 2015) and, therefore, the uncertainties in MERRA-2 can affect the results especially in spring.

The uncertainties associated with the GPCC precipitation product can be divided into two categories: the systematic measuring error and stochastic sampling error (Schneider et al., 2014). The systematic error results from evaporation and from droplets and snowflakes being drifted across the gauge funnel. This error almost always causes underestimation of precipitation (Schneider et al., 2014). The sampling error, in turn, is caused by the sparse network density of the in situ stations and the error increases when the density of the network decreases. A correction method taking into account the weather conditions has been implemented in the GPCC to improve the P estimate (Schneider et al., 2014). In general, estimating precipitation in high latitudes is a challenge. The estimates have improved, but especially regional discrepancies still exist between different precipitation products (Behrangi et al., 2016). This causes uncertainties in assessing the model bias and means that the choice of the precipitation product may influence the results.

We additionally studied the dependency between the residual term and fractional forest cover (Fig. S22), and we did this separately for the entire study area and for non-mountainous regions. In the entire study area, the spread of the residual term is quite large when the fraction of forest cover is small in both winter and spring. When looking at only the non-mountainous areas, the spread decreases compared to the entire study area, especially in spring. This illustrates that the residual term is

particularly large in mountainous areas, indicating that in mountainous areas, T or P can only explain a small fraction of SWE bias, but other factors contribute SWE bias considerably. Also, in mountainous areas, the intermodel variability of the residual term is substantial. The complex terrain and large SWE variability make SWE estimates challenging. Also, the original resolution of CMIP6 models can be too coarse for accurately describing SWE in mountainous terrain; biases in snow cover and snow depth have been found to decrease considerably with downscaled regional climate models in mountainous regions (Matiu et al., 2019). Resampling the datasets to a common grid can cause uncertainties in mountainous areas due to the high topographic variability within a grid cell.

For most models, the residual term in non-mountainous areas in winter depends little on the fractional forest cover, indicating that the residual arises from other sources. For the CESM2 and NorESM2-MM models, however, the residual decreases when forest cover fraction increases, and this dependency is strikingly similar for these two models. Since CESM2 and NorESM2-MM both employ the CLM5 land surface model (Lawrence et al., 2019), this hints that the large positive SWE residual in these models originates from CLM5. Also, a unique feature within CESM2 that may cause high residual values in the coldest regions is that the model allows for a very large maximum SWE to enable the simulation of firn production over ice sheets (van Kampenhout et al., 2017). Even though our study does not cover ice sheets, this feature may cause positive SWE bias in the coldest regions of the study area (e.g. in the northernmost regions of Canada).

In spring in non-mountainous areas, the residual of SWE_{change} depends quasi-systematically on the forest cover fraction. While the level of the residual terms varies between the models, the values increase with increasing forest cover fraction in all models. The residual varies mostly between -25 and -12 kg m^{-2} month^{-1} for open areas, and between -13 and -2 kg m^{-2} month^{-1} when the forest cover fraction is close to 100%. This leaves room for at least two possible interpretations. On one hand, the larger intermodel spread of the residual over open areas suggests that problems in snow albedo parameterization (e.g., how to account for snow metamorphism, Colbeck, 1982) could be at least partly responsible for the large residual term. Due to the shading effect of the canopy (Malle et al., 2019), snow albedo treatment differences in the models should manifest strongest over open tundra regions and less so over dense forest cover. On the other hand, the quasi-systematic dependence of the residual on the forest cover fraction could also point to some subtle issues in the SWE estimates of the SnowCCI dataset in spring.

Overall, the reasons behind the residual term are complex and more detailed model-specific investigations for all participating CMIP6 models are beyond the scope of this study. Further analysis would be required in the future to fully understand the factors behind the residual term.

5 Conclusions

635 We have intercompared NH SWE estimates between CMIP6 models and ~~satellite-based SnowCCI data~~observation-based SWE reference ~~, and~~data and studied whether ~~differences~~model biases in precipitation (P) or temperature (T) ~~between models and observations~~could explain the ~~differences in~~SWE model biases. Our study covered land areas north of 40°N and years between 1982 and 2014. We analyzed separately the SWE in winter (in February) and the SWE change rate in spring (SWE_{change} from February to May). Using regression analysis, we divided the ~~difference in~~SWE ~~between~~model ~~and observation~~bias (ΔSWE

640 and ΔSWE_{change}, model minus observation) into three components: the contribution of P, the contribution of T and the contribution of other factors, such as deficiencies in model parameterizations or inaccuracies in the observational datasets. The main findings in our study are as follows:

- The models generally overestimate SWE, but large variability exists between models. The largest overestimations
645 occur mainly in the northernmost parts of both Eurasia and North America. In winter, the overestimated SWE is mainly concentrated in the same areas in every model, but the magnitude differs between the models. In spring, the snow melt rates vary clearly between the models.

- In winter, the ~~differences in~~SWE biases can be explained mostly with ~~differences in~~the P biases. The contribution of T is clearly smaller than that of P. This is in line with the expected results, as even too cold temperatures cannot
650 cause too high SWE without precipitation. However, other factors contribute to SWE discrepancies as well.

- In spring, T and P explain partly the ~~differences between modeled and observed~~biases in SWE_{change}. Especially cold or warm biases often co-occur with large SWE_{change} ~~differences~~biases, but large spatial and inter-model variability exists. The importance of T in explaining SWE_{change} discrepancies during spring is to be expected, because the increase in T is the main factor that causes snow to melt as spring progresses. Yet it should be noted that the contribution of
655 other factors, such as observation uncertainty or deficiencies in model parameterizations, is more significant in spring than in winter.

Overall, the study showed that the models still need to be improved to accurately describe SWE. However, the analysis also showed that there is a link between T and SWE, especially in the warmer regions of the study area, suggesting that climate
660 models may be able to simulate SWE trends in a warming climate correctly, even if SWE itself is not accurately reproduced. Uncertainties in the observational data also cause uncertainties in the analysis, so by improving the observational data, we can study the models' ability to describe the snow cover more reliably and, thus, further improve the models.

Appendix A: The equations for calculating the ~~differences between models and observations~~model biases and the linear regressions

The steps for calculating the ~~differences~~ model biases in SWE, T and P ~~between models and observations~~, and subsequently the linear regressions in winter are as follows:

1. We calculated cumulative T (T_{cum}) and P (P_{cum}) from November to January for each model and for the observational datasets:

$$T_{cum, model}=T_{Nov}+T_{Dec}+T_{Jan} \tag{A1}$$

$$T_{cum, obs}=T_{Nov}+T_{Dec}+T_{Jan} \tag{A2}$$

$$P_{cum, model}=P_{Nov}+P_{Dec}+P_{Jan} \tag{A3}$$

$$P_{cum, obs}=P_{Nov}+P_{Dec}+P_{Jan} \tag{A4}$$

2. We calculated the ~~difference~~ model bias in cumulative T (ΔT_{cum}) and P (ΔP_{cum})~~between each model and observations~~:

$$\Delta T_{cum}=T_{cum, model}-T_{cum, obs} \tag{A5}$$

$$\Delta P_{cum,}=P_{cum, model}-P_{cum, obs} \tag{A6}$$

3. We calculated the ~~difference~~ model bias in SWE (ΔSWE) in February ~~between each model and SnowCCI~~:

$$\Delta SWE=SWE_{model}-SWE_{obs} \tag{A7}$$

4. We calculated the linear regression for the ~~differences~~ biases using the ordinary least squares method:

$$\Delta SWE = \beta_T \Delta T_{cum} + \beta_P \Delta P_{cum} + C \tag{A8}$$

where β_T and β_P are the regression coefficients, and C is the constant.

The steps for calculating the ~~differences~~ biases in SWE_{change}, T and P ~~between models and observations~~, and subsequently the linear regressions in spring are as follows:

1. We calculated monthly change in SWE (SWE_{change}) for each model and for the SWE reference data~~SnowCCI~~:

$$SWE_{change,1}=SWE_{Mar}-SWE_{Feb} \tag{A9}$$

$$SWE_{change,2}=SWE_{Apr}-SWE_{Mar} \tag{A10}$$

$$SWE_{change,3}=SWE_{May}-SWE_{Apr} \tag{A11}$$

2. We calculated the ~~differences~~ model biases in monthly SWE_{change} (ΔSWE_{change})~~between each model and SnowCCI~~:

$$\Delta SWE_{change}=SWE_{change,1,model}-SWE_{change,1,obs} \tag{A12}$$

$$\Delta SWE_{change}=SWE_{change,~~1~~2,model}-SWE_{change,~~1~~2,obs} \tag{A13}$$

$$\Delta SWE_{change} = SWE_{change,13,model} - SWE_{change,13,obs} \qquad (A14)$$

3. We calculated the ~~differences~~ model biases in T (ΔT) and P (ΔP) ~~between each model and the observations~~ in March, April, and May:

$$\Delta T_1 \cancel{\Delta T} = T_{Mar,model} - T_{Mar,obs} \qquad (A15)$$

$$\Delta T_2 \cancel{\Delta T} = T_{Apr,model} - T_{Apr,obs} \qquad (A16)$$

$$\Delta T_3 \cancel{\Delta T} = T_{May,model} - T_{May,obs} \qquad (A17)$$

$$\Delta P_1 \cancel{\Delta P} = P_{Mar,model} - P_{Mar,obs} \qquad (A18)$$

$$\Delta P_2 \cancel{\Delta P} = P_{Apr,model} - P_{Apr,obs} \qquad (A19)$$

$$\Delta P_3 \cancel{\Delta P} = P_{May,model} - P_{May,obs} \qquad (A20)$$

3. We pooled together the values of ΔSWE_{change}, ΔP and ΔT for the whole spring period (February through May) and calculated the linear regression for the ~~differences~~ biases using the ordinary least squares method:

$$\Delta SWE_{change} = \beta_T \Delta T + \beta_P \Delta P + C \qquad (A21)$$

where β_T and β_P are the regression coefficients, and C is the constant.

Data availability

The CMIP6 model data are available at the Earth System Grid Federation (https://esgf-node.llnl.gov/search/cmip6/). The ~~SnowCCI data~~SWE products used in this study are available at Globsnow data archive (https://www.globsnow.info/swe/archive_v3.0/). The original SnowCCI data and the ESA CCI Land Cover time-series v2.0.7 are available at the ESA CCI data portal (https://climate.esa.int/en/odp/). The MERRA-2 data are available at the Modeling and Assimilation Data and Information Services Center (https://disc.gsfc.nasa.gov/). The GPCC data are available at DWD website (https://opendata.dwd.de/climate_environment/GPCC/html/fulldata-monthly_v2018_doi_download.html).

Author contribution

KK performed the analysis and produced the figures with substantial contributions from AR and PR. KL provided the ~~SnowCCI~~SWE reference data. KK wrote the original draft. All authors contributed to manuscript review and editing.

Acknowledgements

The work of all authors has been funded by the Academy of Finland, Decision 309125. The authors would like to thank the two anonymous referees for their constructive comments.

Competing interests

725   The authors declare that they have no conflict of interest.

References

Bader, D. C., Leung, R., Taylor, M., McCoy, R. B.: E3SM-Project E3SM1.0 model output prepared for CMIP6 CMIP historical. Earth System Grid Federation, https://doi.org/10.22033/ESGF/CMIP6.4497, 2019.

730   Barnett, T., Adam, J. and Lettenmaier, D.: Potential impacts of a warming climate on water availability in snow-dominated regions, Nature 438, 303–309, https://doi.org/10.1038/nature04141, 2005.

Behrangi, A., Christensen, M., Richardson, M., Lebsock, M.,Stephens, G., Huffman, G. J., Bolvin, D., Adler, R. F., Gardner, A., Lambrigtsen, B., and Fetzer, E.: Status of high-latitude precipitation estimates from observations and reanalyses, J. 735   Geophys. Res. Atmos., 121, 4468–4486, https://doi.org/10.1002/2015JD024546, 2016.

Bentsen, M., Oliviè, D. J. L., Seland, Ø., Toniazzo, T., Gjermundsen, A., Graff, L. S., Debernard, J. B., Gupta, A. K., He, Y., Kirkevåg, A., Schwinger, J., Tjiputra, J., Aas, K. S., Bethke, I., Fan, Y., Griesfeller, J., Grini, A., Guo, C., Ilicak, M., Karset, I. H. H., Landgren, O. A., Liakka, J., Moseid, K. O., Nummelin, A., Spensberger, C., Tang, H., Zhang, Z., Heinze, C., Iversen, 740   T., Schulz, M.: NCC NorESM2-MM model output prepared for CMIP6 CMIP historical. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.8040, 2019.

Bormann, K. J., Brown, R. D., Derksen, C., and Painter, T. H.: Estimating snow-cover trends from space, Nat. Clim. Change, 8, 924–928. https://doi.org/10.1038/s41558-018-0318-3, 2018.

745

Bosilovich, M. G., Akella, S., Coy, L., Cullather, R., Draper, C., Gelaro, R., Kovach, R., Liu, Q., Molod, A., Norris, P., Wargan, K., Chao, W., Reichle, R., Takacs, L., Vikhliaev, Y., Bloom, S., Collow, A., Firth, S., Labow, G., Partyka, G., Pawson, S., Reale, O., Schubert, S. D., and Suarez, M.: MERRA-2: Initial evaluation of the climate, Technical Report Series on Global Modeling and Data Assimilation, 2015.

750

Boucher, O., Denvil, S., Levavasseur, G., Cozic, A., Caubel, A., Foujols, M.-A., Meurdesoif, Y., Balkanski, Y., Checa-Garcia, R., Hauglustaine, D., Bekki, S., Marchand, M.: IPSL IPSL-CM5A2-INCA model output prepared for CMIP6 CMIP historical. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.13661, 2020.

755 Boucher, O., Denvil, S., Levavasseur, G., Cozic, A., Caubel, A., Foujols, M.-A., Meurdesoif, Y., Cadule, P., Devilliers, M., Ghattas, J., Lebas, N., Lurton, T., Mellul, L., Musat, I., Mignot, J., Cheruy, F.: IPSL IPSL-CM6A-LR model output prepared for CMIP6 CMIP historical. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.5195, 2018.

Boucher, O., Denvil, S., Levavasseur, G., Cozic, A., Caubel, A., Foujols, M.-A., Meurdesoif, Y., Balkanski, Y., Checa-Garcia, 760 R., Hauglustaine, D., Bekki, S., Marchand, M.: IPSL IPSL-CM6A-LR-INCA model output prepared for CMIP6 CMIP historical. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.13601, 2021.

Brown, R. D., Brasnett, B., and Robinson, D.: Gridded North American monthly snow depth and snow water equivalent for GCM evaluation, Atmos.-Ocean, 41, 1–14, https://doi.org/10.3137/ao.410101, 2003.
765

Brown, R. D. and Mote, P. W.: The response of Northern Hemisphere snow cover to a changing climate, J. Climate, 22, 2124-2145, https://doi.org/10.1175/2008JCLI2665.1, 2009.

Brun, E., Vionnet, V., Boone, A., Decharme, B., Peings, Y., Vallette, R., Karbou, F., and Morin, S.: Simulation of northern 770 Eurasian local snow depth, mass, and density using a detailed snowpack model and meteorological reanalyses, J. Hydrometeorol., 14, 203–219, https://doi.org/10.1175/JHM-D-12-012.1, 2013.

Brutel-Vuilmet, C., Menegoz, M., and Krinner, G.: An analysis of present and future seasonal Northern Hemisphere land snow cover simulated by CMIP5 coupled climate models, The Cryosphere, 7, 67-80, https://doi.org/doi:10.5194/tc-7-67-2013, 2013.
775

Cagnati, A., Crepaz, A., Macelloni, G., Pampaloni, P., Ranzi, R., Tedesco, M., Tomirotti, M., and Valt, M.: Study of the snow melt–freeze cycle using multisensory data and snow modeling, J. Glaciol., 50, 419–426. https://doi.org/10.3189/172756504781830006, 2004.

780 Callaghan, T. V., Johansson, M., Brown, R. D., Groisman, P. Y., Labba, N., Radionov, V., Bradley, R. S., Blangy, S., Bulygina, O. N., Christensen, T. R., Colman, J. E., Essery, R. L. H., Forbes,B. C., Forchhammer, M. C., Golubev, V. N., Honrath, R. E., Ju-day, G. P., Meshcherskaya, A. V., Phoenix, G. K., Pomeroy, J.,Rautio, A., Robinson, D. A., Schmidt, N. M., Serreze, M. C.,Shevchenko, V. P., Shiklomanov, A. I., Shmakin, A. B., Sköld,P., Sturm, M., Woo, M.-K., and Wood, E. F.: Multiple effects ofchanges in arctic snow cover, Ambio, 40, 32–45, https://doi.org/10.1007/s13280-011-0213-x, 2011.
785

Cho, E., Jacobs, J. M., & Vuyovich, C. M.: The value of long-term (40 years) airborne gamma radiation SWE record for evaluating three observation-based gridded SWE data sets by seasonal snow and land cover classifications. Water resources research, 56, https://doi.org/10.1029/2019WR025813, 2020.

790 Cohen, J., Ye, H., and Jones, J.: Trends and variability in rain-on-snow events, Geophys. Res. Lett., 42, 7115-7122, https://doi.org/10.1002/2015GL065320., 2015.

Colbeck, S. C.: An overview of seasonal snow metamorphism. Rev. Geophys., 20(1), 45–61. https://doi.org/10.1029/RG020i001p00045, 1982.

795

Danabasoglu, GokhanG.: NCAR CESM2 model output prepared for CMIP6 CMIP historical,- Earth System Grid Federation, https://doi.org/10.22033/ESGF/CMIP6.7627, 2019a.

Danabasoglu, GokhanG.: NCAR CESM2-WACCM model output prepared for CMIP6 CMIP historical, Earth System Grid
800 Federation, https://doi.org/10.22033/ESGF/CMIP6.10071, 2019b.

Danabasoglu, G.: NCAR CESM2-WACCM-FV2 model output prepared for CMIP6 CMIP historical. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.11298, 2019c.

805 Danek, C., Shi, X., Stepanek, C., Yang, H., Barbi, D., Hegewald, J., Lohmann, G.: AWI AWI-ESM1.1LR model output prepared for CMIP6 CMIP historical. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.9328, 2020.

Derksen, C. and Brown, R.: Spring snow cover extent reductions in the 2008–2012 period exceeding climate model projections, Geophys. Res. Lett., 39, L19504, https://doi.org/10.1029/2012GL053387, 2012.

810

Dix, M., Bi, D., Dobrohotoff, P., Fiedler, R., Harman, I., Law, R., Mackallah, C., Marsland, S., O'Farrell, S., Rashid, H., Srbinovsky, J., Sullivan, A., Trenham, C., Vohralik, P., Watterson, I., Williams, G., Woodhouse, M., Bodman, R., Dias, F. B., Domingues, C., Hannah, N., Heerdegen, A., Savita, A., Wales, S., Allen, C., Druken, K., Evans, B., Richards, C., Ridzwan, S. M., Roberts, D., Smillie, J., Snow, K., Ward, M., Yang, R.: CSIRO-ARCCSS ACCESS-CM2 model output prepared for
815 CMIP6 CMIP historical. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.4271, 2019.

Dong, B., Lenters, J. D., Hu, Q., Kucharik, C. J., Wang, T., Soylu, M. E., and Mykleby, P. M.: Decadal-Scale Changes in the Seasonal Surface Water Balance of the Central United States from 1984 to 2007, J. Hydrometeorol., 21, 1905-1927, https://doi.org/10.1175/JHM-D-19-0050.1, 2020.

820

Douville, H., Chauvin, F., Planton, S., Royer, J. F., Salas-Melia, D., and Tyteca, S.: Sensitivity of the hydrological cycle to increasing amounts of greenhouse gases and aerosols, Clim. Dynam., 20, 45–68. https://doi.org/10.1007/s00382-002-0259-3, 2002.

825   Draper, C. S., Reichle, R. H., and Koster, R. D.: Assessment of MERRA-2 land surface energy flux estimates, J. Climate, 31, 671-691, https://doi.org/10.1175/JCLI-D-17-0121.1, 2018.

EC-Earth Consortium (EC-Earth): EC-Earth-Consortium EC-Earth3 model output prepared for CMIP6 CMIP historical, Earth System Grid Federation, https://doi.org/10.22033/ESGF/CMIP6.4700, 2019a.

830

EC-Earth Consortium (EC-Earth): EC-Earth-Consortium EC-Earth3-Veg model output prepared for CMIP6 CMIP historical, Earth System Grid Federation, https://doi.org/10.22033/ESGF/CMIP6.4706, 2019b.

EC-Earth Consortium (EC-Earth): EC-Earth-Consortium EC-Earth3-AerChem model output prepared for CMIP6 CMIP

835   historical. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.4701, 2020a.

EC-Earth Consortium (EC-Earth): EC-Earth-Consortium EC-Earth-3-CC model output prepared for CMIP6 CMIP historical. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.4702, 2021.

840   EC-Earth Consortium (EC-Earth): EC-Earth-Consortium EC-Earth3-Veg-LR model output prepared for CMIP6 CMIP historical. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.4707, 2020b

ESA: Land Cover CCI Product User Guide Version 2. Tech. Rep., Available at: maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-Ph2-PUGv2_2.0.pdf, 2017.

845

Essery, R., Kim, H., Wang, L., Bartlett, P., Boone, A., Brutel-Vuilmet, C., Burke, E., Cuntz, M., Decharme, B., Dutra, E., Fang, X., Gusev, Y.,Hagemann, S., Haverd, V., Kontu, A., Krinner, G., Lafaysse, M., Lejeune, Y., Marke, T., Marks, D., Marty, C., Menard, C. B., Nasonova, O., Nitta, T., Pomeroy, J., Schädler, G., Semenov, V., Smirnova, T., Swenson, S., Turkov, D., Wever, N., and Yuan, H.: Snow cover duration trends observed at sites and predicted by multiple models, The

850   Cryosphere, 14, 4687–4698, https://doi.org/10.5194/tc-14-4687-2020, 2020.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Model Dev., 9, 1937-1958, https://doi.org/10.5194/gmd-9-1937-2016, 2016.

855   Fierz, C., Armstrong, R. L., Durand, Y., Etchevers, P., Greene, E., McClung, D. M., Nishimura, K., Satyawali, P.K. and Sokratov, S.A.: The International Classification for Seasonal Snow on the Ground,– IHP-VII Technical Documents in Hydrology No. 83, IACS Contribution No. 1, UNESCO-IHP, 2009.

860   Flanner, M. G., Shell, K. M., Barlage, M., Perovich, D. K., and Tschudi, M. A.: Radiative forcing and albedo feedback from the Northern Hemisphere cryosphere between 1979 and 2008, Nat. Geosci., 4, 151-155, https://doi.org/10.1038/ngeo1062, 2011.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G.,
865   Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A. M., Gu, W., Kim, G.-K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J. E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S. D., Sienkiewicz, M., and Zhao, B.: The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). J. Climate, 30, 5419-5454, https://doi.org/10.1175/JCLI-D-16-0758.1, 2017.

870   Global Modeling and Assimilation Office (GMAO): MERRA-2 tavg1_2d_lnd_Nx: 2d, 1-Hourly, Time-Averaged, Single-Level, Assimilation, Land Surface Diagnostics V5.12.4, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), https://doi.org/10.5067/RKPHT8KC1Y1T, 2015a.

Global Modeling and Assimilation Office (GMAO): MERRA-2 tavgM_2d_slv_Nx: 2d, Monthly mean, Time-Averaged,
875   Single-Level, Assimilation, Single-Level Diagnostics V5.12.4, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed 7 April 2020, https://doi.org/10.5067/AP1B0BA5PD2K, 2015b.

Guo, H., John, Jasmin G, Blanton, C., McHugh, C., Nikonov, S., Radhakrishnan, A., Rand, K., Zadeh, N. T., Balaji, V, Durachta, J., Dupuis, C., Menzel, R., Robinson, T., Underwood, S., Vahlenkamp, H., Bushuk, M., Dunne, K. A., Dussin, R.,
880   Gauthier, P. P. G., Ginoux, P., Griffies, S. M., Hallberg, R., Harrison, M., Hurlin, W., Lin, P., Malyshev, S., Naik, V., Paulot, F., Paynter, D. J., Ploshay, J., Reichl, B. G., Schwarzkopf, D. M., Seman, C. J., Shao, A., Silvers, L., Wyman, B., Yan, X., Zeng, Y., Adcroft, A., Dunne, J. P., Held, I. M., Krasting, J. P., Horowitz, L. W., Milly, P. C. D, Shevliakova, E., Winton, M., Zhao, M., and Zhang, R.: NOAA-GFDL GFDL-CM4 model output historical, Earth System Grid Federation, https://doi.org/10.22033/ESGF/CMIP6.8594, 2018.
885

Hall, N. D., Stuntz, B. B., and Abrams, R. H.: Climate change and freshwater resources, Natural Resources & Environment, 22, 30-35, 2008.

Hall, A., and Qu, X.: Using the current seasonal cycle to constrain snow albedo feedback in future climate change, Geophys. Res. Lett., 33, L03502. https://doi.org/10.1029/2005GL025127, 2006.

Henderson, G. R., Peings, Y., Furtado, J. C., and Kushner, P. J.: Snow–atmosphere coupling in the Northern Hemisphere, Nat. Clim. Change, 8, 954–963, https://doi.org/10.1038/s41558-018-0295-6, 2018.

Hernández-Henríquez, M. A., Déry, S. J., and Derksen, C.: Polar amplification and elevation-dependence in trends of Northern Hemisphere snow cover extent, 1971–2014. Environ. Res. Lett., 10, 044010. https://doi.org/10.1088/1748-9326/10/4/044010, 2015.

Huning, L. S., and AghaKouchak, A.: Global snow drought hot spots and characteristics. Proceedings of the National Academy of Sciences, 117, 19753-19759, https://doi.org/10.1073/pnas.1915921117, 2020.

IPCC: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, UK and New York, NY, USA, 2013.

Jungclaus, J., Bittner, M., Wieners, K.-H., Wachsmann, F., Schupfner, M., Legutke, S., Giorgetta, M., Reick, C., Gayler, V., Haak, H., de Vrese, P., Raddatz, T., Esch, M., Mauritsen, T., von Storch, J.-S., Behrens, J., Brovkin, V., Claussen, M., Crueger, T., Fast, I., Fiedler, S., Hagemann, S., Hohenegger, C., Jahns, T., Kloster, S., Kinne, S., Lasslop, G., Kornblueh, L., Marotzke, J., Matei, D., Meraner, K., Mikolajewicz, U., Modali, K., Müller, W., Nabel, J., Notz, D., Peters, K., Pincus, R., Pohlmann, H., Pongratz, J., Rast, S., Schmidt, H., Schnur, R., Schulzweida, U., Six, K., Stevens, B., Voigt, A., and Roeckner, E.: MPI-M MPI-ESM1.2-HR model output prepared for CMIP6 CMIP historical, Earth System Grid Federation, https://doi.org/10.22033/ESGF/CMIP6.6594, 2019.

Kelly, R. E., Chang, A. T., Tsang, L., and Foster, J. L.: A prototype AMSR-E global snow area and snow depth algorithm, IEEE T. Geosci. Remote, 41, 230–242. https://doi.org/10.1109/TGRS.2003.809118, 2003.

Kim, R. S., Kumar, S., Vuyovich, C., Houser, P., Lundquist, J., Mudryk, L., Durand, M., Barros, A., Kim, E. J., Forman, B. A., Gutmann, E. D., Wrzesien, M. L., Garnaud, C., Sandells, M., Marshall, H.-P., Cristea, N., Pflug, J. M., Johnston, J., Cao, Y., Mocko, D., and Wang, S.: Snow Ensemble Uncertainty Project (SEUP): quantification of snow water equivalent uncertainty across North America via ensemble land surface modeling, The Cryosphere, 15, 771–791, https://doi.org/10.5194/tc-15-771-2021, 2021.

Krasting, J. P.; John, J. G; Blanton, C.; McHugh, C.; Nikonov, S.; Radhakrishnan, A.; Rand, K.; Zadeh, N. T.; Balaji, V;
Durachta, J.; Dupuis, C.; Menzel, R.; Robinson, T.; Underwood, S.; Vahlenkamp, H.; Dunne, K. A.; Gauthier, P. P. G.; Ginoux,
P.; Griffies, S. M.; Hallberg, R.; Harrison, M.; Hurlin, W.; Malyshev, S.; Naik, V.; Paulot, F.; Paynter, D. J; Ploshay, J.; Reichl,
B. G.; Schwarzkopf, D. M.; Seman, C. J.; Silvers, L.; Wyman, B.; Zeng, Y.; Adcroft, A.; Dunne, J. P.; Dussin, R.; Guo, H.;
He, J.; Held, I. M.; Horowitz, L. W.; Lin, P.; Milly, P. C. D; Shevliakova, E.; Stock, C.; Winton, M.; Wittenberg, A. T.; Xie,
Y.; and Zhao, M.: NOAA-GFDL GFDL-ESM4 model output prepared for CMIP6 CMIP historical, Earth System Grid
Federation, https://doi.org/10.22033/ESGF/CMIP6.8597, 2018a.

Krasting, J. P.; John, J. G; Blanton, C.; McHugh, C.; Nikonov, S.; Radhakrishnan, A.; Rand, K.; Zadeh, N. T.; Balaji, V;
Durachta, J.; Dupuis, C.; Menzel, R.; Robinson, T.; Underwood, S.; Vahlenkamp, H.; Dunne, K. A.; Gauthier, P. P. G.; Ginoux,
P.; Griffies, S. M.; Hallberg, R.; Harrison, M.; Hurlin, W.; Malyshev, S.; Naik, V.; Paulot, F.; Paynter, D. J; Ploshay, J.; Reichl,
B. G.; Schwarzkopf, D. M.; Seman, C. J.; Silvers, L.; Wyman, B.; Zeng, Y.; Adcroft, A.; Dunne, J. P.; Dussin, R.; Guo, H.;
He, J.; Held, I. M.; Horowitz, L. W.; Lin, P.; Milly, P. C. D; Shevliakova, E.; Stock, C.; Winton, M.; Wittenberg, A. T.; Xie,
Y.; and Zhao, M.: NOAA-GFDL GFDL-ESM4 model output prepared for CMIP6 CMIP esm-hist, Earth System Grid
Federation, https://doi.org/10.22033/ESGF/CMIP6.8597, 2018b.

Kundzewicz, Z. W., Mata, L. J., Arnell, N. W., Döll, P., Jimenez, B., Miller, K., Oki, T., Şen, Z, and Shiklomanov, I.: The
implications of projected climate change for freshwater resources and their management, Hydrolog. Sci. J., 53, 3-10
https://doi.org/10.1623/hysj.53.1.3, 2008.

Kunkel, K. E., Robinson, D. A., Champion, S., Yin, X., Estilow, T., and Frankson, R. M.: Trends and extremes in northern
hemisphere snow characteristics, Current Climate Change Reports, 2, 65–73, https://doi.org/10.1007/s40641-016-0036-8,
2016.

Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., Collier, N., Ghimire, B., van
Kampenhout, L., Kennedy, D., Kluzek, E.,Lawrence, P. J.,Li, F., Li, H.,Lombardozzi, D.,Riley, W. J., Sacks, W. J.,Shi, M.
Vertenstein, M., Wieder, W. R. Xu, C.,Ali, A. A.,Badger, A. M., Bisht, G., van den Broeke, M., Brunke, M. A., Burns, S. P.,
Buzan, J., Clark, M., Craig, A., Dahlin, K., Drewniak, B., Fisher, J. B., Flanner, M., Fox, A. M., Gentine, P., Hoffman, F.,
Keppel-Aleks, G., Knox, R., Kumar, S., Lenaerts, J., Leung, L. R., Lipscomb, W. H., Lu, Y., Pandey, A., Pelletier, J. D.,
Perket, J., Randerson, J. T., Ricciuto, D. M., Sanderson, B. M., Slater, A., Subin, Z. M., Tang, J., Thomas, R. Q., Val Martin,
M., and Zeng, X.: The Community Land Model version 5: Description of new features, benchmarking, and impact of forcing
uncertainty. , J. Adv. Model Earth Sy, 11, 4245– 4287. https://doi.org/10.1029/2018MS001583, 2019.

Lee, W.-L., Liang, H.-C.: AS-RCEC TaiESM1.0 model output prepared for CMIP6 CMIP historical. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.9755, 2020.

960

Loth, B., Graf, H. F., and Oberhuber, J. M.: Snow cover model for global climate simulations, J. Geophys. Res. Atmos., 98, 10451-10464, https://doi.org/10.1029/93JD00324, 1993.

Luojus, K., Pulliainen, J., Takala, M., Lemmetyinen, J., Moisander, M., Mortimer, C., Derksen, C., Hiltunen, M., Smolander,
965   T., Ikonen, J., Cohen, J., Veijola, K., and Venäläinen, P.: GlobSnow v3.0 Northern Hemisphere snow water equivalent dataset, Sci. Data, https://doi.org/10.1038/s41597-021-00939-2, 2021.

Magnusson, J., Nævdal, G., Matt, F., Burkhart, J. F., and Winstral, A.: Improving hydropower inflow forecasts by assimilating snow data, Hydrol. Res., 51, 226-237, https://doi.org/10.2166/nh.2020.025, 2020.

970

Malle, J., Rutter, N., Mazzotti, G., and Jonas, T.: Shading by trees and fractional snow cover control the subcanopy radiation budget. J. Geophys. Res.-Atmos., 124, 3195–3207. https://doi.org/10.1029/2018JD029908, 2019.

Matiu, M., Petitta, M., Notarnicola, C., and Zebisch, M.: Evaluating Snow in EURO-CORDEX Regional Climate Models with
975   Observations for the European Alps: Biases and Their Relationship to Orography, Temperature, and Precipitation Mismatches, Atmosphere, 11(1), 46, https://doi.org/10.3390/atmos11010046, 2020.

Mortimer, C., Mudryk, L., Derksen, C., Luojus, K., Brown, R., Kelly, R., and Tedesco, M.: Evaluation of long-term Northern Hemisphere snow water equivalent products, The Cryosphere, 14, 1579-1594, https://doi.org/10.5194/tc-14-1579-2020, 2020.

980

Mudryk, L. R., Derksen, C., Kushner, P. J., and Brown, R.: Characterization of Northern Hemisphere snow water equivalent datasets, 1981–2010, J. Climate, 28, 8037-8051, https://doi.org/10.1175/JCLI-D-15-0229.1, 2015.

Mudryk, L. R., Kushner, P. J., Derksen, C., and Thackeray, C.: Snow cover response to temperature in observational and
985   climate model ensembles. Geophys. Res. Lett., 44, 919–926. https://doi.org/10.1002/2016GL071789, 2017.

Mudryk, L., Santolaria-Otín, M., Krinner, G., Ménégoz, M., Derksen, C., Brutel-Vuilmet, C., Brady, M., and Essery, R. (2020). Historical Northern Hemisphere snow cover trends and projected changes in the CMIP6 multi-model ensemble. The Cryosphere, 14, 2495-2514, https://doi.org/10.5194/tc-14-2495-2020, 2020.

990

NASA Goddard Institute for Space Studies (NASA/GISS): NASA-GISS GISS-E2.1G model output prepared for CMIP6 CMIP historical. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.7127, 2018.

995   NASA Goddard Institute for Space Studies (NASA/GISS): NASA-GISS GISS-E2-1-G-CC model output prepared for CMIP6 CMIP esm-hist. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.11723, 2019a.

NASA Goddard Institute for Space Studies (NASA/GISS): NASA-GISS GISS-E2.1H model output prepared for CMIP6 CMIP historical. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.7128, 2019b.

1000   Neubauer, D., Ferrachat, S., Siegenthaler-Le Drian, C., Stoll, J., Folini, D. S., Tegen, I., Wieners, K.-H., Mauritsen, T., Stemmler, I., Barthel, S., Bey, I., Daskalakis, N., Heinold, B., Kokkola, H., Partridge, D., Rast, S., Schmidt, H., Schutgens, N., Stanelle, T., Stier, P., Watson-Parris, D., Lohmann, U.: HAMMOZ-Consortium MPI-ESM1.2-HAM model output prepared for CMIP6 CMIP historical. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.5016, 2019.

1005   Park, H., Yabuki, H., and Ohata, T.: Analysis of satellite and model datasets for variability and trends in Arctic snow extent and depth, 1948–2006, Polar Sci., 6, 23-37, https://doi.org/10.1016/j.polar.2011.11.002, 2012.

Park, S., and Shin, J.: SNU SAM0-UNICON model output prepared for CMIP6 CMIP historical, Earth System Grid Federation https://doi.org/10.22033/ESGF/CMIP6.7789, 2019.

1010

Pulliainen, J., Luojus, K., Derksen, C., Mudryk, L., Lemmetyinen, J., Salminen, M., Ikonen, J., Takala, M., Cohen, J., Smolander, T., and Norberg, J.: Patterns and trends of Northern Hemisphere snow mass from 1980 to 2018, Nature, 581, 294-298, https://doi.org/10.1038/s41586-020-2258-0, 2020.

1015   Qu, X. and Hall, A.: Surface contribution to planetary albedo variability in cryosphere regions, J. Climate, 18, 5239–5252. https://doi.org/10.1175/JCLI3555.1, 2005.

Qu, X., and Hall, A.: On the persistent spread in snow-albedo feedback. Climate Dynamics, 42(1-2), 69–81. https://doi.org/10.1007/s00382-013-1774-0, 2014.

1020

Räisänen, J.: Warmer climate: less or more snow?, Clim. Dyn., 30, 307-319 https://doi.org/10.1007/s00382-007-0289-y, 2008.

Räisänen, P., Luomaranta, A., Järvinen, H., Takala, M., Jylhä, K., Bulygina, O. N., Luojus, K., Riihelä., A., Koskinen, J., and Pulliainen, J.: Evaluation of North Eurasian snow-off dates in the ECHAM5.4 atmospheric general circulation model, Geosci. Model Dev., 7, 3037-3057, https://doi.org/10.5194/gmd-7-3037-2014, 2014.

Santolaria-Otín, M. and Zolina, O.: Evaluation of snow cover and snow water equivalent in the continental Arctic in CMIP5 models, Clim. Dyn., 55, 2993-3016, https://doi.org/10.1007/s00382-020-05434-9, 2020.

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Ziese, M, and Rudolf, B.: GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. Theor Appl Climatol 115, 15–40, https://doi.org/10.1007/s00704-013-0860-x, 2014.

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., and Ziese, M.: GPCC Full Data Monthly Product Version 2018 at 0.5°: Monthly Land-Surface Precipitation from Rain-Gauges built on GTS-based and Historical Data, https://doi.org/10.5676/DWD_GPCC/FD_M_V2018_050, 2018

Seland, Ø., Bentsen, M., Oliviè, D. J. L., Toniazzo, T., Gjermundsen, A., Graff, L. S., Debernard, J. B., Gupta, A. K., He, Y., Kirkevåg, A., Schwinger, J., Tjiputra, J., Aas, K. S., Bethke, I., Fan, Y., Griesfeller, J., Grini, A., Guo, C., Ilicak, M., Karset, I. H. H., Landgren, O. A., Liakka, J., Moseid, K. O., Nummelin, A., Spensberger, C., Tang, H., Zhang, Z., Heinze, C., Iversen, T., Schulz, M.: NCC NorESM2-LM model output prepared for CMIP6 CMIP historical. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.8036, 2019a.

Seland, Ø., Bentsen, M., Oliviè, D. J. L., Toniazzo, T., Gjermundsen, A., Graff, L. S., Debernard, J. B., Gupta, A. K., He, Y., Kirkevåg, A., Schwinger, J., Tjiputra, J., Aas, K. S., Bethke, I., Fan, Y., Griesfeller, J., Grini, A., Guo, C., Ilicak, M., Karset, I. H. H., Landgren, O. A., Liakka, J., Moseid, K. O., Nummelin, A., Spensberger, C., Tang, H., Zhang, Z., Heinze, C., Iversen, T., Schulz, M.: NCC NorESM2-LM model output prepared for CMIP6 CMIP esm-hist. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.7924, 2019b.

Simmons, A. J., Berrisford, P., Dee, D. P., Hersbach, H., Hirahara, S., and Thépaut, J. N.: A reassessment of temperature variations and trends from global reanalyses and monthly surface climatological datasets, Q. J. Roy. Meteor. Soc., 143, 101-119. https://doi.org/10.1002/qj.2949, 2017.

Stuefer, S., Kane, D. L., and Liston, G. E.: In situ snow water equivalent observations in the US Arctic. Hydrol. Res., 44, 21-34, https://doi.org/10.2166/nh.2012.177, 2013.

Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., and Hsu, K.-L.: A review of global precipitation data sets: Data sources, estimation, and intercomparisons, Rev. Geophys., 56, 79–107. https://doi.org/10.1002/2017RG000574, 2018.

1060 Swart, N. C., Cole, Jason N.S., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., Anstey, J., Arora, V., Christian, J. R., Jiao, Y., Lee, W. G., Majaess, F., Saenko, O. A., Seiler, C., Seinen, C., Shao, A., Solheim, L., von Salzen, K., Yang, D., Winter, B., Sigmond, M.: CCCma CanESM5 model output prepared for CMIP6 CMIP historical. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.3610, 2019a.

1065 Swart, N. C., Cole, Jason N.S., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., Anstey, J., Arora, V., Christian, J. R., Jiao, Y., Lee, W. G., Majaess, F., Saenko, O. A., Seiler, C., Seinen, C., Shao, A., Solheim, L., von Salzen, K., Yang, D., Winter, B., Sigmond, M.: CCCma CanESM5 model output prepared for CMIP6 CMIP esm-hist. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.3583, 2019b.

1070 Tatebe, H. and Watanabe, M.: MIROC MIROC6 model output prepared for CMIP6 CMIP historical. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.5603, 2018.

Thackeray, C. W., Fletcher, C. G., and Derksen, C.: Quantifying the skill of CMIP5 models in simulating seasonal albedo and snow cover evolution. J. Geophys. Res.-Atmos., 120(12), 5831-5849. https://doi.org/10.1002/2015JD023325, 2015.
1075
Thackeray, C. W., Fletcher, C. G., Mudryk, L. R., and Derksen, C.: Quantifying the uncertainty in historical and future simulations of Northern Hemisphere spring snow cover, J. Climate, 29, 8647-8663, https://doi.org/10.1175/JCLI-D-16-0341.1, 2016.

1080 Thackeray, C. W., Qu, X., and Hall, A.: Why do models produce spread in snow albedo feedback?. Geophys. Res. Lett., 45(12), 6223-6231. https://doi.org/10.1029/2018GL078493, 2018.

Thackeray, C. W., Hall, A., Zelinka, M. D., and Fletcher, C. G.: Assessing prior emergent constraints on surface albedo feedback in CMIP6. J. Climate, 34(10), 3889-3905, https://doi.org/10.1175/JCLI-D-20-0703.1, 2021.
1085
Trenberth, K. E. and Fasullo, J. T.: Global warming due to increasing absorbed solar radiation, Geophys. Res. Lett., 36, L07706. https://doi.org/10.1029/2009GL03752, 2009.

van Kampenhout, L., Lenaerts, J. T. M., Lipscomb, W. H., Sacks, W. J., Lawrence, D. M., Slater, A. G., and van den Broeke, M. R.: Improving the representation of polar snow and firn in the Community Earth System Model, J. Adv. Model Earth Sy., 9, 2583-2600. https://doi.org/10.1002/2017MS000988, 2017.

Wieners, K.-H., Giorgetta, M., Jungclaus, J., Reick, C., Esch, M., Bittner, M., Legutke, S., Schupfner, M., Wachsmann, F., Gayler, V., Haak, H., de Vrese, P., Raddatz, T., Mauritsen, T., von Storch, J.-S., Behrens, J., Brovkin, V., Claussen, M., Crueger, T., Fast, I., Fiedler, S., Hagemann, S., Hohenegger, C., Jahns, T., Kloster, S., Kinne, S., Lasslop, G., Kornblueh, L., Marotzke, J., Matei, D., Meraner, K., Mikolajewicz, U., Modali, K., Müller, W., Nabel, J., Notz, D., Peters-von Gehlen, K., Pincus, R., Pohlmann, H., Pongratz, J., Rast, S., Schmidt, H., Schnur, R., Schulzweida, U., Six, K., Stevens, B., Voigt, A., Roeckner, E.: MPI-M MPI-ESM1.2-LR model output prepared for CMIP6 CMIP historical. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.6595, 2019a.

Wieners, K.-H., Giorgetta, M., Jungclaus, J., Reick, C., Esch, M., Bittner, M., Legutke, S., Schupfner, M., Wachsmann, F., Gayler, V., Haak, H., de Vrese, P., Raddatz, T., Mauritsen, T., von Storch, J.-S., Behrens, J., Brovkin, V., Claussen, M., Crueger, T., Fast, I., Fiedler, S., Hagemann, S., Hohenegger, C., Jahns, T., Kloster, S., Kinne, S., Lasslop, G., Kornblueh, L., Marotzke, J., Matei, D., Meraner, K., Mikolajewicz, U., Modali, K., Müller, W., Nabel, J., Notz, D., Peters-von Gehlen, K., Pincus, R., Pohlmann, H., Pongratz, J., Rast, S., Schmidt, H., Schnur, R., Schulzweida, U., Six, K., Stevens, B., Voigt, A., Roeckner, E.: MPI-M MPI-ESM1.2-LR model output prepared for CMIP6 CMIP esm-hist. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.6545, 2019b

Wrzesien, M. L., Durand, M. T., Pavelsky, T. M., Kapnick, S. B., Zhang, Y., Guo, J., & Shum, C. K.: A new estimate of North American mountain snow accumulation from regional climate model simulations, Geophys. Res. Lett., 45, 1423–1432, https://doi.org/10.1002/2017GL076664, 2018.

Wu, T.; Chu, M.; Dong, M.; Fang, Y.; Jie, W.; Li, J.; Li, W.; Liu, Q.; Shi, X.; Xin, X.; Yan, J.; Zhang, F.; Zhang, J.; Zhang, L.; and Zhang, Y.: BCC BCC-CSM2MR model output prepared for CMIP6 CMIP historical, Earth System Grid Federation, https://doi.org/10.22033/ESGF/CMIP6.2948, 2018a.

Wu, T.; Chu, M.; Dong, M.; Fang, Y.; Jie, W.; Li, J.; Li, W.; Liu, Q.; Shi, X.; Xin, X.; Yan, J.; Zhang, F.; Zhang, J.; Zhang, L.; and Zhang, Y.: BCC BCC-CSM2MR model output prepared for CMIP6 CMIP esm-hist, Earth System Grid Federation, https://doi.org/10.22033/ESGF/CMIP6.2901, 2018b.

Yukimoto, S.; Koshiro, T.; Kawai, H.; Oshima, N.; Yoshida, K.; Urakawa, S.; Tsujino, H.; Deushi, M.; Tanaka, T.; Hosaka, M.; Yoshimura, H.; Shindo, E.; Mizuta, R.; Ishii, M.; Obata, A.; Adachi, Y.: MRI MRI-ESM2.0 model output prepared for CMIP6 CMIP historical, Earth System Grid Federation, https://doi.org/10.22033/ESGF/CMIP6.6842, 2019.

1125 Zamani, Y., Hashemi Monfared, S., Azhdari moghaddam, M., and Hamidianpour, M.: A comparison of CMIP6 and CMIP5 projections for precipitation to observational data: the case of Northeastern Iran. Theor. Appl. Climatol., 142, 1613–1623, https://doi.org/10.1007/s00704-020-03406-x, 2020.

Ziehn, T., Chamberlain, M., Lenton, A., Law, R., Bodman, R., Dix, M., Wang, Y., Dobrohotoff, P., Srbinovsky, J., Stevens, L., Vohralik, P., Mackallah, C., Sullivan, A., O'Farrell, S., Druken, K.: CSIRO ACCESS-ESM1.5 model output prepared for
1130 CMIP6 CMIP historical. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.4272, 2019a

Ziehn, T., Chamberlain, M., Lenton, A., Law, R., Bodman, R., Dix, M., Wang, Y., Dobrohotoff, P., Srbinovsky, J., Stevens, L., Vohralik, P., Mackallah, C., Sullivan, A., O'Farrell, S., Druken, K.: CSIRO ACCESS-ESM1.5 model output prepared for
1135 CMIP6 CMIP esm-hist. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.4242, 2019b.

Zhang, J., Wu, T., Shi, X., Zhang, F., Li, J., Chu, M., Liu, Q., Yan, J., Ma, Q., Wei, M.: BCC BCC-ESM1 model output prepared for CMIP6 CMIP historical. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.2949, 2018.

1140 Zhu, X., Lee, S. Y., Wen, X., Wei, Z., Ji, Z., Zheng, Z., and Dong, W.: Historical evolution and future trend of Northern Hemisphere snow cover in CMIP5 and CMIP6 models, Environ. Res. Lett., 16, 065013, https://doi.org/10.1088/1748-9326/ac0662, 2021.