Impact of topography and meteorological forcing on snow simulation in the Canadian Land Surface Scheme Including Biogeochemical Cycles (CLASSIC)

Libo Wang[1], Lawrence Mudryk[1], Joe R. Melton[2], Colleen Mortimer[1], Jason Cole[3], Gesa Meyer[2], Paul Bartlett[1], and Mickaël Lalande[4,5]

1 Climate Processes Section, Climate Research Division, Environment and Climate Change Canada, Toronto, ON, Canada

2 Climate Processes Section, Climate Research Division, Environment and Climate Change Canada, Victoria, BC, Canada

3 Canadian Centre for Climate Modelling and Analysis, Climate Research Division, Environment Canada, Victoria, BC, Canada

4 Centre for Research on Watershed-Aquatic Ecosystem Interactions, Environmental Sciences Department, Université du Québec à Trois-Rivières, Trois-Rivières, QC, Canada

5 Centre for Northern Studies, Université Laval, QC, Canada

Corresponding author: Libo Wang, libo.wang@ec.gc.ca

Abstract

Our study evaluates the impacts of an alternate snow cover fraction (SCF) parameterization on snow simulation in the Canadian Land Surface Scheme Including Biogeochemical Cycles (CLASSIC). Three reanalysis-based meteorological datasets are used to drive the model to account for uncertainties in the forcing data. While the default parameterization assumes a simple linear relationship between SCF and snow depth with no dependence on topography, the alternate parameterization accounts for the topographic effects of sub-grid terrain on SCF. We show that the alternate parameterization improves SCF simulated in CLASSIC during winter and spring in mountainous areas for all three choices of meteorological datasets. Annual mean bias, unbiased root mean squared area, and correlation improve by 75 %, 32 %, and 7 % when evaluated with MODIS SCF observations over the Northern Hemisphere. We also demonstrate that the improvements to simulated SCF lead to further improvements in variables related to surface radiation, energy fluxes, and the water cycle. Finally, we link relative biases in the meteorological forcing data to differences in simulated snow water equivalent and SCF. Assessment of simulations with different combinations of SCF parameterizations and meteorological datasets reveals the large impact of meteorological forcing on snow simulation in CLASSIC. Two out of the three meteorological datasets were bias-adjusted using observation-based datasets. However, simulations forced by the dataset without bias correction outperform relative to simulations forced by datasets with bias correction, suggesting that there are large uncertainties in the observation-based datasets and/or methods used for bias correction. This study underscores the importance of accounting for topographic effects of sub-grid terrain and accurate meteorological forcing on snow simulation in land surface models.

1. Introduction

Snow cover exists from six to nine months of the year at the high latitudes and high elevations of mountainous regions. The seasonal transition from snow covered to snow free conditions can have a large impact on the stability of permafrost, the length of the active growing season, and surface water and energy balances due to the much higher albedo of snow cover than other land surfaces (e.g., Myneni et al., 1997; Betts et al., 1998; Osterkamp and Romanovsky, 1999; Frolking et al., 2006). Snow cover plays an important role in the regional and global climate system because of the snow-albedo feedback mechanism (Fletcher et al., 2009; Qu and Hall, 2013). Any uncertainty in the magnitude of this climate feedback decreases our ability to reduce uncertainty in climate sensitivity (Roe and Baker, 2007). Therefore, accurate simulation of snow cover is crucial for future climate predictions in climate and Earth system models (ESMs).

In principle, snow depth (SND) should vary considerably at sub-grid scales of global climate models as a result of multiple heterogeneities in land cover, terrain, and meteorological conditions (Liston 2004). ~~Snow depth (SND) varies at scales from about 10 to 100 m as a result of heterogeneity in land cover, terrain, and meteorological conditions (Liston 2004), while a typical ESM grid cell size is about 100 km or more.~~ Most land surface models (LSMs) explicitly treat only some of this heterogeneity, for example by accounting for different land cover types

2

72 within a grid cell (Verseghy et al., 2017). Snow cover fraction (SCF) parameterizations are
73 commonly used to account for unresolved (sub-grid scale) snow depth variability. However, most
74 models from the Coupled Model Intercomparison Project (CMIP) phase 5 (Taylor et al., 2012)
75 and phase 6 (Eyring et al., 2016) have been found to overestimate SCF in mountainous regions,
76 often with a corresponding cold bias in surface air temperature (Su et al., 2013; Lalande et al.,
77 2021). These biases are also present in the most recent Canadian Earth System Models
78 (CanESM5, Swart et al., 2019; Sigmond et al., 2023) and the latest version of its land surface
79 component, the Canadian Land Surface Scheme Including biogeochemical Cycles (CLASSIC,
80 Melton et al., 2020; Seiler et al., 2021). The SCF overestimation has been attributed to many
81 potential causes, such as too much precipitation and/or overly simplistic SCF parameterizations
82 in ESMs (Lalande et al., 2021; Miao et al., 2022).

83 Some early SCF parameterizations assumed a linear increase in snow cover with snow depth or
84 snow water equivalent (SWE), reaching 100% SCF once a specified threshold was met (e.g.,
85 Verseghy, 1991; Bonan, 1996). Other approaches incorporated surface roughness length into the
86 SCF–SND (or SWE) relationships (e.g., Dickinson et al., 1986; Marshall and Oglesby, 1994),
87 and distinguished SCF estimates between bare ground and vegetated areas (Douville et al., 1995;
88 Yang et al., 1997). Large uncertainties in modeled SCF from these early schemes motivated
89 efforts to refine parameterizations by accounting for terrain heterogeneity or incorporating sub-
90 grid snow distribution (Roesch et al., 2001; Liston, 2004). More recent SCF parameterizations
91 have included snow density (e.g., Niu and Yang, 2007; Lalande et al., 2023) and land cover type
92 (e.g., He et al., 2023), with some schemes adopting separate formulations for snow accumulation
93 and melt periods (Swenson and Lawrence, 2012). ~~Most SCF parameterizations in LSMs estimate~~
94 ~~SCF based on its relationship with SND or snow water equivalent (SWE) (e.g., Marshall and~~
95 ~~Oglesby,1994; Douville et al., 1995; Yang et al., 1997; Roesch et al., 2001; Niu and Yang, 2007;~~
96 ~~Swenson and Lawrence, 2012; Lalande et al., 2023).~~ Some of these parameterizations ~~also~~
97 account for topographic effects of sub-grid terrain on SCF (e.g., Douville et al., 1995; Roesch et
98 al., 2001; Swenson and Lawrence, 2012; Lalande et al., 2023), which have been shown to be
99 crucial for accurate SCF simulation in mountainous regions (Miao et al., 2022).

100 In CLASSIC, the default parameterization historically used is a linear relationship between SCF
101 and SND with no dependence on topography. A grid cell is considered fully snow-covered when
102 the diagnosed SND reaches 0.1 m (Verseghy, 1991). Melton et al. (2019) investigated the impact
103 of two alternative SCF parameterizations on SCF and permafrost area simulated by CLASSIC.
104 The first was to change the SCF-SND linear relationship to a hyperbolic tangent function (Yang
105 et al., 1997), and the second was to change the SCF-SND linear form to an exponential form
106 (Brown et al., 2003). Both alternative SCF parameterizations worsened performance in terms of
107 the global permafrost area and active layer thickness, so that neither was implemented.

108 Here we consider another option previously developed by Swenson and Lawrence (2012). Their
109 parameterization (hereafter referred as SL12) qualitatively reproduces the hysteresis present in
110 the observational data (SCF-SND relationship) between snow accumulation and ablation seasons
111 while also accounting for the topographic effects of sub-grid terrain. The SL12 parameterization
112 was implemented in the Community Land Model version 5 (CLM5, Lawrence et al., 2019)), the

113 land surface component in the Community Earth System Model version 2 (CESM2,
114 Danabasoglu et al., 2020). Notably, CESM2 was one of the models that showed the lowest
115 surface air temperature and SCF biases over the High Mountainous Asia (HMA) region among
116 the CMIP6 models (Lalande et al., 2021). Based on these results, the SL12 parameterization was
117 implemented in the CLASSIC model and here we evaluate the impact of this change on SCF,
118 SWE, and other snow-related land surface variables. Our evaluation is based on offline
119 CLASSIC simulations forced by historical temperature and precipitation from reanalyses.
120 Because there is uncertainty in these historical values, especially in mountainous regions, we use
121 three different reanalysis-based meteorological datasets to drive CLASSIC. For each
122 meteorological forcing datasets we perform two CLASSIC simulations, one with the default SCF
123 parameterization and one with the SL12 parameterization. The two parameterization schemes are
124 compared with observed SCF and SWE, and the other snow-related land surface variables are
125 evaluated using the Automated Model Benchmarking R package (AMBER, Seiler et al., 2021).
126 The remainder of this paper is organized as follows. In Section 2, we describe the CLASSIC
127 model, the two SCF parameterizations, the forcing data, and model setup. In Section 3, we
128 describe the observation data and our evaluation methods. Results are detailed in Section 4 and
129 discussion points in Section 5. We present conclusions in Section 6.

130

2. CLASSIC model, SCF parameterization methods, and model setup

2.1 CLASSIC description and snow model characteristics

133 CLASSIC is an open-source community land model that is designed to address research
134 questions that explore the role of the land surface in the climate system. It is the successor to the
135 coupled modelling framework based on the Canadian Land Surface Scheme (CLASS; Verseghy,
136 1991; 1993) and the Canadian Terrestrial Ecosystem Model (CTEM; Arora and Boer, 2005;
137 Melton and Arora, 2016). The physics and biogeochemistry modules of CLASSIC are based on
138 CLASS and CTEM models, respectively. Older versions of CLASSIC are (under the name
139 CLASS-CTEM) have served as the land component in the family of Canadian Earth System
140 Models (CanESM) which contribute to the CMIPs (Swart et al., 2019). The CLASSIC model
141 simulations can be performed at point, regional, and global scales both in coupled and offline
142 modes. CLASSIC has been applied in an offline context, i.e. forced with observed meteorology
143 (e.g. Bailey et al., 2000; Bartlett et al., 2006; Melton et al., 2019), as the physical land surface
144 component of regional climate models, e.g. CRCM (Wang et al., 2014; Ganji et al., 2015) and
145 CanRCM (Scinocca et al., 2016), and integrated into each version of the Canadian Atmospheric
146 Model (CanAM; von Salzen et al., 2013), and Earth System Model (CanESM; Arora et al.,
147 2011; Swart et al., 2019) since the early 1990s.

148 The physics component of CLASSIC modelsss energy and water balances separately for the
149 vegetation canopy, snow, and soil (Verseghy, 1991; Melton et al., 2019). As a first-order
150 treatment of subgrid-scale heterogeneity, each grid cell is divided into four sub-areas:, consisting
151 of vegetated, bare soil, vegetated with snow cover, and snow cover over bare soil vegetated and
152 bare soil areas, each with and without snow cover. Snow is represented as a single layer, and,

153 ~~which includes~~ canopy snow processes such as interception, unloading, sublimation and melt <u>are</u>
154 <u>included</u> (Bartlett et al., 2006; Verseghy et al., 2017). The grid cell albedo is computed as a
155 weighted mean based on the fractional coverages for each surface type. In previous versions of
156 CLASSIC, the snow albedo decreases exponentially with time from fresh snow values according
157 to empirically derived functions (Verseghy, 1991). In more recent versions, a new physics-based
158 snow albedo parameterization is available, which accounts for contributions of black carbon
159 snow mixing ratio and the effective snow grain size on snow albedo (Namazi et al., 2015). The
160 new snow albedo scheme is the default scheme in CanESM models and is used in this study.
161 Further details on the CLASSIC model can be found in Melton et al. (2020).

162

2.2 SCF parameterization methods

2.2.1 The current default SCF parameterization

165 In CLASSIC, the thicknesses of all ~~vertical~~ layers <u>(snow and soil)</u> are recommended to be greater
166 than 0.1 m to avoid numerical instability problems. Therefore, the local SND over the snow-
167 covered portion of a grid cell is not allowed to decrease below this threshold (0.1 m), instead, the
168 fractional snow cover decreases to conserve snow mass. Snow cover is considered complete
169 when SND reaches 0.1 m; when SND < 0.1 m, SCF is computed as SCF = SND/0.1, and SND is
170 reset to 0.1 m. Hereafter we refer to the current default SCF parameterization as the Control
171 (CTL) parameterization. Previous analysis has shown that increasing or decreasing this threshold
172 value by 50 % has little effect on the simulated SWE or SCF (Verseghy et al., 2017).

173

2.2.2 The SL12 SCF parameterization

175 Based on snow cover datasets at relatively high spatial and temporal resolution, Swenson and
176 Lawrence (2012) demonstrated that the relationship between SCF and SND depends not only on
177 the amount of snow, but also <u>whether snow mass is increasing (accumulation) or decreasing</u>
178 <u>(ablation).</u>~~on month of the year.~~ This dependence is hypothesized to stem from differences in
179 how accumulation versus ablation processes alter the correlation of the two variables. Based on
180 this they proposed separate formulations for snow accumulation and melt periods as follows.

181 During snow accumulation:

$$f_{sno}^n = 1 - ((1 - \tanh(k_{acc}\Delta W))(1 - f_{sno}^{n-1}) \qquad (1)$$

183 Where f_{sno}^n and f_{sno}^{n-1} are SCF from the current and the previous time step, k_{acc} is a scale
184 parameter (mm^{-1}) and ΔW (mm) is the amount of new snow that falls within the current time
185 step. <u>Eq. (1) assumes that precipitation is randomly distributed across the region, which may be</u>
186 <u>questionable in mountainous areas where snowfall tends to preferentially accumulate at higher</u>
187 <u>elevations. Nevertheless, SCF simulated using the SL12 parameterization from coarse-resolution</u>
188 <u>climate models shows reasonable agreement with observations (e.g. Lalande et al., 2023).</u> Note
189 Eq. (1) is the formulation used in CLM5 code (and implemented in CLASSIC), which is
190 different from that in Swenson and Lawrence (2012). In most LSMs including CLASSIC, SND

191 is diagnostically computed through snow water equivalent (W in Eq. (1)-(4)) and snow density
192 (ρ_s): SND=W/ρ_s. Swenson and Lawrence (2012, their Fig. 7) illustrated that the rate of SCF
193 increase with SND depends on the k_{acc} parameter, such that a larger k_{acc} parameter would result
194 in faster SCF increase with SND. The default value from Swenson and Lawrence 2012) is 0.1
195 mm^{-1}, which is also used in our study. The impact of this choice will be discussed in Section 5.2.
196 During snowmelt:

197
$$f_{sno} = 1 - \left[\frac{1}{\pi}\mathrm{acos}\left(2\frac{W}{W_{max}} - 1\right)\right]^{N_{melt}} \tag{2}$$

198
$$N_{melt} = \frac{200}{max(10,\ \sigma_{topo})} \tag{3}$$

199
$$W_{max} = \frac{W}{0.5\left(cos\left(\pi(1-f_{sno})^{\frac{1}{N_{melt}}}\right)+1\right)} \tag{4}$$

200 where the W and W_{max} are the current and the maximum accumulated snow water equivalent
201 (mm), and N_{melt} (unitless) is a parameter determined from the standard deviation of topography,
202 σ_{topo} (m). Eq. (4) is used to reconcile the relationship during periods of mixed accumulation and
203 melt. Eq. (2) and Eq. (3) suggest that the rate of SCF decrease with SND depends on the N_{melt}
204 parameter, such that SCF decreases faster with (normalized) SND in mountainous areas (small
205 N_{melt}) than flat areas (large N_{melt}, Fig.9 in Swenson and Lawrence, 2012).

206 In our implementation we do not distinguish the use of these two formulations by time of year
207 but based on whether SWE is increasing or decreasing with respect to the previous time step
208 (Wang et al., 2025). To avoid the numerical instability issues mentioned above (Section 2.2.1),
209 the SL12 parameterization is only used when the local SND over the snow-covered portion of a
210 grid cell is greater than 0.1 m. When SND < 0.1 m, SCF is computed in the same way as in the
211 default parameterization. Therefore, the largest difference in SCF between the default and SL12
212 parameterization as implemented in CLASSIC is expected in mountainous areas during the melt
213 period. In these regions and times the topographic effects of sub-grid terrain are accounted for in
214 SL12 but not in CTL.

215

2.3 Forcing data and simulation setup

217 The modeling domain chosen for this study is a global land-only latitude-longitude grid at 1°
218 resolution (Fig. 1a). Three gridded meteorological datasets are used to drive CLASSIC in this
219 study: CRUJRA, ERA5, and GSWP3-W5E5, described below. CRUJRA is regularly used to
220 drive LSMs participating the annual Global Carbon Project which provides analysis of the land
221 carbon sink (Friedlingstein et al., 2025). It was constructed by regridding data from the Japanese
222 reanalysis (JRA, Kobayashi et al., 2015) and adjusting where possible to align with the Climatic
223 Research Unit (CRU) TS4 data (Harris, 2020; 2023). The blended product spanning January
224 1901 to December 2020 has the 6-hourly temporal resolution of the JRA reanalysis product but
225 monthly means adjusted to match the CRU data at 0.5° spatial resolution.

226 ERA5 is the fifth generation European Centre for Medium-Range Weather Forecasts atmospheric
227 reanalysis of the global climate covering the period from January 1940 to present (Hersbach et
228 al., 2020). ERA5 data are available at hourly temporal and 0.25º spatial resolution. Currently it
229 has the highest spatial and temporal resolutions available among all global reanalysis products.

230 GSWP3-W5E5 (here after referred as GSWP3W5) is a combination of two datasets: GSWP3
231 v1.09 (Dirmeyer et al., 2006; Kim 2017) from 1901-1978 and W5E5 v2.0 (Cucchi et al. 2020;
232 Lange et al. 2021) from 1979-2019. It is one of the forcings used in the Inter-Sectoral Impact
233 Model Intercomparison Project (ISIMIP). The GSWP3 dataset is a dynamically downscaled
234 version of the Twentieth Century Reanalysis version 2 (20CRv2; Compo et al. 2011), bias-
235 corrected using three separate observational data sets (see Kim 2017 for details). The W5E5
236 dataset is an interpolated version of ERA5 reanalysis, bias-corrected using CRU TS4. W5E5
237 also provides a second set of precipitation forcing data, bias-corrected with observations from the
238 Global Precipitation Climatology Project (GPCP; Adler et al., 2003). The GPCP dataset includes
239 around 3–4 times as many precipitation stations as CRU, thus we use this version of the
240 precipitation forcing in our experiments. The GSWP3W5 data are available at daily temporal and
241 0.5º spatial resolution.

242 The three meteorological forcing datasets are regridded using the first order conservative
243 remapping method to the 1º model grid via Climate Data Operators. They are disaggregated on
244 the fly within CLASSIC into half-hourly data following the methodology of Melton and Arora
245 (2016) for the following seven meteorological variables that are used to force the model: 2 m air
246 temperature, total precipitation, specific humidity, downward solar radiation flux, downward
247 longwave radiation flux, surface pressure, and wind speed. In CLASSIC, the phase of
248 precipitation is determined by a threshold surface air temperature according to three possible
249 options described in (Bartlett et al., 2006). Jennings et al. (2018) showed that the snowfall-
250 rainfall transition temperature varied from -0.4°C to 2.4°C across the NH. Based on this, we used
251 the option where the partitioning between rainfall and snowfall varies linearly between all
252 rainfall at temperatures above 2º C, and all snowfall at temperatures below 0°C.

253 The plant functional types used in CLASSIC are derived from the Climate Change Initiative land
254 cover product produced by the European Space Agency (Wang et al., 2023). The atmospheric
255 CO_2 concentration values are provided by the Global Carbon Project (Le Quere et al., 2018). The
256 soil texture information consists of the percentage of sand, clay, and organic matter and is
257 derived from the SoilGrids250m dataset (Hengl et al., 2017), and the permeable soil depth is
258 based on Shangguan et al. (2017).

259 CLASSIC simulations use either the CTL or the SL12 parameterization forced by the CRUJRA,
260 ERA5, and GSWP3W5 respectively, yielding six simulations over the historical period. We refer
261 to these simulations hereafter as: CRUJRA-CTL, CRUJRA-SL12, ERA5-CTL, ERA5-SL12,
262 GSWP3W5-CTL, and GSWP3W5-SL12. Pre-industrial spin-up simulations were performed to
263 allow the model to equilibrate carbon fluxes to conditions corresponding to the first year of the
264 forcing data. During spin-up, we loop climate data from the earliest 25 years available for
265 CRUJRA/ERA5 and 100 years of spin-up data for GSWP3W5 (Lange et al., 2022), and hold
266 atmospheric CO2 concentrations at the pre-industrial level (286.46 ppm). The transient runs use

267   time-varying CO2 concentrations and climate. The period from 2005 to 2014 is selected for
268   analyzing the simulated results, when there is overlap with the three observational SCF datasets
269   (see Section 3.1).

270

271   **3. Observation data and evaluation methods**

272

273   **3.1 Study area and evaluation methods**

274   Our analysis will include evaluation of SCF, SWE, meteorological forcings, and other land
275   surface variables. Assessment of SCF, SWE, and meteorological forcings will focus on the
276   mountain (σ_{topo} > 200 m) and flat (σ_{topo} <= 200 m) regions over the Northern Hemisphere (NH),
277   and sub-regions of North America (NA), Eurasia (EA), and HMA. Classification of mountain
278   and flat regions is based on standard deviation of the sub-grid terrain from the ETOPO1
279   elevation data at 1 arc-minute resolution (NOAA, 2009, Fig. 1a). In the SL12 parameterization,
280   the topographic effects of sub-grid terrain are considered via the Nmelt parameter (Eq. (2)),
281   which is inversely related to σ_{topo} (Eq. (3)). Figure 1a shows that at 1° resolution, the magnitudes
282   of σ_{topo} are around 200 m – 600 m for most of the mountainous regions except for the HMA and
283   the Andes where the magnitude of σ_{topo} can reach 1200 m or more.
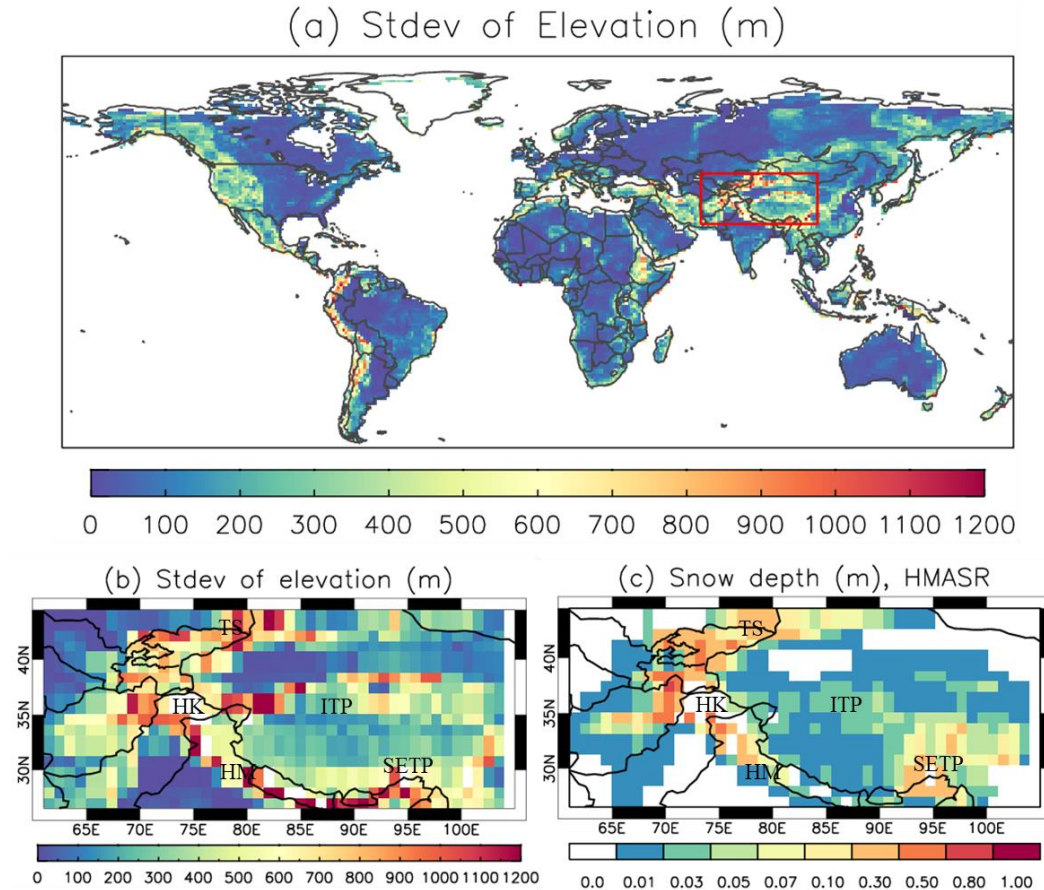

284

Figure1. (a) The standard deviation of elevation over the whole model domain; (b) the standard deviation of elevation in the HMA region (red rectangle box in (a)); (c) HMA mean snow depth during the main snow season (Sep – May) over the 2005-2014 period. Labels in (b) and (c) represent: Tibetan Plateau (TP), interior TP (ITP), southeastern TP (SETP), Tian Shan (TS), Hindu Kush–Karakoram (HK), and western Himalayas (HM).

The HMA region is one of the most complex topographic areas on Earth, with very high sub-grid scale variability (Fig. 1b). It surrounds the Tibetan Plateau (TP), with an average elevation of 4000 m (Du and Qingsong, 2000). Considering the large SCF biases found in CanESM5 and other CMIP models in this region (e.g. Lalande et al., 2021), we will present results for HMA separately. Different regions of HMA exhibit different spatiotemporal patterns in snowfall and SWE due to its unique topography (Yao et al., 2012; Bolch et al., 2019). According to the High Mountainous Asia Snow Reanalysis (HMASR) dataset (see Section 3.2), during Sep. to May over 2005 to 2014 period, SND is only a few centimeters over most of the interior TP, with relatively deeper snow in southeastern TP (Fig. 1c). Deeper snow (SND > 0.2 m) is concentrated at the high elevations of the mountains where σ_{topo} is usually greater than 500 m, such as Tian Shan, Hindu Kush–Karakoram, and western Himalayas (Fig. 1c).

Gridded data are regridded using the first order conservative remapping method to the 1º latitude-longitude grid. In addition to the SCF and SWE data detailed below, the monthly air temperature and precipitation from CRU TS4 (Harris et al., 2020) are used as references to compare with the three meteorological forcing datasets. Evaluation metrics for SCF, SWE and meteorological forcing include the mean bias, unbiased root mean squared error (uRMSE) and Pearson correlation. The uRMSE is defined as the square root of the mean square error minus the squared bias: uRMSE = sqrt (RMSE^2 - Bias^2). Evaluation of other land surface variables is according to AMBER and detailed in Section 3.4.

3.2 SCF observations

The monthly SCF was obtained from the Moderate Resolution Imaging Spectroradiometer (MODIS) /Terra snow cover monthly L3 0.05º Climate Modeling Grid product (MOD10CM, version 61). This dataset provides monthly mean SCF based on the clearest views of the surface from 28 – 31 days of MOD10C1 daily observations and are available from the National Snow and Ice Data Center (Hall and Riggs, 2021). The MODIS snow detection algorithm, which is based on the Normalized Difference Snow Index (NDSI), applies processing steps to alleviate snow detection commission errors and to flag uncertain snow detection (Hall et al., 2002). Due to spectral similarities between cloud and snow, cloud/snow confusion situations remain in MODIS version 6.1 snow products despite continued efforts in improving cloud masking and snow mapping algorithms (Riggs et al., 2019). Regardless of these inherent challenges, the NDSI-based snow detection technique has proven to be a robust indicator of snow presence under diverse situations, as demonstrated by numerous studies reporting accuracy statistics in the range of 88–93% (Riggs et al., 2019).

To mitigate the uncertainties in the MODIS product due to frequent cloud cover and/or complex terrains, SCF from the Interactive Multisensor Snow and Ice Mapping System (IMS) produced by the U.S. National Ice Center (2008) was also used as a reference in our analysis. The IMS snow cover analysis system consists of an interactive workstation for snow cover mapping by a snow analyst (Ramsay,1998; Helfrich et al., 2007). It relies mainly on visible satellite imagery (including MODIS data) but is augmented by station observations and passive microwave data. The IMS dataset consists of binary snow/no snow information on a 4 km resolution polar stereographic projection grid (Helfrich et al. 2007). Though the binary format of this dataset is not ideal for SCF estimation, especially in areas around the snow line, SCF estimates from IMS are included because the resolution of our model is coarse (1º) and IMS data has been used to evaluate modelled SCF in previous studies (e.g. Wang et al., 2014; Orsolini et al., 2019). Daily IMS data were converted to monthly snow cover duration fraction (SCF = total number of days with snow cover in a month divided by the number of days in the month) following the method in Brown et al. (2010).

Previous studies suggested that there were large uncertainties in the SCF data from MODIS and IMS datasets in the HMA region (Hao et al., 2019; Orsolini et al., 2019). Thus, the daily SCF from the HMASR dataset (Liu et al., 2021a) is used as an additional reference for the HMA region in this study. HMASR is based on a Bayesian snow reanalysis framework with model-based snow estimates refined through the assimilation of high resolution SCF data from MODIS (500 m) and Landsat (30 m) sensors (Liu et al., 2021b). The framework also accounts for a priori uncertainties in meteorological forcings and utilizes an ensemble approach (Margulis et al, 2019). The dataset provides daily data of posterior snow estimates at ~500 m spatial resolution over the HMA region. Ensemble mean values of SCF and SND are used in this study. The method used for HMASR is best suited for seasonal snow characterization (Liu et al. (2021a), thus grid cells with semi-permanent snow and ice greater than 30% are masked out in our analysis. The monthly SCF data from MODIS, IMS, and HMASR over the 2005-2014 period are used to evaluate modelled SCF.

3.3 SWE measurements

As shown in Eq. (1) and Eq. (2) simulated SCF is calculated from SWE directly in the SL12 parameterization, and from SND in the CTL parameterization (Section 2.2.1). Therefore, to better understand the sources of bias in simulated SCF, we also evaluate simulated SWE using snow course and airborne gamma SWE observations from Mortimer and Vionnet (2024) covering 1980 – 2014 (Fig. 2). Both types of in situ SWE information have previously been used to evaluate gridded products (e.g. Cho et al. 2019; Mortimer et al. 2020; Mudryk et al. 2025) and details of these data are described elsewhere (Mortimer et al. 2024, Mortimer and Vionnet 2025). Briefly, snow courses generally consist of multiple snow depth and density measurements collected along a predefined transect several hundred meters to several kilometers in length averaged together to obtain a single SWE value for each transect on a given date (WMO, 2018). Airborne gamma SWE estimates are calculated by differencing snow-free and snow-covered measurements of gamma radiation collected along a 15-20 km long flight line with a 300 m wide

footprint after accounting for background soil moisture (Carroll, 2001). Spatial distribution and measurement frequency of the observations varies by measurement method and jurisdiction (e.g. Fig. 2 in Mortimer and Vionnet, 2025). These measurements are better able to capture the larger-scale average compared to single point observations and have been shown capable of discerning subtle differences in SWE products (Mortimer et al. 2022) and of ranking such products based on their relative performance (Mudryk et al. 2025).
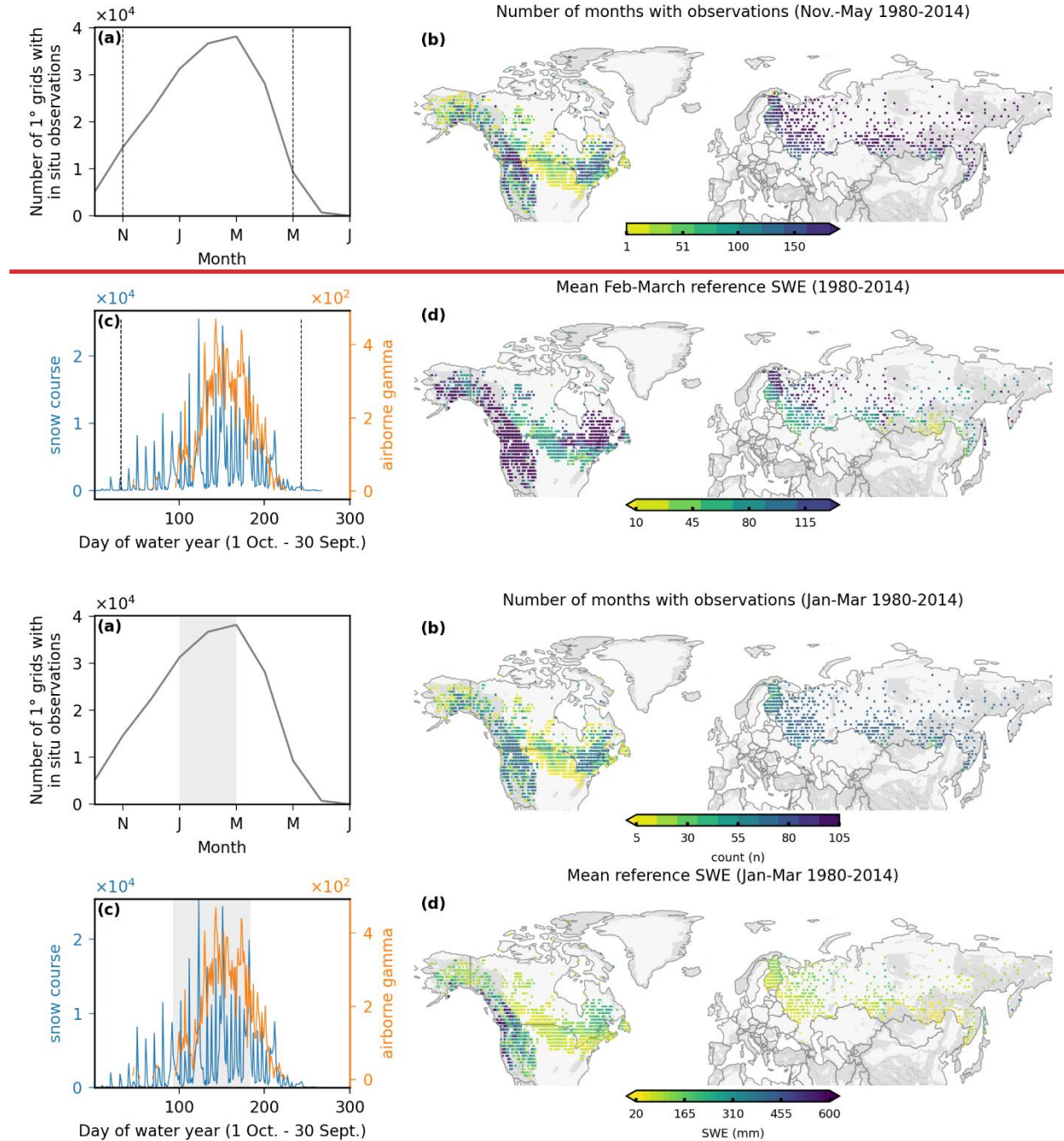
Figure 2. Distribution of in situ reference data. (a) Number of monthly 1°x1° grid cells with reference data during 1980-2014 (each monthly 1° grid with reference data is a data point), (b) Number of months

11

during Nov-May 1980-2014 with reference observations by 1° grid. (c) Temporal distribution of raw in situ SWE observations. (d) Mean February-March reference SWE for grid cells with at least 5 months of data. Vertical lines in (a) and (c) indicate Nov-May period used in the analysis.

To evaluate monthly model output with reference observations from a specific date, we first match reference SWE observations to the model grid estimate from the corresponding month. Next, from these matched data, we calculate the mean reference SWE for each month. If there were multiple reference SWE observations within the same product grid cell on the same date, they were averaged prior to calculating the monthly mean. Metrics were calculated separately for mountainous and flat regions (see definition below) for each month (all years pooled together), for each year (all months pooled together), and for the full time period (all years and months pooled together), and for each product grid (all years pooled together). The analysis is limited to non-zero values with SWE < 3000 mm in both the observation and model outputs, and to the months November through May.

The reference SWE observations do not account for snow-free periods because they are only conducted when there is snow. During the accumulation and ablation seasons, the monthly mean of available reference SWE will therefore often overestimate the true monthly mean value. For this reason, we restrict the comparisons of product SWE with reference SWE to January-March. Additionally, the infrequent sampling of the reference data (Fig. 2c; see also Table 4 in Mortimer and Vionnet, 2025) means that, even when there is continuous snow cover, the monthly value calculated from the available dates with observations may not be representative of the true monthly mean. Investigation of the timing of the in-situ measurements within a month showed that, for the full domain, the timing of the observations is fairly well distributed across a month. However, this varies regionally and by network with some networks (e.g. Canada) biased towards the first half of the month and others (e.g. Russia) slightly biased towards the latter two thirds of the month (Fig. A1). We are unable to account for these biases in our analysis. The statistics calculated from comparisons with in-situ data are not intended to be used as absolute performance measures. Rather, we are interested in assessing how the relative performance of CLASSIC SWE varies under the three choices of forcings; as Mortimer et al. (2024) demonstrates, the reference data is well able to discern relative performance of SWE products.

To evaluate monthly model output with reference observations from a specific date, we first match reference SWE observations to the model grid cell estimate from the corresponding month. Next, from these matched data, we calculate the mean reference SWE for each month. If there were multiple reference SWE observations within the same product grid cell on the same date, they were averaged prior to calculating the monthly mean. Metrics were calculated separately for mountainous and flat regions (as defined in Section 3.1) for each month (all years pooled together), for each year (all months pooled together), for the full time period (all years and months pooled together), and for each product grid cell (all years pooled together). The analysis is limited to non-zero values with SWE ≤ 3000 mm in both the observation and model outputs, and to the months January to March.

3.4 Reference datasets used to evaluate land surface variables in AMBER

Spatial and temporal variations of snow cover account for most of the variations in surface albedo due to its much higher reflectivity relative to underlying land surfaces. Changes in SCF thereby lead to changes in surface albedo, which in turn lead to changes in surface radiation and energy fluxes. To illustrate the impact of the SL12 parameterization on the simulated radiation, energy fluxes, and the water cycle in CLASSIC, we computed skill scores using the AMBER package (Seiler et al., 2021) for the global 1º simulations. AMBER assesses model performance against a collection of observation-based reference datasets based on five scores: bias (S_{bias}), root-mean-square-error (S_{rmse}), phase (S_{phase}), interannual variability (S_{iav}), and spatial distribution (S_{dist}). An overall score ($S_{overall}$) is calculated by averaging the five scores. The scores are dimensionless and on a scale from 0 to 1 where a higher value implies better model performance. Lower values are, however, not necessarily a product of poor model performance as the scores are also affected by uncertainties in the forcing and the reference data. Further details regarding the AMBER package as well as the skill score equations are presented in Seiler et al. (2021) and Seiler (2019). Table 1 shows the 21 reference datasets used in AMBER in this study, which contain information about seven variables relevant to the radiation, energy, and water cycle including net surface radiation (RNS), net surface shortwave radiation (RSS), net surface longwave radiation (RLS), surface albedo (ALBS), latent heat flux (HFLS), sensible heat flux (HFSS), and runoff (MRRO). These datasets include monthly mean values, and more details can be found in Seiler et al. (2021).

Table 1. Overview of the reference datasets used in AMBER, including the following variables: net surface radiation (RNS), net surface shortwave radiation (RSS), net surface longwave radiation (RLS), surface albedo (ALBS), latent heat flux (HFLS), sensible heat flux (HFSS), and runoff (MRRO).

Dataset	Variables	Method	Period	References
CERES	ALBS, RSS, RLS, RNS	Radiative transfer model	2000-2013	Kato et al. (2013)
CLASSr	RNS, HFLS, HFSS, MRRO	Blended product	2003-2009	Hobeichi et al. (2020)
FLUXCOM	RNS, HFLS, HFSS	Machine learning ensemble	1980-2013	Jung et al. (2019)
FLUXNET	RNS, HFLS, HFSS	eddy covariance (204)	1997–2014	Pastorello et al. (2017)
GEWEXSRB	ALBS, RSS, RLS, RNS	radiative transfer model	1984-2007	Stackhouse et al. (2011)
GRDC	MRRO	gauge records (50)	1980–2010	Dai and Trenberth (2002)
GRUN	MRRO	Reconstruction via machine learning	1902-2014	Ghiggi et al. (2019)
MODIS	ALBS	Bidirectional Reflectance Distribution function	2000-2014	Schaaf and Wang (2015)

4. Results

4.1 Comparison of air temperature and precipitation in meteorological datasets

To better understand biases in the simulated snow cover, we first compare air temperature and precipitation from the three meteorological datasets with respect to CRU over the NH and HMA during the 1980-2014 period (Fig. 3 and Fig. A2). Because the CRUJRA data is already bias-corrected to CRU temperature and precipitation, it exhibits very small biases in both variables in all regions relative to this product. By comparison, both ERA5 and GSWP3W5 are colder during most of the months in the NH (Fig. 3a). The magnitude of the cold bias is larger in the mountainous than in the flat regions and larger in GSWP3W5 than in ERA5. Likewise, both ERA5 and GSWP3W5 have more precipitation than CRUJRA over the whole snow season. This difference is especially pronounced in ERA5 in the mountainous regions during the fall and spring months (Fig. 3b and Fig. A2b). In HMA, the bias patterns in temperature and precipitation are similar to those for mountainous regions across the full NH. However, the magnitude of the cold bias (with respect to CRU) is larger in ERA5 than in GSWP3W5 (Fig. 3c). Because different reference datasets were used to bias-adjust precipitation in CRUJRA (CRU) and GSWP3W5 (GPCP), we also compare the monthly precipitation from CRU and GPCP in the above regions and over the same period. This analysis (not shown) indicates that the differences between CRU and GPCP are within 2 % and 3 % for NH flat and mountainous regions respectively, but up to 21 % in HMA.
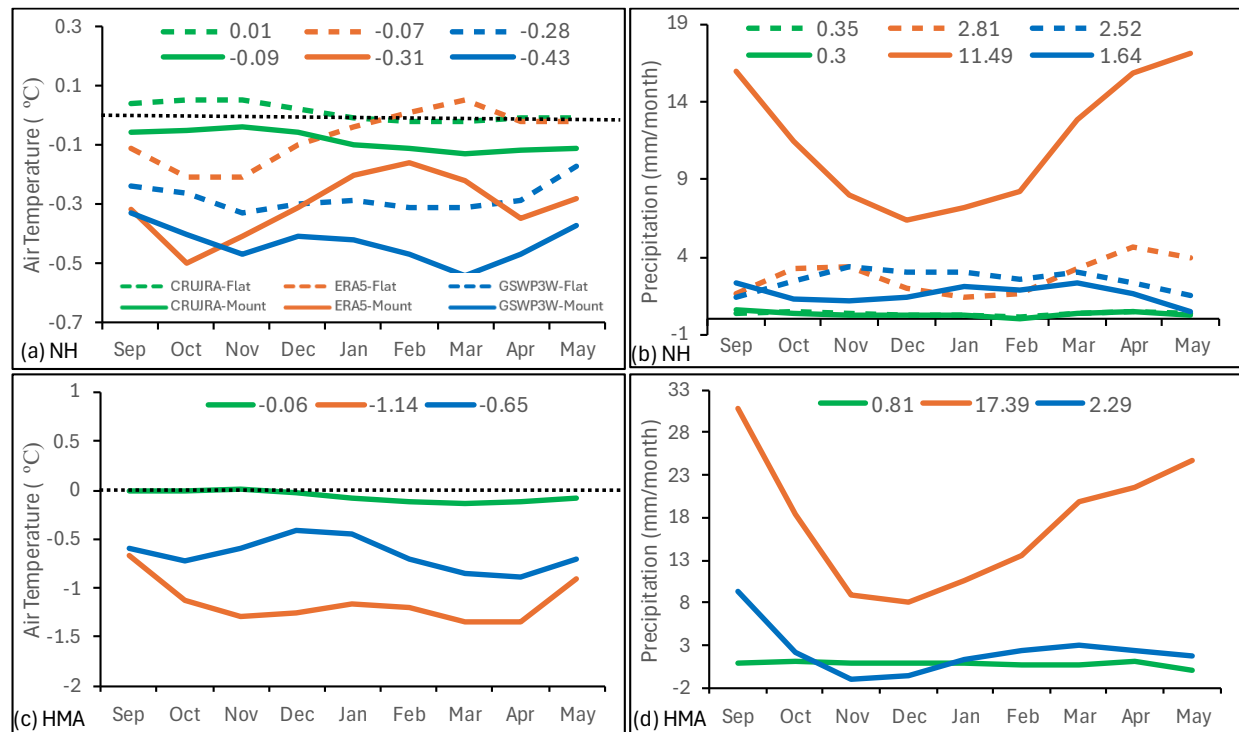

Figure 3. Bias in monthly mean air temperature (a and c) and precipitation (b and d) in the NH mountainous (solid line) and flat (dashed line) regions (a and b) and the HMA mountainous region (c and d) over the 1980-2014 period. Values shown at the top of each plot are the mean temperature or precipitation during Sep-May period for each dataset.

4.2 Evaluation of SWE

Large differences in SWE from the model runs using the CTL and SL12 parameterizations are limited to small areas near grid cells with land ice because the runs are forced by the same three sets of meteorological datasets, and there is limited~~no~~ feedback in offline runs. Thus, we only present results for SWE from the model runs using the SL12 parameterization. The SWE reference measurements (Section 3.2) indicate that for all choices of meteorological forcing, CLASSIC underestimates SWE in mountainous regions (Fig. 4a) and overestimates SWE in flat regions (Fig. 4b) over the 1980-2014 period. For both types of regions, the magnitudes of the biases increase as the snow season progresses. In the mountainous regions, the biases are similar ~~for~~
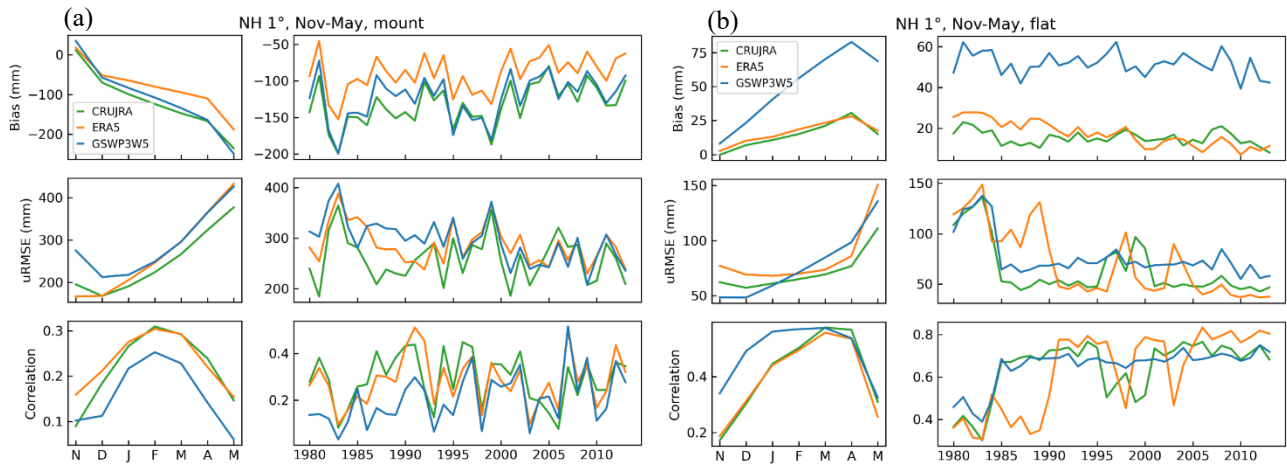

~~Figure 4. Annual and interannual evolution of bias, uRMSE, and correlation for modelled SWE in model runs using the SL12 parameterization forced by CRUJRA, ERA5, and GSWP3-W5E5 in (a) NH mountainous regions and (b) NH flat regions over the 1980-2014 period.~~

for

GSWP3W5-SL12 (-129.4) and CRUJRA-SL12 (-136.6) and the lowest~~slightly smaller~~ for ERA5-SL12 (-90.8). In flat regions, GSWP3W5-SL12 (50.3) has more than twice the SWE bias seen in either CRUJRA-SL12 (15.0) or ERA5-SL12 (17.5), which is mainly due to SWE overestimation in eastern NA and northern Europe (Fig. A3~~1~~). ~~While in mountainous regions the uRMSE increases nearly linearly from Dec. to May, in flat regions, uRMSE remains under 100 mm from Nov to April but then increases sharply during April and May.~~ Overall, ERA5-SL12 outperforms the other two model runs with lower bias ~~and better correlation~~ in mountainous regions and it shows similar performance as CRUJRA-SL12 in flat regions.
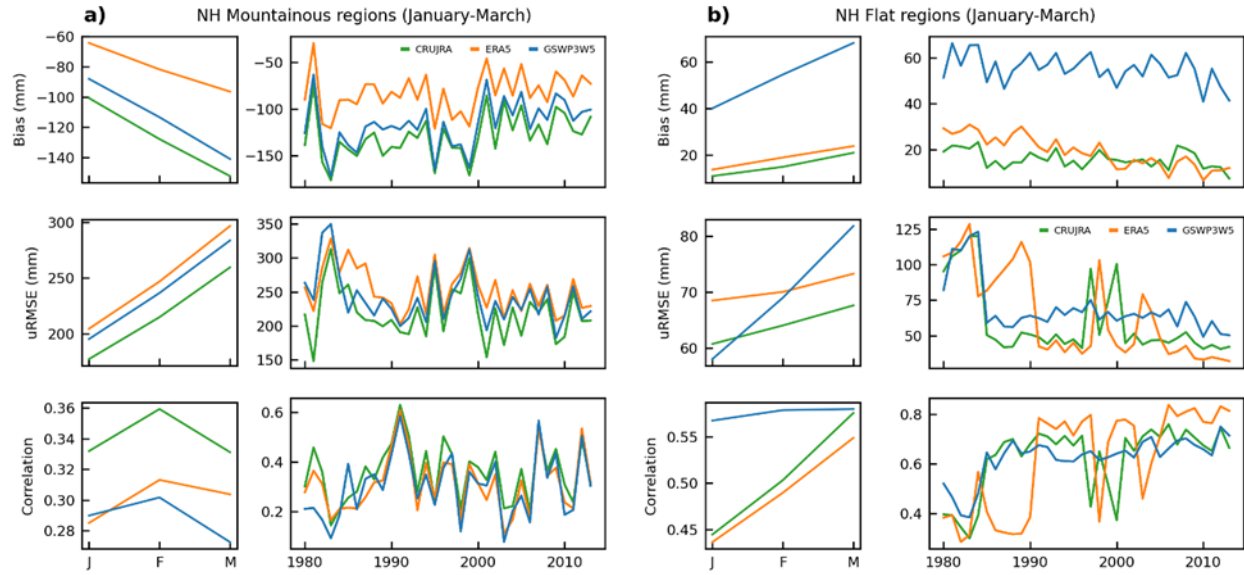
15

Figure 4. Annual and interannual evolution of bias, uRMSE, and correlation for modelled SWE in model runs using the SL12 parameterization forced by CRUJRA, ERA5, and GSWP3-W5E5 in (a) NH mountainous regions and (b) NH flat regions over the 1980-2014 period.

Table 2. The seasonal mean SCF bias, uRMSE, and Pearson correlation coefficient (r) for the Control and SL12 simulations over the (a) NH mountainous regions (σ_{topo} >200 m), (b) NH flat regions (σ_{topo} <= 200 m). The observed SCF from MODIS is used as the reference.

(a) NH Mountain	SON			DJF			MAM			Annual		
Met-Scheme	Bias	uRMSE	r	Bias	uRMSE	r	Bias	uRMSE	r	Bias	uRMSE	r
CRUJRA - CTL	-0.01	0.08	0.55	0.06	0.08	0.24	0.10	0.12	0.45	0.04	0.13	0.59
CRUJRA - SL12	-0.04	0.07	0.56	0.01	0.07	0.31	0.01	0.08	0.55	-0.01	0.09	0.62
ERA5 - CTL	0.01	0.06	0.59	0.07	0.07	0.28	0.09	0.11	0.48	0.05	0.12	0.62
ERA5 - SL12	-0.02	0.06	0.60	0.02	0.06	0.38	0.01	0.06	0.60	0.00	0.08	0.66
GSWP3W5 - CTL	-0.02	0.07	0.57	0.03	0.08	0.29	0.05	0.11	0.48	0.03	0.13	0.59
GSWP3W5-SL12	-0.04	0.07	0.58	-0.02	0.07	0.35	-0.03	0.07	0.56	-0.02	0.09	0.64
(b) NH Flat	SON			DJF			MAM			Annual		
Met-Scheme	Bias	uRMSE	r	Bias	uRMSE	r	Bias	uRMSE	r	Bias	uRMSE	r
CRUJRA - CTL	-0.02	0.07	0.57	0.02	0.05	0.20	0.09	0.12	0.44	0.03	0.11	0.59
CRUJRA - SL12	-0.04	0.08	0.57	0.01	0.06	0.24	0.08	0.11	0.47	0.02	0.11	0.59
ERA5 - CTL	-0.02	0.07	0.58	0.01	0.05	0.24	0.07	0.09	0.50	0.02	0.10	0.61
ERA5 - SL12	-0.04	0.08	0.58	0.00	0.05	0.27	0.06	0.09	0.52	0.01	0.10	0.61
GSWP3W5 - CTL	0.00	0.08	0.57	0.02	0.06	0.19	0.10	0.13	0.41	0.04	0.13	0.58
GSWP3W5-SL12	-0.02	0.08	0.57	0.01	0.06	0.23	0.09	0.13	0.45	0.03	0.12	0.58

16

4.3 Evaluation of SCF

4.3.1 NH regions

Figure 5 shows the monthly mean SCF (area weighted) from all six simulations along with the MODIS and IMS observations over different regions. SCF from MODIS and IMS generally agree well with each other in all regions except for HMA, where IMS shows ~3 % - 6 % more SCF than MODIS in the winter months (Fig. 5g). In the NH, NA, and EA mountainous regions (Fig. 5a-5c and Table 2), both the CTL and the SL12 parameterizations underestimate SCF in the fall (SON), with the SL12 parameterization performing slightly worse than the CTL parameterization. However, during winter (DJF) and spring (MAM), the SL12 parameterization greatly outperforms the CTL parameterization for all three meteorological datasets. For example, in the NH mountains during the spring, the mean biases are 0.1, 0.09, and 0.05 with the CTL parameterization for model runs forced by CRUJRA, ERA5, and GSWP3W5 respectively; they are 0.01, 0.01, and -0.03 with the SL12 parameterization (Table 2a). The uRMSEs are 0.12, 0.11, and 0.11 with the CTL parameterization, and 0.08, 0.06, and 0.07 with the SL12
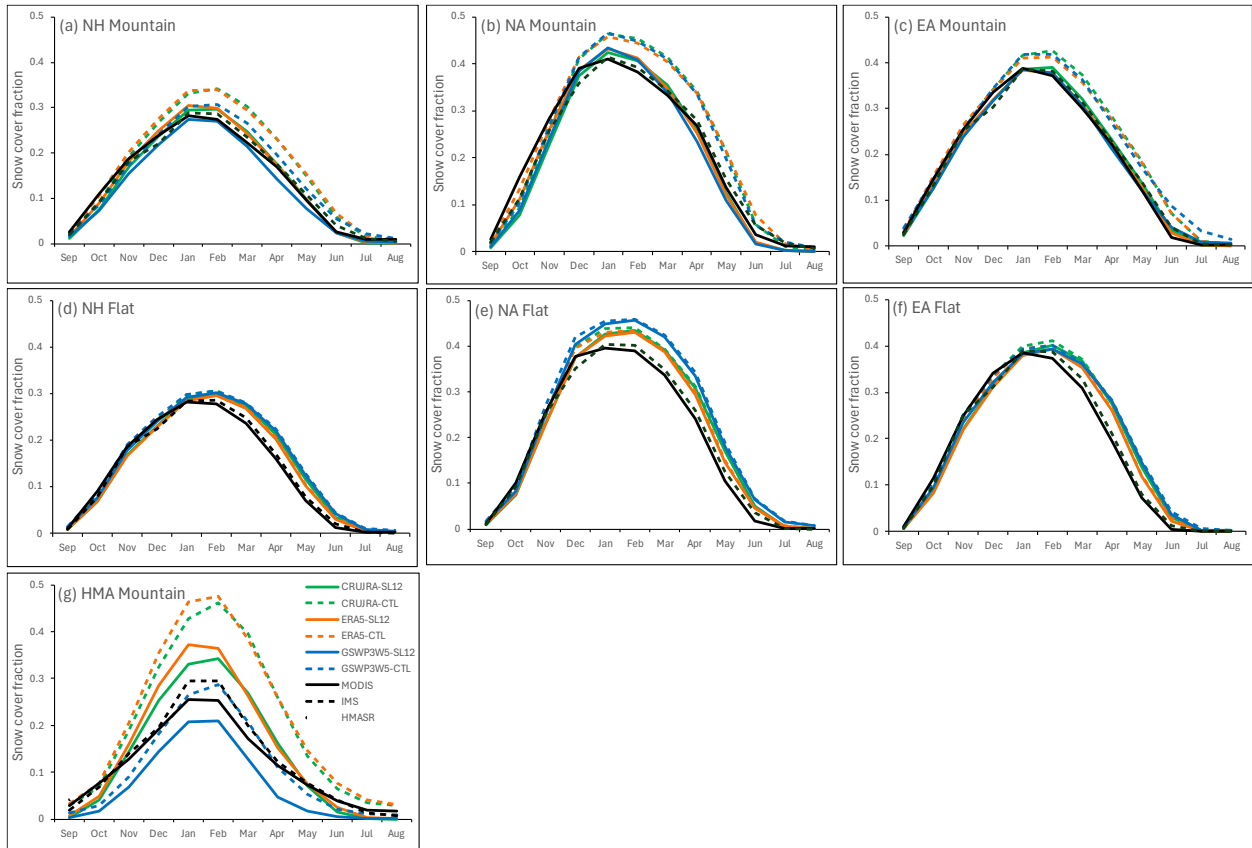

Figure 5. The monthly mean SCF from model runs using the Control (dashed line) and SL12 (solid line) parameterizations for NH, NA, and EA mountainous (σtopo $\geqslant$ 200 m, a-c) and flat (σtopo < 200 m, d-e) regions, and (g) shows the monthly mean SCF for the HMA mountainous region. The black lines represent observed SCF from MODIS (solid), IMS (dashed), and HMASR (dotted).

528 parameterization; and the correlation coefficients are 0.45, 0.48, and 0.48 with the CTL
529 parameterization, and 0.55, 0.60, 0.56 with the SL12 parameterization (Table 2a). On average for
530 all three meteorological forcing choices, the annual mean bias, uRMSE, and correlation improve
531 by 75 %, 32 %, and 7 % when evaluated with MODIS SCF observations over the NH
532 mountainous regions.

533 In flat regions (all domains), as expected, the performance is similar regardless of the
534 parameterization with a 2-4 % SCF underestimation in the fall, but a 1-2 % and 6-10 % SCF
535 overestimation during the winter and spring seasons, respectively (Fig. 5d-5f and Table 2b).
536 Among the six simulations, ERA5-SL12 has the lowest annual bias (0.0) and uRMSE (0.08), and
537 the highest correlation (0.66) in the NH mountainous regions, as well as in the flat regions
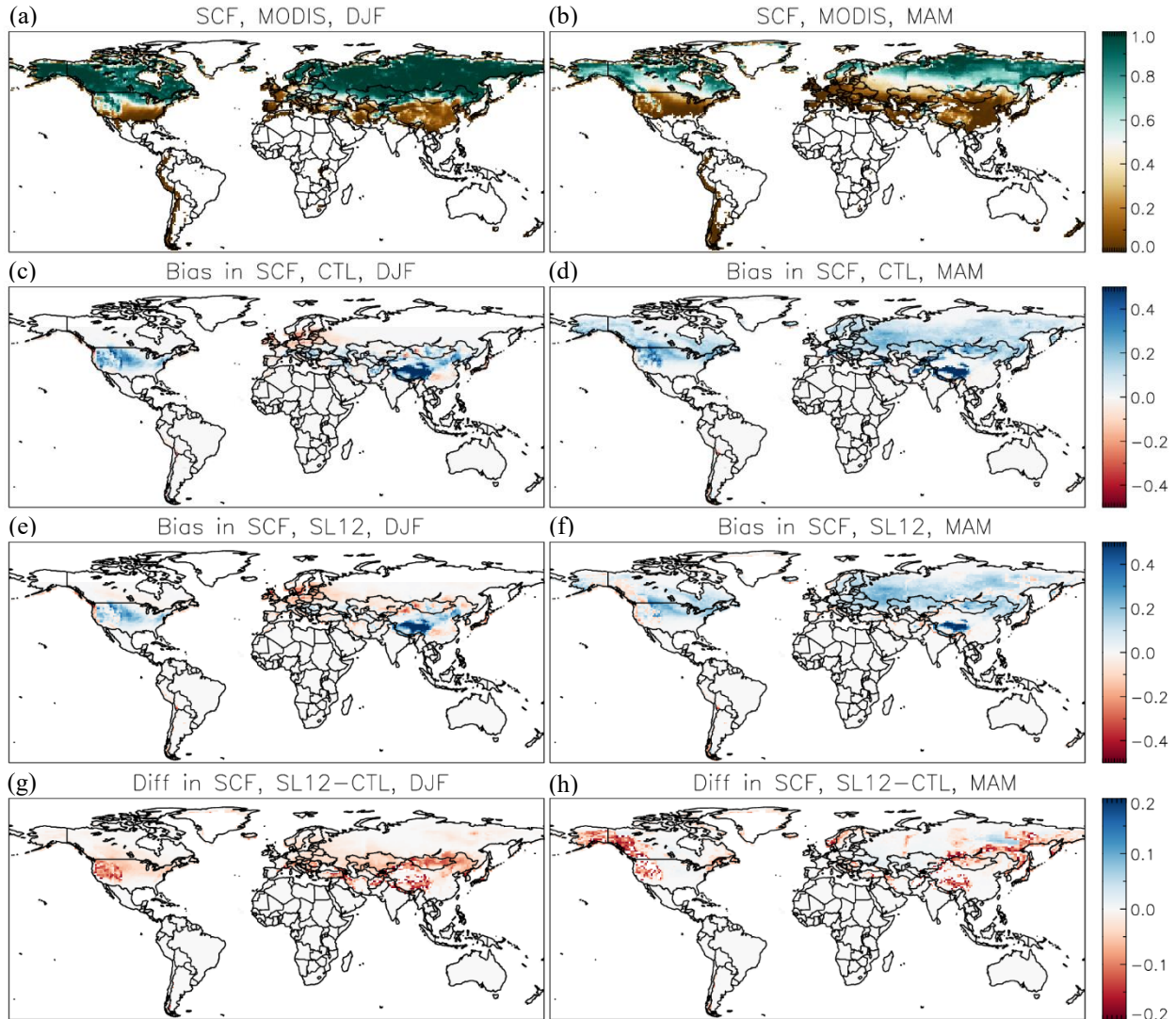538 (bias=0.01, uRMSE=0.1, and r=0.61) (Table 2).

539

540
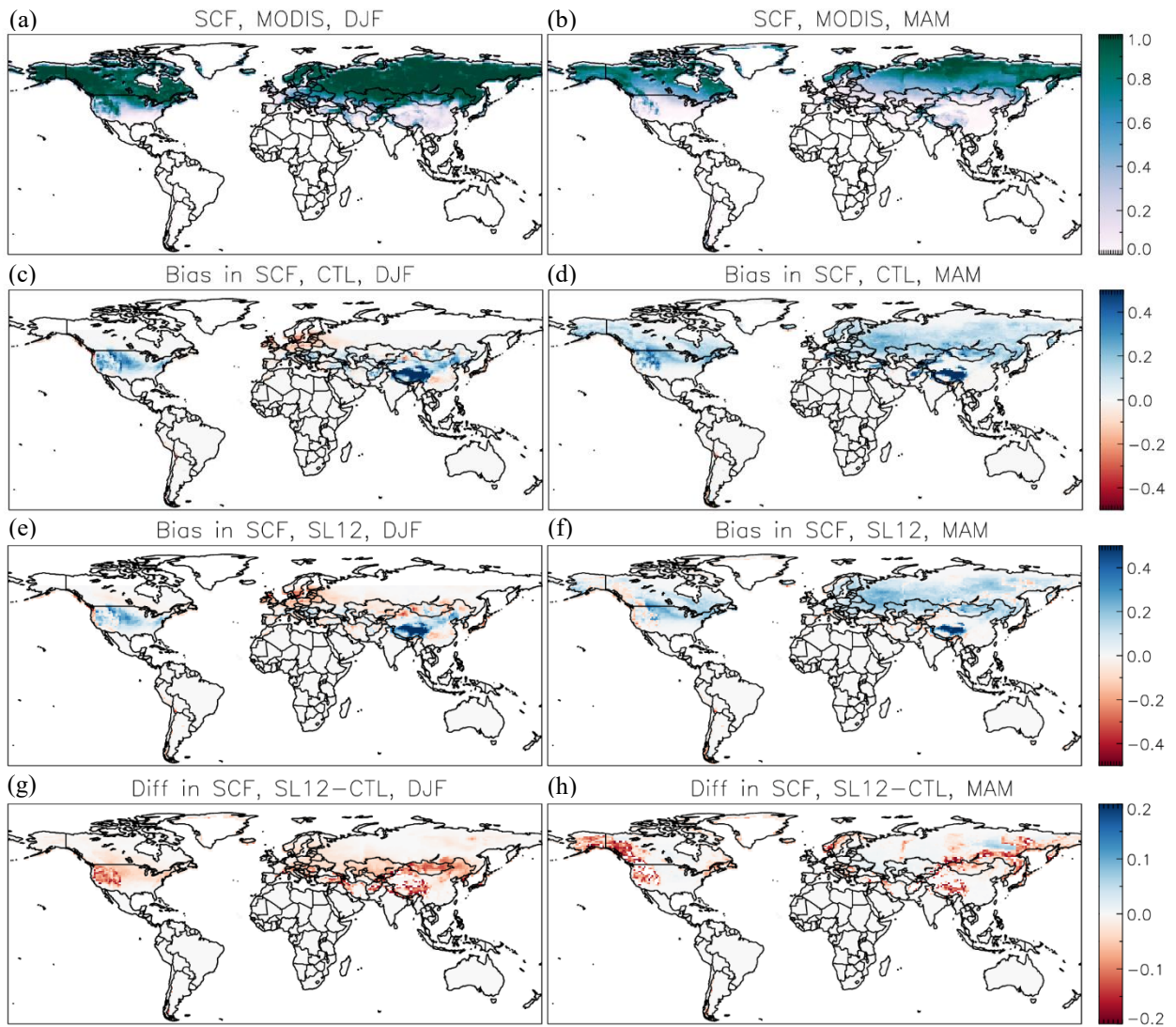
Figure 6. Snow cover fraction from MODIS (a and b), SCF bias in model runs using the Control (c and d) and SL12 (e and f) parameterizations, and difference in SCF between SL12 and Control (g and h) during the winter (left) and spring (right) season.

On the global scale, the spatial patterns of SCF bias are similar for all three meteorological forcing choices. Figure 6 shows an example of the spatial pattern in SCF bias from the model runs forced by ERA5 during the winter and spring seasons. Compared to observed SCF from MODIS, model runs tend to overestimate SCF in areas where SCF is less than 100 % in both the winter and spring seasons. In the winter, both parameterizations have areas with SCF underestimation, such as in the western NA mountainous areas, northern Europe, and some areas of Asia (Fig. 6c and 6e). In the spring, the CTL parameterization overestimates SCF in most NH regions except for some limited areas in western NA (Fig. 6d). The SCF overestimation is reduced in the run using the SL12 parameterization, and replaced with some SCF

573 underestimation, such as in the western NA mountains (Fig. 6f). Overall, the SL12
574 parameterization produces less SCF and thus reduces the SCF overestimation found in the model
575 runs using the CTL parameterization over all major mountain ranges across the globe (Fig. 6g
576 and 6h).

577

578 **4.3.2 HMA region**

579 In HMA, large uncertainties have been found in SCF from the MODIS and IMS datasets (Hao et
580 al., 2019; Orsolini et al., 2019), thus SCF from the HMASR dataset is also included as a
581 reference along with MODIS and IMS. Results are only shown for the mountainous region (Fig.
582 5g) because there are limited flat areas with snow cover (Fig. 1b and 1c). HMASR has a single
583 peak in Feb., while MODIS, IMS, and all the model runs have peaks in both Jan. and Feb. Over
584 this region, simulations using either parameterization exhibit large SCF overestimations during
585 the winter and spring compared to all three reference datasets especially when forced by
586 CRUJRA or ERA5 (Fig. 5g). Compared to SCF from HMASR, the mean biases are 0.30 and
587 0.35 in CRUJRA-CTL and ERA5-CTL respectively during the winter (Table 3). In contrast, the
588 model runs driven by GSWP3W5 have much lower SCF and smaller biases (Fig. 5g and Table
589 3). Overall, the SL12 parameterization exhibits improved performance compared to the CTL
590 parameterization. On average from all three meteorological forcing choices, the annual mean
591 bias, uRMSE, and correlation improve by 48 %, 30 %, and 5 % when evaluated with HMASR
592 SCF data over the HMA mountainous areas.

593 **Table 3.** Same as Table 2 but for the HMA region. SCF from the HMASR dataset is used as the reference.

HMA Mountain	SON			DJF			MAM			Annual		
Met-Scheme	Bias	uRMSE	r	Bias	uRMSE	r	Bias	uRMSE	r	Bias	uRMSE	r
CRUJRA - CTL	0.02	0.12	0.35	0.30	0.15	0.31	0.20	0.17	0.39	0.13	0.21	0.42
CRUJRA - SL12	-0.03	0.09	0.37	0.16	0.11	0.35	0.06	0.10	0.45	0.04	0.15	0.44
ERA5 - CTL	0.05	0.12	0.43	0.35	0.14	0.36	0.23	0.15	0.40	0.16	0.22	0.45
ERA5 - SL12	-0.01	0.09	0.45	0.22	0.11	0.42	0.08	0.09	0.51	0.06	0.15	0.48
GSWP3W - CTL	-0.06	0.08	0.40	0.08	0.14	0.39	0.01	0.12	0.44	0.00	0.14	0.45
GSWP3W - SL12	-0.08	0.07	0.41	0.00	0.11	0.39	-0.08	0.08	0.48	-0.05	0.10	0.46

594

595 In HMA, areas with high SCF (> 40 %) are mainly found along the western mountain ranges
596 (e.g. Tian Shan, Hindu Kush–Karakoram, and western Himalayas) and southeast portion of the
597 TP (Fig.7a-7c). SCF is less than 20 % in most of the interior TP, even during the winter (Fig. 7a).
598 On average, maximum SCF occurs in winter in western HMA (i.e. Tian Shan and Hindu Kush–
599 Karakoram), but it occurs in spring in interior TP and southeast TP. Among the model runs using
600 the CTL parameterization, there are significant SCF overestimations in most of HMA when
601 forced by CRUJRA or ERA5 (Fig. 7d, 7e). The run forced by GSWP3W5 still overestimates
602 SCF in the mountainous areas of western HMA but underestimates SCF in the interior TP and
603 southeast of TP (Fig. 7f). Given that all three simulations use the same CTL parameterization

(Fig. 7d–f), the substantial differences in simulated SCF, particularly in the GSWP3W5-forced run, suggest that the primary source of the discrepancy lies in the forcing data. This will be discussed further in Section 5. In the model runs using the SL12 parameterization (Fig. 7g-7i), the SCF overestimations are much reduced in the western mountainous areas while across the rest of the plateau the SCF underestimations are very similar for both parameterizations.
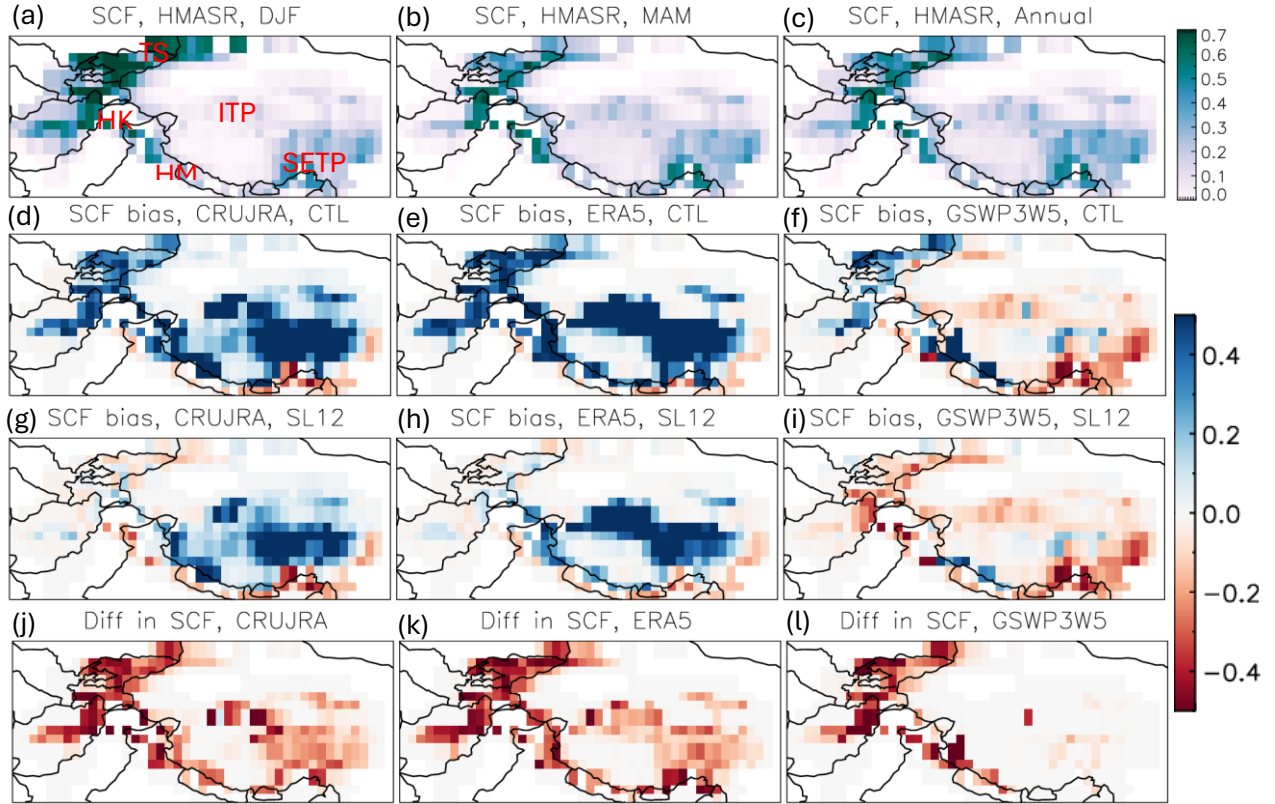

Figure 7. The top panel shows SCF from HMASR for (a) winter, (b) spring, and (c) annual mean. The second and third panel shows SCF biases from model runs using the CTL (d-f) and SL12 (g-i) parameterizations forced by the three meteorological datasets respectively during spring. The bottom panel (j-k) shows the difference in SCF between the model runs using the SL12 and CTL parameterizations.

4.4 Evaluation of other land surface variables

Evaluation of other land surface variables (besides SCF and SWE) via AMBER scores (Section 3.3) is shown in Fig. 8 for each of the six CLASSIC simulations. Model runs using the SL12 parameterization have the best score for 101 of 119 diagnostic tests while they have the worst score for only 16 of 119 diagnostic tests (Fig. 8c and 8d). CRUJRA-SL12 (ID=2) and ERA5-SL12 (ID=4) have the highest overall scores for five radiation reference datasets (one RNS, two RSS, two RLS), and three surface albedo (ALBS) reference datasets with improvements ranging from 0.01 to 0.06 when compared to the runs with the lowest scores (Fig. 8b and 8c). The relatively large score differences in the interannual variability score (S_{iav}) for net surface

radiation (RNS) suggest improved interannual variability of net surface radiation when using the

SL12 parameterization (Fig. 8b). For surface albedo, relatively large differences are observed in

the spatial distribution score (S_{dist}), suggesting better characterization of the spatial patterns in

surface albedo when using the SL12 parameterization (Fig. 9). Figure 9 shows that surface

albedo is generally overestimated by the control scheme (Fig. 9a), with this overestimation

notably reduced in the mountainous regions when the SL12 scheme is applied (Fig. 9b),

consistent with the improvements seen in SCF. Previous studies have indicated that the MODIS

surface albedo product may exhibit biases due to the absence of shading corrections in

mountainous areas and underestimation of snow cover in dense forest regions (Hall et al., 2002;

Bair et al., 2022). These limitations may have contributed, at least in part, to the albedo

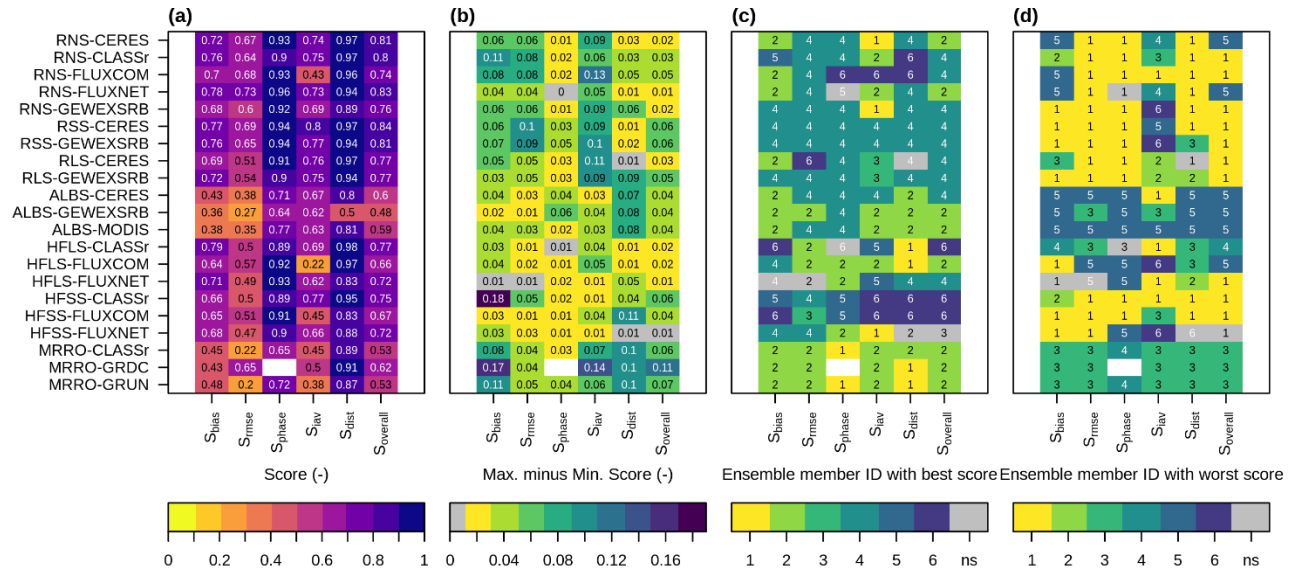overestimation shown in Figure 9a.

Figure 8. AMBER results for other land surface variables from the six model runs, (a) mean ensemble
score, (b) maximum score difference among ensemble members, (c) ensemble member with the highest
score, and (d) ensemble member with the lowest score. Comparisons are grayed out in panels (b–d) when
the difference between the maximum and minimum scores is less than 0.01. Ensemble member IDs
represent the following model runs: 1: CRUJRA-CTL, 2: CRUJRA-SL12, 3: ERA5-CTL, 4: ERA5-SL12,
5: GSWP3W5-CTL, 6: GSWP3W5-SL12.

Though GSWP3W5-SL12 (ID=6) has the lowest frequency of the model runs with the best

scores (Fig. 8c), it has the highest overall performance for some of the heat fluxes datasets - one

out of the three HFLS and two out of the three HFSS reference datasets. For surface runoff,

model runs with the best scores are all forced by CRUJRA, while model runs with the worst

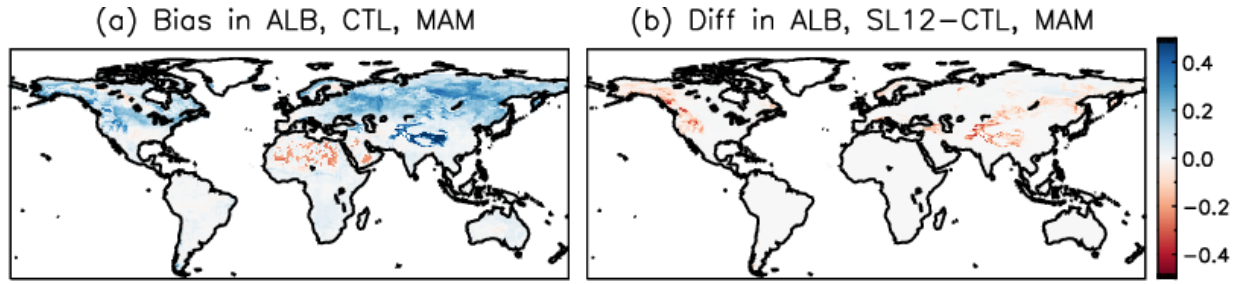scores are all forced by ERA5 (Fig. 8c and 8d).

(a) Bias in ALB, CTL, MAM

(b) Diff in ALB, SL12−CTL, MAM

To isolate the impact of meteorological forcing data and SCF parameterization on these snow-related variables, we also calculate AMBER scores for the three model runs separately for the SL12 (Fig. 10) and the CTL (Fig. A4) parameterizations. The results show that regardless of the parameterization, overall model runs forced by ERA5 (ID = 2) perform best for most radiation fluxes, while model runs forced by CRUJRA (ID = 1) perform best for the rest of the variables except for some heat fluxes where model runs forced by GSWP3W5 (ID = 3) perform best (Fig. 10c). These are generally consistent with results shown in Fig. 8 with both parameterizations included, suggesting that the score differences among ensemble members are largely due to differences in the meteorological forcing. However, the overall scores with the SL12 parameterization (Fig. 10a) are slightly larger for most variables than those with the CTL parameterization (Fig. A4a). Among the three model runs using the SL12 parameterization, ERA5-SL12 has the most (43/99) frequency in the model runs with the best scores (Fig. 10c), followed by CRUJRA-SL12 (38), with GSWP3W5-SL12 having the least frequency (18).
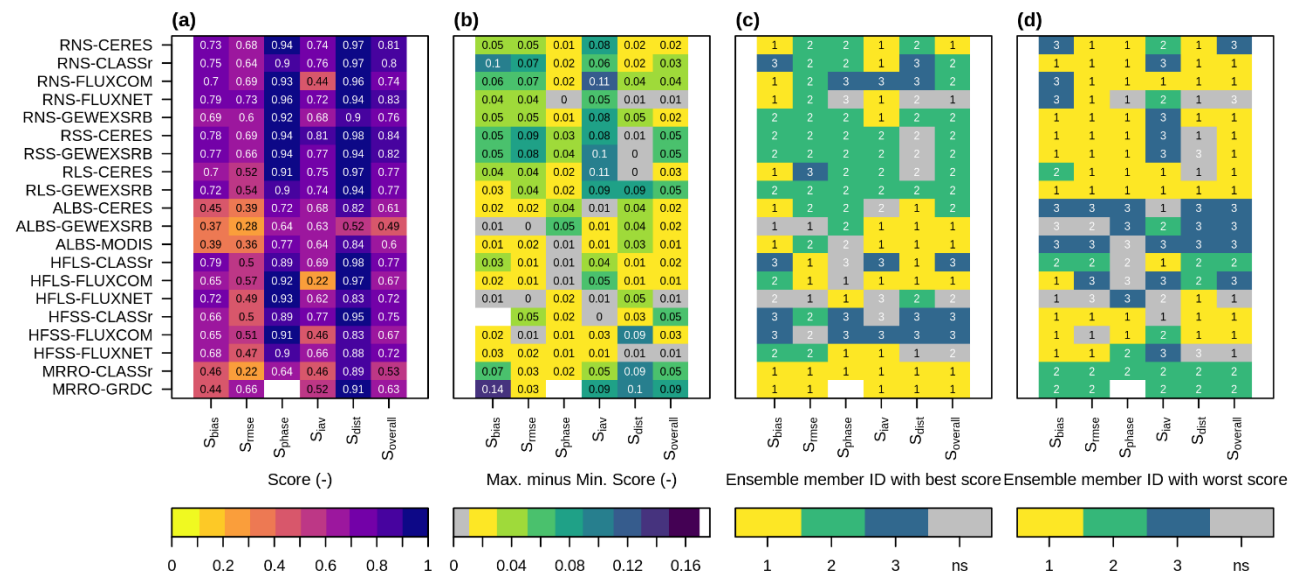
(a) Score (-)

	S_{bias}	S_{rmse}	S_{phase}	S_{iav}	S_{dist}	$S_{overall}$
RNS-CERES	0.73	0.68	0.94	0.74	0.97	0.81
RNS-CLASSr	0.75	0.64	0.9	0.76	0.97	0.8
RNS-FLUXCOM	0.7	0.69	0.93	0.44	0.96	0.74
RNS-FLUXNET	0.79	0.73	0.96	0.72	0.94	0.83
RNS-GEWEXSRB	0.69	0.6	0.92	0.68	0.9	0.76
RSS-CERES	0.78	0.69	0.94	0.81	0.98	0.84
RSS-GEWEXSRB	0.77	0.66	0.94	0.77	0.94	0.82
RLS-CERES	0.7	0.52	0.91	0.75	0.97	0.77
RLS-GEWEXSRB	0.72	0.54	0.9	0.74	0.94	0.77
ALBS-CERES	0.45	0.39	0.72	0.68	0.82	0.61
ALBS-GEWEXSRB	0.37	0.28	0.64	0.63	0.52	0.49
ALBS-MODIS	0.39	0.36	0.77	0.64	0.84	0.6
HFLS-CLASSr	0.79	0.5	0.89	0.69	0.98	0.77
HFLS-FLUXCOM	0.65	0.57	0.92	0.22	0.97	0.67
HFLS-FLUXNET	0.72	0.49	0.93	0.62	0.83	0.72
HFSS-CLASSr	0.66	0.5	0.89	0.77	0.95	0.75
HFSS-FLUXCOM	0.65	0.51	0.91	0.46	0.83	0.67
HFSS-FLUXNET	0.68	0.47	0.9	0.66	0.88	0.72
MRRO-CLASSr	0.46	0.22	0.64	0.46	0.89	0.53
MRRO-GRDC	0.44	0.66		0.52	0.91	0.63

(b) Max. minus Min. Score (-)

	S_{bias}	S_{rmse}	S_{phase}	S_{iav}	S_{dist}	$S_{overall}$
RNS-CERES	0.05	0.05	0.01	0.08	0.02	0.02
RNS-CLASSr	0.1	0.07	0.02	0.06	0.02	0.03
RNS-FLUXCOM	0.06	0.07	0.02	0.11	0.04	0.04
RNS-FLUXNET	0.04	0.04	0	0.05	0.01	0.01
RNS-GEWEXSRB	0.05	0.05	0.01	0.08	0.05	0.02
RSS-CERES	0.05	0.09	0.03	0.08	0.01	0.05
RSS-GEWEXSRB	0.05	0.08	0.04	0.1	0	0.05
RLS-CERES	0.04	0.04	0.02	0.11	0	0.03
RLS-GEWEXSRB	0.03	0.04	0.02	0.09	0.09	0.05
ALBS-CERES	0.02	0.02	0.04	0.01	0.04	0.02
ALBS-GEWEXSRB	0.01	0	0.05	0.01	0.04	0.02
ALBS-MODIS	0.01	0.02	0.01	0.01	0.03	0.01
HFLS-CLASSr	0.03	0.01	0.01	0.04	0.01	0.02
HFLS-FLUXCOM	0.02	0.01	0.01	0.05	0.01	0.01
HFLS-FLUXNET	0.01	0	0.02	0.01	0.05	0.01
HFSS-CLASSr		0.05	0.02	0	0.03	0.05
HFSS-FLUXCOM	0.02	0.01	0.01	0.03	0.09	0.03
HFSS-FLUXNET	0.03	0.02	0.01	0.01	0.01	0.01
MRRO-CLASSr	0.07	0.03	0.02	0.05	0.09	0.05
MRRO-GRDC	0.14	0.03		0.09	0.1	0.09

(c) Ensemble member ID with best score

	S_{bias}	S_{rmse}	S_{phase}	S_{iav}	S_{dist}	$S_{overall}$
RNS-CERES	1	2	2	1	2	1
RNS-CLASSr	3	2	2	1	3	2
RNS-FLUXCOM	1	2	3	3	3	2
RNS-FLUXNET	1	2	3	1	2	1
RNS-GEWEXSRB	2	2	2	1	2	2
RSS-CERES	2	2	2	2	2	2
RSS-GEWEXSRB	2	2	2	2	2	2
RLS-CERES	1	3	2	2	2	2
RLS-GEWEXSRB	2	2	2	2	2	2
ALBS-CERES	1	2	2	2	1	2
ALBS-GEWEXSRB	1	1	2	1	1	1
ALBS-MODIS	1	2	2	1	1	1
HFLS-CLASSr	3	1	3	3	1	3
HFLS-FLUXCOM	2	1	1	1	1	1
HFLS-FLUXNET	2	1	2	1	1	1
HFSS-CLASSr	3	2	3	3	3	3
HFSS-FLUXCOM	3	2	3	3	3	3
HFSS-FLUXNET	2	2	1	1	1	2
MRRO-CLASSr	1	1	1	1	1	1
MRRO-GRDC	1	1		1	1	1

(d) Ensemble member ID with worst score

	S_{bias}	S_{rmse}	S_{phase}	S_{iav}	S_{dist}	$S_{overall}$
RNS-CERES	3	1	1	2	1	3
RNS-CLASSr	1	1	1	3	1	1
RNS-FLUXCOM	3	1	1	1	1	1
RNS-FLUXNET	3	1	1	2	1	3
RNS-GEWEXSRB	1	1	1	3	1	1
RSS-CERES	1	1	1	3	1	1
RSS-GEWEXSRB	1	1	1	3	3	1
RLS-CERES	2	1	1	1	1	1
RLS-GEWEXSRB	1	1	1	1	1	1
ALBS-CERES	3	3	3	1	3	3
ALBS-GEWEXSRB	3	2	3	2	3	3
ALBS-MODIS	3	3	3	3	3	3
HFLS-CLASSr	2	2	2	1	2	2
HFLS-FLUXCOM	1	3	3	3	2	3
HFLS-FLUXNET	1	1	3	2	1	1
HFSS-CLASSr	1	1	1	1	1	1
HFSS-FLUXCOM	1	1	1	2	1	1
HFSS-FLUXNET	1	1	2	3	1	3
MRRO-CLASSr	2	2	2	2	2	2
MRRO-GRDC	2	2		2	2	2

Figure 10~~9~~. Same as in Fig. 8 except for the three model runs using the SL12 parameterization. Ensemble member IDs represent the following model runs: 1: CRUJRA-SL12, 2: ERA5-SL12, 3: GSWP3W5-SL12.

5. Discussion

This study evaluates the SL12 SCF parameterization against the current default (CTL) parameterization on snow simulation in CLASSIC. To account for uncertainties in the meteorological forcing data, three reanalysis-based datasets are used to drive the model. Biases in modeled SCF vary between flat and mountainous regions for both SCF parameterizations (Table 2, Fig. 5, and Fig. 6). Previous studies have highlighted the importance of accounting for sub-grid topography on SCF simulations in mountainous regions (Swenson and Lawrence, 2012; Miao et al., 2022). Are the modelled SCF biases related to topographic complexity in this study? To explore this, we generated scatter plots and examined the correlations between SCF biases and the standard deviation of sub-grid topography during the winter and spring seasons for each simulation. As expected, significant correlations were found in all simulations, indicating that SCF biases tend to increase with increasing topographic complexity. However, this relationship is notably reduced under the SL12 scheme, particularly in spring. An example of these scatter plots, based on model runs forced by CRUJRA, is presented in Figure A5. Below we discuss the possible factors contributing to biases in the simulated SWE and SCF including potential biases in the meteorological forcing datasets.

5.1 Impacts of meteorological forcing datasets on modelled SWE

Evaluation based on measurements from snow course and airborne gamma data indicates that the magnitude of SWE bias and uRMSE seen in CLASSIC are comparable to those from other gridded SWE products and LSMs (Brown et al., 2018; Mortimer et al., 2024; Cho et al., 2022) intended to represent historical snow conditions. However, for all three choices of meteorological forcing SWE is underestimated in mountainous regions (Fig. 4a) and overestimated in flat regions (Fig. 4b) throughout the snow season (with subsequent impacts on SCF). Since SCF is directly linked to SWE in the SL12 scheme (see Eq.1 and Eq. 2), these SWE biases can exert a large impact on simulated SCF in the fall and spring seasons in the model (limited impact during the peak SWE period because SCF is usually saturated). The consistent SCF biases shown in Figure 5 are linked to these consistent SWE biases for all three forcing choices in the model.

Naively, the bias-adjustments applied to temperature and precipitation in both the CRUJRA and GSWP3W5 forcing data might be expected to result in more accurate simulations. Yet among the three choices of forcing we used, the unadjusted ERA5 data yielded the lowest bias when evaluating the simulated SWE in ~~both~~ mountainous ~~and flat~~ regions (Fig. 4, Fig. A3~~1~~). In mountain regions, this discrepancy may result because the CRU and GPCP data used to adjust the precipitation values are biased towards locations with less precipitation (e.g. outside of regions with orographic features; e.g. Nijssen et al., 2001; Adler et al, 2003; Shi et al., 2017).

24

Mountain precipitation underestimation was also linked to negative SWE biases based on precipitation observations from the Snowpack Telemetry stations over western U.S. (Cho et al. 2022).

In NH flat regions, precipitation values from CRU and GPCP are expected to be more accurate than in mountainous regions (Adler et al., 2003), so it is less clear why GSWP3W5 has a much larger SWE bias despite having a precipitation bias similar to ERA5. The fact that GSWP3W5 is colder in flat regions compared to the other two forcings could play a role (Fig. 3a). This may reduce its ability to simulate mid-season ablation events (e.g., Brown et al., 2006; Slater et al., 2001) and/or alter the timing and location of snowfall. The reason that GSWP3W5 is colder than CRUJRA is also not immediately clear since both products use CRU TS4 for bias-adjusting their temperature (see Section 2.3.2). Differences between the interpolation and bias-adjustment methods may be responsible for the differences since they are more complex for GSWP3W5 (see Cucchi et al., 2020 and Weedon et al., 2010) than CRUJRA (Harris, 2023). For example, a constant lapse rate of 6.5 K km^{-1} was applied to temperature correction in GSWP3W5 but not in CRUJRA.

These results highlight that there is uncertainty in the accuracy of both temperature and precipitation forcing even when bias-adjusted to observations. These uncertainties can propagate to uncertainty in simulated SWE directly through precipitation amounts or in the case of temperature through phase partitioning of rainfall versus snowfall or direct melt. Even with perfectly constrained bias-adjustments for temperature and precipitation individually, there may still be spread in simulated SWE stemming from uncertainties in the joint distribution of temperature and precipitation that determines when snowfall occurs. Although measurements from snow course and airborne gamma data used in this study can better sample the subgrid-scale variability than a single-point measurement, we acknowledge that there are still uncertainties in our evaluation results, e.g. in situ sites may be biased towards locations with more snow cover.

5.2 Factors contributing to residual bias in modelled SCF

Although SCF overestimation in the mountainous regions is much reduced by the SL12 parameterization compared to the CTL parameterization (Fig. 5a – 5c and 5g), there are still areas with notable SCF biases. For example, much of the western NA mountainous areas have negative biases during the spring with the SL12 parameterization (Fig. 6d and 6f). Furthermore, in flat areas, all model runs overestimate SCF (Fig. 5d – 5f). These remaining SCF biases may be at least partly attributable to SWE underestimation in mountainous regions and SWE overestimation in flat regions (Fig. 4). The fact that in flat regions, there are larger SWE biases (Fig. 4b) and correspondingly larger SCF overestimation (Fig. 5d – 5f) in the model runs forced by GSWP3W5 supports this argument (see Section 5.1). Below we present some evidence on the link between differences in meteorological forcing datasets and choices of parameter values in the SL12 parameterization and the bias in modelled SCF.

762 Overall NH performance for model runs driven by ERA5 is comparable or slightly better than
763 the runs driven by CRUJRA in terms of simulated SWE and SCF (Fig. 3, Fig. 5, and Table 2),
764 while model runs driven by GSWP3W5 are worse everywhere except for HMA. In HMA, there
765 is significant SCF overestimation in model runs forced by CRUJRA and ERA5, while model
766 runs forced by GSWP3W5 have comparable SCF to observations (Fig. 5g and Table 3). For
767 model runs forced by ERA5, this is consistent with the cold temperature bias and large
768 precipitation overestimation in ERA5 (Fig. 3c and 3d). However, CRUJRA and GSWP3W5
769 exhibit similar biases in temperature and precipitation (Fig. 3c and 3d), yet model runs forced by
770 them have contrasting SCF biases (Fig. 5g). Therefore, biases in temperature and precipitation
771 cannot explain the SCF biases here. Instead, we found that the number of wet days (days with
772 precipitation >= 0.1 mm) differs in each of the three datasets, especially in the HMA region (Fig.
773 1$\underline{1}$$\theta$). Figure 1$\underline{1}$$\theta$ shows that on average ERA5 has near-daily precipitation events in the
774 mountainous areas (e.g. Tian Shan, Hindu Kush–Karakoram, and Himalayas) and southeast of
775 TP, while GSWP3W5 has the fewest wet days over the whole HMA region, especially over the
776 interior TP. The number of wet days in CRUJRA falls between the other two. This is consistent
777 with differences in the SCF annual cycles (Fig. 5g) and the SCF bias patterns (Fig. 7) found
778 among the three sets of model runs, suggesting that the different number of wet days in the
779 forcings contributes most to the difference in modelled SCF in this region. This conclusion is
780 also consistent with findings in previous studies (Liu et al., 2022; Orsolini et al., 2019), which
781 suggested that excessive snowfall in ERA5 contributes to overestimation of SND, SWE, and
782 SCF across HMA. In CLASSIC, the large number of wet days in ERA5 would lead to prolonged
783 periods with fresh snow and therefore high snow albedo. In coupled simulations this could lead
784 to or reinforce an existing cold bias. GSWP3W5 also has a smaller number of wet days in some
785 other regions of the globe, such as the middle to high latitudes of NA and eastern Siberia (not
786 shown).

787
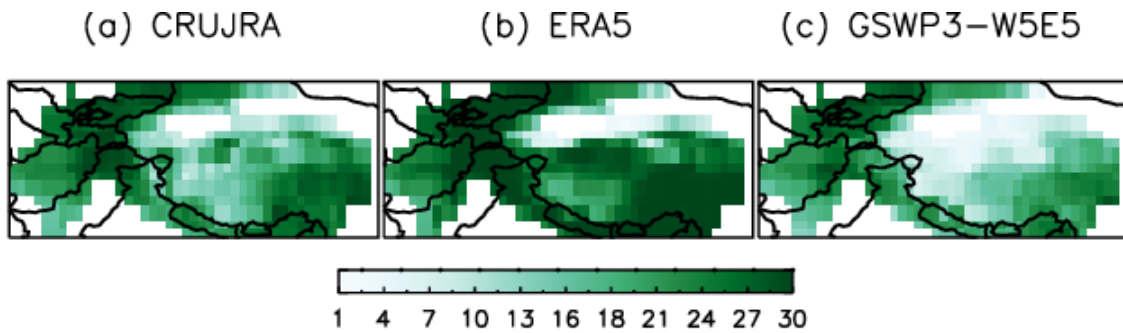
788 **Figure 1$\underline{1}$$\theta$.** The monthly mean number of wet days (days with total Pr >= 0.1 mm) in (a) CRUJRA, (b)
789 ERA5, and (c) GSWP3W5 during the main snow season (Sep – May) in HMA over the 2005-2014
790 period.

791

792 Besides biases in the meteorological datasets, the choice of parameter values in the SL12
793 parameterization can also contribute to uncertainties in modelled SCF. As illustrated in Swenson
794 and Lawrence (2012, their Fig. 7), choosing a larger k_{acc} parameter in Eq. (1) would result in
795 faster SCF increase with SND during accumulation events. All the previously discussed

26

796　simulations have used the default value of 0.1 for this parameter. We also performed sensitivity
797　experiments where the k_{acc} parameter was changed to 0.18 and 0.26. In these simulations, SCF
798　increases faster with SND especially in the fall, thereby resulting in higher SCF over NH
799　mountainous regions during that time of year. Notably, increasing k_{acc} to 0.26 produces less
800　biased SCF values during the fall (similar to those seen in the CTL simulations) while still
801　maintaining the improvements already presented during winter and spring (Fig. A63).

802　Likewise, the ablation portion of the SL12 parameterization (Eq. (2)) can be altered via the N_{melt}
803　parameter, which controls the rate at which SCF decreases as a function of SND. SCF decreases
804　faster with (normalized) SND in mountainous areas (small N_{melt}) than flat areas (large N_{melt}, Fig.
805　9 in Swenson and Lawrence, 2012). We adjusted the N_{melt} parameter by increasing the numerator
806　in Eq. (3) from 200 to 300, thereby increasing the N_{melt} value in mountain regions for the same
807　value of sub-grid topographic variability and resulting in slower SCF decrease. Results of the test
808　run show reduced SCF bias in the NA mountains in the spring compared to simulations with the
809　default N_{melt} value (Fig. A74).

810　The adjustments to k_{acc} and N_{melt} parameters described above provide ways to fine-tune the
811　agreement in simulated SCF with observations. However, because none of the three
812　meteorological forcing datasets used in this study are exempt from biases, there is a limit to how
813　well optimal parameter values can be chosen for use in CLASSIC. In addition, it may not be
814　ideal to over-tune the model to a specific observational estimate which may still have
815　uncertainties (Section 5.3).

816

817　**5.3 Other uncertainties**

818　SCF derived from satellite optical sensors such as MODIS represents the viewable snow cover
819　from space during cloud-free overpasses (i.e., from above the canopy). Dense forests and steep
820　terrain may obscure the MODIS sensor's view of snow-covered ground, leading to
821　underestimation of SCF (Hall et al., 2002; Marchane et al., 2015). For example, Stillinger et al.
822　(2023) found a consistent negative bias of approximately 10% under intermediate canopy cover
823　when comparing MODIS SCF with high-resolution airborne lidar data in parts of the western
824　U.S. The SCF overestimation in flat regions (Fig. 5d–f) may be partially attributable to this
825　underestimation by MODIS. However, as noted by Riggs et al. (2019), snow commission errors,
826　often related to residual cloud contamination, are among the most common sources of error in
827　MODIS snow products. As a result, the SCF derived from MODIS in this study may be subject
828　to both underestimation and overestimation.

829　While the IMS snow system primarily relies on visible satellite imagery, it also incorporates
830　surface station observations and passive microwave data. Therefore, SCF derived from IMS is
831　generally less affected by cloud cover and forest canopy than that from MODIS. Previous studies
832　have shown that IMS tends to report higher SCF than MODIS (e.g., Brown et al., 2010), which is
833　consistent with our results (Fig. 5). Nevertheless, SCF estimates from MODIS and IMS are
834　largely consistent across all regions except the HMA, suggesting that our evaluation results are
835　reasonably robust despite known uncertainties.

In LSMs, snow depth is typically diagnosed from SWE and snow density. As a result, uncertainties in modeled snow density can propagate to uncertainties in SCF, particularly when the SCF parameterization depends on snow density and/or snow depth, as demonstrated by Abolafia-Rosenzweig et al. (2024). In CLASSIC, these uncertainties influence SCF simulated by the control parameterization but do not directly affect SCF in the SL12 parameterization (Section 2.2). Since our focus is on the SL12 parameterization in this study, we do not explore this issue further.

Additional uncertainties may arise from the elevation data used to compute the standard deviation of sub-grid topography (σ_{topo}), particularly related to its spatial resolution. We compared σ_{topo} derived from two elevation datasets: ETOPO1 (1-arc-minute resolution, used in this study) and ETOPO2022 (15-arc-second resolution). The results indicate that the differences are limited in spatial extent and are primarily concentrated along edges of mountain ranges. To assess the impact on model results, we conducted a test simulation using σ_{topo} derived from ETOPO2022 and compared the simulated SCF with that from a run based on σ_{topo} derived from ETOPO1. The maximum difference in SCF between the two runs was less than 5% (not shown). These findings suggest that the resolution of the elevation data has a limited effect on the calculation of sub-grid topographic variability and simulated SCF, consistent with sensitivity tests reported by Lalande et al. (2023).

6. Conclusions

Our results demonstrate that implementing the SL12 parameterization in CLASSIC improves simulated SCF in mountainous regions. This confirms that the lack of topographic dependency in the current default parameterization is at least partly responsible for the SCF overestimation and cold bias in the coupled model configuration, CanESM5 (Lalande et al., 2021; Swart et al. 2019; Sigmond et al., 2023). The improved simulation of SCF also improves the simulation of surface albedo, which in turn leads to improved simulation of the surface radiation, energy fluxes, and water cycle in CLASSIC.

The results also demonstrate that the choice of meteorological forcing data can have a large impact on snow simulation in offline LSM runs. Based on our analysis, we suggest that at least part of the SWE underestimation in mountainous areas and SWE overestimation in flat areas can be linked to relative biases in temperature and precipitation from the meteorological forcing datasets. The SWE biases then propagate to biases in modelled SCF. In addition, we highlighted that bias-adjustment methods that improve temperature or precipitation separately may not result in more accurately simulated SWE, with consequences for downstream components of the water and energy cycles related to snow. These meteorological forcing datasets are regularly used to drive LSMs in various projects, such as the Global Carbon Project and ISIMIP, but for snow simulations it is important to better understand how inaccuracies in temperature and precipitation can propagate to errors in modelled SWE and SCF.

Based on the evaluation results presented in this study along with preliminary test results in fully coupled CanESM runs, the SL12 parameterization has been adopted in CLASSIC and will be

876 used in CanESM simulations for CMIP7 submission. Future work will focus on the evaluation of
877 the SL12 parameterization in fully coupled CanESM simulations where a full analysis of
878 feedbacks will be possible.

879

891

892 **References**

893 Abolafia‑Rosenzweig, R., He, C., Chen, F., & Barlage, M: Evaluating and enhancing snow
894 compaction process in the Noah‑MP land surface model, Journal of Advances in modeling
895 Earth Systems, 16, e2023MS003869, https://doi.org/10.1029/2023MS003869, 2024.
896

897 Adler, R. F., Huffman, G. J., Chang, A., Ferraro, R., Xie, P., Janowiak, J., Rudolf, B., Schneider,
898 U., Curtis, S., Bolvin, D., Gruber, A., Susskind, J., Arkin, P., & Nelkin, E.: The Version-2 Global
899 Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979–Present),
900 Journal of Hydrometeorology, 4(6), 1147-1167. https://doi.org/10.1175/1525-
901 7541(2003)004%3C1147:TVGPCP%3E2.0.CO;2, 2003.

902 Arora, V. K. and Boer, G. J.: A parameterization of leaf phenology for the terrestrial ecosystem
903 component of climate models, Glob. Chang. Biol., 11, 39–59, 2005.

904 Arora, V. K., Scinocca, J. F., Boer, G. J., Christian, J. R., Denman, K. L., Flato, G. M., Kharin, V.
905 V., Lee, W. G., and Merryfield, W. J.: Carbon emission limits required to satisfy future
906 representative concentration pathways of greenhouse gases, Geophys. Res. Lett., 38, L05805,
907 https://doi.org/10.1029/2010GL046270, 2011.

908 AUER, A. H.: The rain versus snow threshold temperatures. Weatherwise, 27: 67, 1974.

909 Bailey, W. G., Saunders, I. R., Bowers, J. D., and Verseghy, D. L.: Application of the Canadian
910 land surface scheme to a full canopy crop during a drying cycle, Atmos.-Ocean, 38, 57–80,
911 https://doi.org/10.1080/07055900.2000.9649640, 2000.

Bair, E. H., Dozier, J., Stern, C., LeWinter, A., Rittger, K., Savagian, A., Stillinger, T., and Davis, R. E.: Divergence of apparent and intrinsic snow albedo over a season at a sub-alpine site with implications for remote sensing, The Cryosphere, 16, 1765–1778, https://doi.org/10.5194/tc-16-1765-2022, 2022.

Bartlett, P.A., MacKay, M. D., and Verseghy, D. L.: Modified snow algorithms in the Canadian Land Surface Scheme:Model runs and sensitivity analysis at three boreal forest stands. Atmos.–Ocean, 44, 207–222, doi:10.3137/ao.440301, 2006.

Betts, A. K., Viterbo, P., Beljaars, A., Pan, H. L., Hong, S. Y., Goulden, M., and Wofsy, S: Evaluation of land-surface interaction in ECMWF and NCEP/NCAR reanalysis models over grassland (FIFE) and boreal forest (BOREAS). J. Geophys. Res., 103 (D18), 23079–23085, doi:10.1029/98JD02023, 1998.

Bolch, T., Shea, J. M., Liu, S., Azam, F. M., Gao, Y., Gruber, S., Immerzeel, W. W., Kulkarni, A., Li, H., Tahir, A. A., Zhang, G., and Zhang, Y.: Status and Change of the Cryosphere in the Extended Hindu Kush Himalaya Region, in: The Hindu Kush Himalaya Assessment: Mountainous, Climate Change, Sustainability and People, edited by: Wester, P., Mishra, A., Mukherji, A., and Shrestha, A. B., Springer International Publishing, Cham, 209–255, 2019.

Brown, R., Brasnett, B., and Robinson, D: Gridded North American monthly snow depth and snow water equivalent for GCM evaluation. ATMOSPHERE-OCEAN, 41: 1–14, 2003.

Brown, R., Bartlett, P. A., MacKay, M., and Verseghy, D. L.: Evaluation of snow cover in CLASS for SnowMIP. Atmos.–Ocean, 44, 223–238, doi:10.3137/ao.440302, 2006.

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

Brown, R., Tapsoba, D., and Derksen, C.: Evaluation of snow water equivalent datasets over the Saint-Maurice river basin region of southern Québec, Hydrol. Process., 32, 2748–2764, 2018.

Carroll, T.R. Airborne Gamma Radiation Snow Survey Program: A user's guide, Version 5.0. National Operational Hydrologic Remote Sensing Center (NOHRSC), Chanhassen, 14, 2001.

Cho, E., Vuyovich, C. M., Kumar, S. V., Wrzesien, M. L., Kim, R. S., and Jacobs, J. M.: Precipitation biases and snow physics limitations drive the uncertainties in macroscale modeled snow water equivalent, Hydrol. Earth Syst. Sci., 26, 5721–5735, https://doi.org/10.5194/hess-26-5721-2022, 2022.

Clark, M. P., Hendrikx, J., Slater, A. G., Kavetski, D., Anderson, B., Cullen, N.J., Kerr, T., Hreinsson, E. O., and Woods, R. A.: Representing spatial variability of snow water equivalent in hydrologic and land-surface models: A review, Water Resour. Res., 47, W07539, doi:10.1029/2011WR010745, 2011.

Compo, G.P., Whitaker, J.S., Sardeshmukh, P.D., Matsui, N., Allan, R.J., Yin, X., Gleason, B.E., Vose, R.S., Rutledge, G., Bessemoulin, P., Brönnimann, S., Brunet, M., Crouthamel, R.I., Grant,

A.N., Groisman, P.Y., Jones, P.D., Kruk, M.C., Kruger, A.C., Marshall, G.J., Maugeri, M., Mok, H.Y., Nordli, Ø., Ross, T.F., Trigo, R.M., Wang, X.L., Woodruff, S.D. and Worley, S.J.: The Twentieth Century Reanalysis Project. Q.J.R. Meteorol. Soc., 137: 1-28, https://doi.org/10.1002/qj.776, 2011

Cucchi, M., Weedon, G. P., Amici, A., Bellouin, N., Lange, S., Müller Schmied, H., Hersbach, H. and Buontempo, C.: WFDE5: bias-adjusted ERA5 reanalysis data for impact studies. Earth System Science Data, 12, 2097–2120, 2020.

Dai, A. and Trenberth, K. E.: Estimates of Freshwater Discharge from Continents: Latitudinal and Seasonal Variations, J. Hydrometeorol., 3, 660–687, 2002.

Danabasoglu, G., Lamarque, J.-F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., Emmons, L. K., Fasullo, J. T., García, R. R., Gettelman, A., Hannay, C., Holland, M. M., Large, W. G., Lauritzen, P. H., Lawrence, D. M., Lenaerts, J. T. M., Lindsay, K., Lipscomb, W. H., Mills, M. J., Neale, R., Oleson, K. W., Otto-Bliesner, B., Phillips, A. S., Sacks, W., Tilmes, S., Kampenhout, L., Vertenstein, M., Bertini, A., Dennis, J., Deser, C., Fischer, C., Fox-Kemper, B., Kay, J. E., Kinnison, D., Kushner, P. J., Larson, V. E., Long, M. C., Mickelson, S., Moore, J. K., Nienhouse, E., Polvani, L., Rasch, P. J., Strand, W. G., ... Lawrence, D.: The Community Earth System Model Version 2 (CESM2). Journal of Advances in Modeling Earth Systems, 12(2). https://doi.org/10.1029/2019MS001916, 2020.

Dickinson, R.E., Henderson-Sellers, A., Kennedy, P. J., Wilson, M. F.: NCAR Technical Note: Surface physics parameterization package for the NCAR Community Climate Model. National Center for Atmospheric Research, Boulder, Colorado, 1986.

Dirmeyer, P. A., Gao, X., Zhao, M., Guo, Z., Oki, T. and Hanasaki, N.: GSWP-2: Multimodel Analysis and Implications for Our Perception of the Land Surface. Bulletin of the American Meteorological Society, 87(10), 1381–98, 2006.

Douville, H., Royer, J. F., and Mahfouf, J. F.: A new snow parameterization for the Météo-France climate model, Clim. Dynam.,12, 37–52, https://doi.org/10.1007/BF00208761, 1995.

Du, Z. and Qingsong, Z.: Introduction, in: Mountainous Geoecology and Sustainable Development of the Tibetan Plateau, Chap. 1, Springer, Dordrecht, 1–17, ISBN 978-94-010-3800-3, https://doi.org/10.1007/978-94-010-0965-2_1, 2000.

Elder, K., Dozier, J., and Michaelsen, J.: Snow accumulation and distribution in an alpine watershed, Water Resour. Res., 27, 1541–1552, doi:10.1029/91WR00506, 1991.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. Geoscientific Model Development, 9(5), 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016, 2016.

Fletcher, C. G., Kushner, P. J., Hall, A., and Qu, X.: Circulation responses to snow albedo feedback in climate change. Geophys. Res. Lett., 36, L09702, doi:10.1029/2009GL038011, 2009.

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Hauck, J., Landschützer, P., Le Quéré, C., Li, H., Luijkx, I. T., Olsen, A., Peters, G. P., Peters, W., Pongratz, J., Schwingshackl, C., Sitch, S., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S. R., Arneth, A., Arora, V., Bates, N. R., Becker, M., Bellouin, N., Berghoff, C. F., Bittig, H. C., Bopp, L., Cadule, P., Campbell, K., Chamberlain, M. A., Chandra, N., Chevallier, F., Chini, L. P., Colligan, T., Decayeux, J., Djeutchouang, L. M., Dou, X., Duran Rojas, C., Enyo, K., Evans, W., Fay, A. R., Feely, R. A., Ford, D. J., Foster, A., Gasser, T., Gehlen, M., Gkritzalis, T., Grassi, G., Gregor, L., Gruber, N., Gürses, Ö., Harris, I., Hefner, M., Heinke, J., Hurtt, G. C., Iida, Y., Ilyina, T., Jacobson, A. R., Jain, A. K., Jarníková, T., Jersild, A., Jiang, F., Jin, Z., Kato, E., Keeling, R. F., Klein Goldewijk, K., Knauer, J., Korsbakken, J. I., Lan, X., Lauvset, S. K., Lefèvre, N., Liu, Z., Liu, J., Ma, L., Maksyutov, S., Marland, G., Mayot, N., McGuire, P. C., Metzl, N., Monacci, N. M., Morgan, E. J., Nakaoka, S.-I., Neill, C., Niwa, Y., Nützel, T., Olivier, L., Ono, T., Palmer, P. I., Pierrot, D., Qin, Z., Resplandy, L., Roobaert, A., Rosan, T. M., Rödenbeck, C., Schwinger, J., Smallman, T. L., Smith, S. M., Sospedra-Alfonso, R., Steinhoff, T., Sun, Q., Sutton, A. J., Séférian, R., Takao, S., Tatebe, H., Tian, H., Tilbrook, B., Torres, O., Tourigny, E., Tsujino, H., Tubiello, F., van der Werf, G., Wanninkhof, R., Wang, X., Yang, D., Yang, X., Yu, Z., Yuan, W., Yue, X., Zaehle, S., Zeng, N., and Zeng, J.: Global Carbon Budget 2024, Earth Syst. Sci. Data, 17, 965–1039, https://doi.org/10.5194/essd-17-965-2025, 2025.

Frolking, S., Milliman, T., McDonald, K., Kimball, J., Zhao, M., & Fahnestock, M.: Evaluation of the SeaWinds scatterometer for regional monitoring of vegetation phenology. Journal of Geophysical Research, 111, D17302. doi:10.1029/2005JD006588, 2006.

Ganji, A., Sushama, L., Verseghy, D., and Harvey, R.: On improving cold region hydrological processes in the Canadian Land Surface Scheme, Theor. Appl. Climatol., 127, 45–59, https://doi.org/10.1007/s00704-015-1618-4, 2015.

Ghiggi, G., Humphrey, V., Seneviratne, S. I., and Gudmundsson, L.: GRUN: an observation-based global gridded runoff dataset from 1902 to 2014, Earth Syst. Sci. Data, 11, 1655–1674, https://doi.org/10.5194/essd-11-1655-2019, 2019.

Hall, D. K. and Riggs, G. A.: MODIS/Terra Snow Cover Monthly L3 Global 0.05Deg CMG, Version 61. [Indicate subset used]. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. https://doi.org/10.5067/MODIS/MOD10CM.061, 2021. [Last access September 2021].

Hao, S., Jiang, L., Shi, J., Wang, G., Liu, X.: Assessment of MODIS Based Fractional Snow Cover Products Over the Tibetan Plateau. IEEE J Sel Top Appl Earth Observations Remote Sens12:533–548, 2019.

Harris, I., Osborn, T. J., Jones, P., and Lister, D.: Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset, Sci. Data, 7, 109, https://doi.org/10.1038/s41597-020-0453-3, 2020.

Harris, I.C.: CRU JRA v2.4: A forcings dataset of gridded land surface blend of Climatic Research Unit (CRU) and Japanese reanalysis (JRA) data; Jan.1901 - Dec.2022. NERC EDS

Centre for Environmental Data Analysis, last access Sep. 2023.
https://catalogue.ceda.ac.uk/uuid/aed8e269513f446fb1b5d2512bb387ad/, 2023.

He, C., P. Valayamkunnath, M. Barlage, F. Chen, D. Gochis, R. Cabell, T. Schneider, R. Rasmussen, G.-Y. Niu, Z.-L. Yang, D. Niyogi, and M. Ek: The Community Noah-MP Land Surface Modeling System Technical Description Version 5.0. NCAR Technical Note NCAR/TN-575+STR, doi:10.5065/ew8g-yr95, 2023.

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

Hengl, T., Mendes de Jesus, J., Heuvelink, G. B. M., Ruiperez Gonzalez, M., Kilibarda, M., Blagoti´c, A., Shangguan, W., Wright, M. N., Geng, X., Bauer-Marschallinger, B., Guevara, M. A., Vargas, R., MacMillan, R. A., Batjes, N. H., Leenaars, J. G. B., Ribeiro, E., Wheeler, I., Mantel, S., and Kempen, B.: SoilGrids250m: Global gridded soil information based on machine learning, PLOS ONE, 12, 1–40, https://doi.org/10.1371/journal.pone.0169748, 2017

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Q. J. Roy. Meteor. Soc., online first, https://doi.org/10.1002/qj.3803, 2020.

Hobeichi, S., Abramowitz, G. and Evans, J.: Conserving Land–Atmosphere Synthesis Suite (CLASS), Journal of climate, 33(5), pp. 1821–1844, 2020.

Jennings, K. S., Winchell, T. S., Livneh, B., & Molotch, N. P.: Spatial variation of the rain–snow temperature threshold across the Northern Hemisphere. Nature Communications, 9(1), 1–9. https://doi.org/10.1038/s41467-018-03629-7, 2018.

Jung, M., Koirala, S., Weber, U., Ichii, K., Gans, F., Camps-Valls, G., et al. : The FLUXCOM ensemble of global land-atmosphere energy fluxes. Scientific Data, 6(1), 74. https://doi.org/10.1038/s41597-019-0076-8, 2019.

Kato, S., Loeb, N. G., Rose, F. G., Doelling, D. R., Rutan, D. A., Caldwell, T. E., Yu, L., and Weller, R. A.: Surface Irradiances Consistent with CERES-Derived Top-of-Atmosphere Shortwave and Longwave Irradiances, J. Climate, 26, 2719–2740, 2013.

Kim, H.: Global Soil Wetness Project Phase 3 Atmospheric Boundary Conditions (Experiment 1) [Data set]. Data Integration and Analysis System (DIAS). https://doi.org/10.20783/DIAS.501, 2017.

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., Onogi, K., Kamahori, H., Kobayashi, C., Endo, H., Miyaoka, K., Takahashi, K.: The JRA-55 Reanalysis: General

Specifications and Basic Characteristics. J. Met. Soc. Jap., 93(1), 5-48, https://dx.doi.org/10.2151/jmsj.2015-001, 2015.

Lalande, M., Ménégoz, M., Krinner, G., Naegeli, K., and Wunderle, S.: Climate change in the High Mountainous Asia in CMIP6, Earth System Dynamics, 12, 1061–1098, https://doi.org/10.5194/esd-12-1061-2021, 2021.

Lalande, M., Ménégoz, M., Krinner, G., Ottlé, C., and Cheruy, F.: Improving climate model skill over High Mountainous Asia by adapting snow cover parameterization to complex-topography areas, The Cryosphere, 17, 5095–5130, https://doi.org/10.5194/tc-17-5095-2023, 2023.

Lange, S.: Trend-preserving bias adjustment and statistical downscaling with ISIMIP3BASD (v1.0), Geosci. Model Dev., 12, 3055–3070, https://doi.org/10.5194/gmd-12-3055-2019, 2019.

Lange, S., Menz, C., Gleixner, S., Cucchi, M., Weedon, G.P., Amici, A., Bellouin, N., Müller-Schmied, H., Hersbach, H., Buontempo, C., Cagnazzo, C.: WFDE5 over land merged with ERA5 over the ocean (W5E5 v2.0), ISIMIP Repository, https://doi.org/10.48364/ISIMIP.342217, 2021.

Lange, S., Mengel, M., Treu, S., Büchner, M: ISIMIP3a atmospheric climate input data (v1.1), ISIMIP Repository, https://doi.org/10.48364/ISIMIP.982724.1, 2022.

Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., et al.: The Community Land Model version 5: Description of new features, benchmarking, and impact of forcing uncertainty. Journal of Advances in Modeling Earth Systems, 11(12), 4245–4287. https://doi.org/10.1029/2018ms001583, 2019.

Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Hauck, J., Pongratz, J., Pickers, P. A., Korsbakken, J. I., Peters, G. P., Canadell, J. G., Arneth, A., Arora, V. K., Barbero, L., Bastos, A., Bopp, L., Chevallier, F., Chini, L. P., Ciais, P., Doney, S. C., Gkritzalis, T., Goll, D. S., Harris, I., Haverd, V., Hoffman, F. M., Hoppema, M., Houghton, R. A., Hurtt, G., Ilyina, T., Jain, A. K., Johannessen, T., Jones, C. D., Kato, E., Keeling, R. F., Goldewijk, K. K., Landschützer, P., Lefèvre, N., Lienert, S., Liu, Z., Lombardozzi, D., Metzl, N., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S., Neill, C., Olsen, A., Ono, T., Patra, P., Peregon, A., Peters, W., Peylin, P., Pfeil, B., Pierrot, D., Poulter, B., Rehder, G., Resplandy, L., Robertson, E., Rocher, M., Rödenbeck, C., Schuster, U., Schwinger, J., Séférian, R., Skjelvan, I., Steinhoff, T., Sutton, A., Tans, P. P., Tian, H., Tilbrook, B., Tubiello, F. N., van der Laan-Luijkx, I. T., van der Werf, G. R., Viovy, N., Walker, A. P.,Wiltshire, A. J., Wright, R., Zaehle, S., and Zheng, B.: Global Carbon Budget 2018, Earth Syst. Sci. Data, 10, 2141–2194, https://doi.org/10.5194/essd-10-2141-2018, 2018.

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

Liu, Y., Fang, Y., and Margulis, S. A.: High Mountainous Asia UCLA Daily Snow Reanalysis, Version 1, Boulder, Colorado USA, NASA National Snow and Ice Data Center Distributed Active Archive Center [data set], https://doi.org/10.5067/HNAUGJQXSCVU, 2021a.

1103   Liu, Y., Fang, Y., and Margulis, S. A.: Spatiotemporal distribution of seasonal snow water
1104   equivalent in High Mountainous Asia from an 18-year Landsat–MODIS era snow reanalysis
1105   dataset, The Cryosphere, 15, 5261–5280, https://doi.org/10.5194/tc-15-5261-2021, 2021b.

1106   Liu, Y., Fang, Y., Li, D., and Margulis, S. A.: How well do global snow products characterize
1107   snow storage in High Mountainous Asia? Geophysical Research Letters, 49, e2022GL100082.
1108   https://doi.org/10.1029/2022GL100082, 2022.

1109   Margulis, S. A., Liu, Y., and Baldo, E.: A Joint Landsat- and MODIS-Based Reanalysis Approach
1110   for Midlatitude Montane Seasonal Snow Characterization, Front. Earth Sci., 7, 1–23,
1111   https://doi.org/10.3389/feart.2019.00272, 2019.

1112   Marshall, S. and Oglesby, R. J.: An improved snow hydrology for GCMs. Part 1: snow cover
1113   fraction, albedo, grain size, and age, Clim. Dynam., 10, 21–37,
1114   https://doi.org/10.1007/BF00210334, 1994.

1115   Melton, J. R. and Arora, V. K.: Competition between plant functional types in the Canadian
1116   Terrestrial Ecosystem Model (CTEM) v. 2.0, Geosci. Model Dev., 9, 323–361,
1117   https://doi.org/10.5194/gmd-9-323-2016, 2016.

1118   Melton, J. R., Verseghy, D. L., Sospedra-Alfonso, R., and Gruber, S.: Improving permafrost
1119   physics in the coupled Canadian Land Surface Scheme (v.3.6.2) and Canadian Terrestrial
1120   Ecosystem Model (v.2.1) (CLASS-CTEM), Geosci. Model Dev., 12, 4443–4467,
1121   https://doi.org/10.5194/gmd-12-4443-2019, 2019.

1122   Melton, J. R., Arora, V. K., Wisernig-Cojoc, E., Seiler, C., Fortier, M., Chan, E., and Teckentrup,
1123   L.: CLASSIC v1.0: the open-source community successor to the Canadian Land Surface Scheme
1124   (CLASS) and the Canadian Terrestrial Ecosystem Model (CTEM) – Part 1: Model framework
1125   and site level performance, Geosci. Model Dev., 13, 2825–2850, https://doi.org/10.5194/gmd-13-
1126   2825-2020, 2020.

1127   Miao, X., Guo, W., Qiu, B., Lu, S., Zhang, Y., Xue, Y., and Sun, S.: Accounting for Topographic
1128   Effects on Snow Cover Fraction and Surface Albedo Simulations Over the Tibetan Plateau in
1129   Winter, J. Adv. Model. Earth Sy., 14, e2022MS003035, https://doi.org/10.1029/2022MS003035,
1130   2022.

1131   Monteiro, D. and Morin, S.: Multi-decadal analysis of past winter temperature, precipitation and
1132   snow cover data in the European Alps from reanalyses, climate models and observational
1133   datasets, The Cryosphere, 17, 3617–3660, https://doi.org/10.5194/tc-17-3617-2023, 2023.

1134   Myneni, R. B., Keeling, C. D., Tucker, C. J., Asrar, G., and Nemani, R. R.: Increased plant
1135   growth in the northern high latitudes from 1981 to 1991, Nature, 386, 698–702,
1136   doi:10.1038/386698a0, 1997.

1137   Monteiro, D. and Morin, S.: Multi-decadal analysis of past winter temperature, precipitation and
1138   snow cover data in the European Alps from reanalyses, climate models and observational
1139   datasets, The Cryosphere, 17, 3617–3660, https://doi.org/10.5194/tc-17-3617-2023, 2023.

1140 Mortimer, C., Mudryk, L., Derksen, C., Luojus, K., Brown, R., Kelly, R., and Tedesco, M.:
1141 Evaluation of long-term Northern Hemisphere snow water equivalent products, The Cryosphere,
1142 14, 1579–1594, https://doi.org/10.5194/tc-14-1579-2020, 2020.

1143 Mortimer, C., Mudryk, L., Derksen, C., Brady, M., Luo jus, K., Venäläinen, P., Moisander, M.,
1144 Lemmetyinen, J., Takala, M., Tanis, C., and Pulliainen, J.: Benchmarking algorithm changes to
1145 the Snow CCI+ snow water equivalent product, Remote Sens. Environ., 274, 112988,
1146 https://doi.org/10.1016/j.rse.2022.112988, 2022.

1147 Mortimer, C., and Vionnet, V.: Northern Hemisphere historical in-situ Snow Water Equivalent
1148 dataset (NorSWE, 1979-2021) (2.0) [data set]. Zenodo.
1149 https://doi.org/10.5281/zenodo.14503592, 2024.

1150 Mortimer, C., Mudryk, L., Cho, E., Derksen, C., Brady, M., and Vuyovich, C.: Use of multiple
1151 reference data sources to cross-validate gridded snow water equivalent products over North
1152 America, The Cryosphere, 18, 5619–5639, https://doi.org/10.5194/tc-18-5619-2024, 2024.

1153 Mortimer, C. and Vionnet, V.: Northern Hemisphere in situ snow water equivalent dataset
1154 (NorSWE, 1979–2021), Earth Syst. Sci. Data Discuss. [preprint], https://doi.org/10.5194/essd-
1155 2024-602, in review, 2025.

1156 Mudryk, L., Mortimer, C., Derksen, C., Elias Chereque, A., and Kushner, P.: Benchmarking of
1157 snow water equivalent (SWE) products based on outcomes of the SnowPEx+ Intercomparison
1158 Project, The Cryosphere, 19, 201–218, https://doi.org/10.5194/tc-19-201-2025, 2025.

1159 Namazi, M., von Salzen, K., and Cole, J. N. S.: Simulation of black carbon in snow and its
1160 climate impact in the Canadian Global Climate Model, Atmos. Chem. Phys., 15, 10887–10904,
1161 https://doi.org/10.5194/acp-15-10887-2015, 2015.

1162 Nijssen, B., O'Donnell, G. M., Lettenmaier, D. P., Lohmann, D., and Wood, E. F.: Predicting the
1163 discharge of global rivers. J. Climate, 14, 3307–3323, 2001.

1164 Niu, G.-Y. and Yang, Z.-L.: An observation-based formulation of snow cover fraction and its
1165 evaluation over large North American river basins, J. Geophys. Res., 112, D21101,
1166 https://doi.org/10.1029/2007JD008674, 2007.

1167 NOAA National Geophysical Data Center: ETOPO1 1 Arc-Minute Global Relief Model, NOAA
1168 National Centers for Environmental Information, 2009. Accessed in Sep. 2021.

1169 Orsolini, Y., Wegmann, M., Dutra, E., Liu, B., Balsamo, G., Yang, K., de Rosnay, P., Zhu, C.,
1170 Wang, W., Senan, R., and Arduini, G.: Evaluation of snow depth and snow cover over the
1171 Tibetan Plateau in global reanalyses using in situ and satellite remote sensing observations, The
1172 Cryosphere, 13, 2221–2239, https://doi.org/10.5194/tc-13-2221-2019, 2019.

1173 Osterkamp, T. E., & Romanovsky, V. E.: Evidence for warming and thawing of discontinuous
1174 permafrost in Alaska. Permafrost and Periglacial Processes, 10(1), 17−37, 1999.

Pastorello, G., Papale, D., Chu, H., Trotta, C., Agarwal, D., Canfora, E., Baldocchi, D., and Torn, M.: A new data set to keep a sharper eye on land-air exchanges, Eos T. Am. Geophys. Union, 98, https://doi.org/10.1029/2017EO071597, 2017.

Qu, X., and Hall, A.: On the persistent spread in snow-albedo feedback, Clim. Dyn., 42, 69–81, doi:10.1007/s00382-013-1774-0, 2013.

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

Riggs, G.A., Hall, D.K. and Roman, M.O. 2019. MODIS Snow Products Collection 6.1 User Guide. NASA Goddard Space Flight Center, Greenbelt, MD. https://modis-snow-ice.gsfc.nasa.gov/uploads/snow_user_guide_C6.1_final_revised_april.pdf.

Roe, G. H., and Baker, M. B.: Why is climate sensitivity so unpredictable? Science, 318, 629–632, doi:10.1126/science.1144735, 2007.

Roesch, A., Wild, M., Gilgen, H., and Ohmura, A.: A new snow cover fraction parameterization for the ECHAM4 GCM, Clim. Dynam., 17, 933–946, https://doi.org/10.1007/s003820100153, 2001.

Schaaf, C., Wang, Z.: MCD43C3 MODIS/Terra+Aqua BRDF/Albedo Daily L3 0.05Deg CMG. NASA LP DAAC, http://doi.org/10.5067/MODIS/MCD43C3.006, 2015.

Scinocca, J. F., Kharin, V. V., Jiao, Y., Qian, M. W., Lazare, M., Solheim, L., Flato, G. M., Biner, S., Desgagne, M., and Dugas, B.: Coordinated Global and Regional Climate Modeling, J. Climate, 29, 17–35, https://doi.org/10.1175/JCLI-D-15-0161.1, 2016.

Seiler, C.: amber: Automated model benchmarking package for the Canadian land surface scheme, 2019.

Seiler, C., Melton, J. R., Arora, V. K., and Wang, L.: CLASSIC v1.0: the open-source community successor to the Canadian Land Surface Scheme (CLASS) and the Canadian Terrestrial Ecosystem Model (CTEM) – Part 2: Global Benchmarking, Geoscientific Model Development, https://doi.org/10.5194/gmd-14-2371-2021, 2021.

Shangguan, W., Hengl, T., Mendes de Jesus, J., Yuan, H., and Dai, Y.: Mapping the global depth to bedrock for land surface modeling, J. Adv. Model. Earth Syst., 9, 65–88, https://doi.org/10.1002/2016MS000686, 2017.

Shi, H., Li, T., Wei, J.: Evaluation of the gridded CRU TS precipitation dataset with the point raingauge records over the Three-River Headwaters Region, Journal of Hydrology, 548, 322-332, https://doi.org/10.1016/j.jhydrol.2017.03.017, 2017.

Sigmond, M., Anstey, J., Arora, V., Digby, R., Gillett, N., Kharin, V., Merryfield, W., Reader, C., Scinocca, J., Swart, N., Virgin, J., Abraham, C., Cole, J., Lambert, N., Lee, W.-S., Liang, Y., Malinina, E., Rieger, L., von Salzen, K., Seiler, C., Seinen, C., Shao, A., Sospedra-Alfonso, R.,

Wang, L., and Yang, D.: Improvements in the Canadian Earth System Model (CanESM) through systematic model analysis: CanESM5.0 and CanESM5.1, Geosci. Model Dev., 16, 6553–6591, https://doi.org/10.5194/gmd-16-6553-2023, 2023.

Slater, A. G., Schlosser, C. A., Desborough, C. E., Pitman, A. J., Henderson-Sellers, A., Robock, A., Vinnikov, K. Y., Entin, J., Mitchell, K., Chen, F., Boone, A., Etchevers, P., Habets, F., Noilhan, J., Braden, H., Cox, P. M., de Rosnay, P., Dickinson, R. E., Yang, Z.-L., Dai, Y.-J., Zeng, Q., Duan, Q., Koren, V., Schaake, S., Gedney, N., Gusev, Y. M., Nasonova, O. N., Kim, J., Kowalczyk, E. A., Shmakin, A. B., Smirnova, T. G., Verseghy, D.,Wetzel, P., and Xue, Y.: The representation of snow in land surface schemes: Results from PILPS 2(d), J. Hydrometeorol., 2, 7–25, 2001.

Lange, S., Menz, C., Gleixner, S., Cucchi, M., Weedon, G. P., Amici, A., Bellouin, N., Schmied, H. M., Hersbach, H., Buontempo, C., Cagnazzo, C.: WFDE5 over land merged with ERA5 over the ocean (W5E5 v2.0). ISIMIP Repository. https://doi.org/10.48364/ISIMIP.342217, 2021.

Stackhouse, Jr, P. W., Gupta, S. K., Cox, S. J., Zhang, T., Mikovitz, J. C., and Hinkelman, L. M.: The NASA/GEWEX surface radiation budget release 3.0: 24.5-year dataset, Gewex news, 21, 10–12, 2011.

Stillinger, T., Rittger, K., Raleigh, M. S., Michell, A., Davis, R. E., and Bair, E. H.: Landsat, MODIS, and VIIRS snow cover mapping algorithm performance as validated by airborne lidar datasets, The Cryosphere, 17, 567–590, https://doi.org/10.5194/tc-17-567-2023, 2023.

Su, F., Duan, X., Chen, D., Hao, Z., and Cuo, L.: Evaluation of the Global Climate Models in the CMIP5 over the Tibetan Plateau, J. Climate, 26, 3187–3208, https://doi.org/10.1175/JCLI-D-12-00321.1, 2013.

Swart, N. C., Cole, J. N. S., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., Anstey, J., Arora, V., Christian, J. R., Hanna, S., Jiao, Y., Lee, W. G., Majaess, F., Saenko, O. A., Seiler, C., Seinen, C., Shao, A., Sigmond, M., Solheim, L., von Salzen, K., Yang, D., and Winter, B.: The Canadian Earth System Model version 5 (CanESM5.0.3), Geosci. Model Dev., 12, 4823–4873, https://doi.org/10.5194/gmd-12-4823-2019, 2019.

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

Taylor, K. E., Stouffer, R. J., & Meehl, G. A.: An overview of CMIP5 and the experiment design. Bulletin of the American Meteorological Society, 93(4), 485–498, https://doi.org/10.1175/BAMS-D-11-00094.1, 2012.

U.S. National Ice Center: IMS Daily Northern Hemisphere Snow and Ice Analysis at 1 km, 4 km, and 24 km Resolutions. (G02156, Version 1). [Data Set]. Boulder, Colorado USA. National Snow and Ice Data Center. https://doi.org/10.7265/N52R3PMC, 2008. Date Accessed Oct. 22, 2024.

Verseghy, D. L.: CLASS – A Canadian Land Surface Scheme for GCMs, I. Soil model, Int. J. Climatol., 11, 111–133, https://doi.org/10.1002/joc.3370110202, 1991.

Verseghy, D., McFarlane, N., and Lazare, M.: Class – A Canadian land surface scheme for GCMs, II: Vegetation model and coupled runs, Int. J. Climatol., 13, 347–370, https://doi.org/10.1002/joc.3370130402, 1993.

Verseghy, D., Brown, R., & Wang, L.: Evaluation of CLASS Snow Simulation over Eastern Canada. Journal of Hydrometeorology, 18(5), 1205–1225, https://doi.org/10.1175/JHM-D-16-0153.1, 2017.

von Salzen, K., Scinocca, J. F., McFarlane, N. A., Li, J., Cole, J. N. S., Plummer, D., Verseghy, D., Reader, M. C., Ma, X., Lazare, M., and Solheim, L.: The Canadian Fourth Generation Atmospheric Global Climate Model (CanAM4). Part I: Representation of Physical Processes, Atmos.-Ocean, 51, 104–125, https://doi.org/10.1080/07055900.2012.755610, 2013.

Wang, L., MacKay, M., Brown, R., Bartlett, P., Harvey, R., and Langlois, A: Application of satellite data for evaluating the cold climate performance of the Canadian Regional Climate model over Québec, Canada, Journal of Hydrometeorology, 15, 614-630, https://doi.org/10.1175/JHM-D-13-086.1, 2014.

Wang, L., Arora, V. K., Bartlett, P., Chan, E., and Curasi, S. R.: Mapping of ESA-CCI land cover data to plant functional types for use in the CLASSIC land model, Biogeosciences, 20, 2265–2282, https://doi.org/10.5194/bg-20-2265-2023, 2023.

Wang, L., Mudryk, L., Melton, J. R., Mortimer, C., Cole, J., Meyer, G., Bartlett, P., & Lalande, M.: Impact of topography and meteorological forcing on snow simulation in the Canadian Land Surface Scheme Including Biogeochemical Cycles (CLASSIC). https://doi.org/10.5281/zenodo.15032447, 2025.

Weedon, G. P., Gomes, S., Viterbo, P., Österle, H., Adam, J. C., Bellouin, N., Boucher, O., and Best, M.: The WATCH Forcing Data 1958–2001: A meteorological forcing dataset for land surface and hydrological-models, Tech. rep., WATCH Technical Report 22, available at: http://www.eu-watch.org/publications/technical-reports (last access: 26 August 2020), 2010.

Weedon, G. P., Balsamo, G., Bellouin, N., Gomes, S., Best, M. J., and Viterbo, P.: The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERAInterim reanalysis data, Water Resour. Res., 50, 7505–7514, https://doi.org/10.1002/2014WR015638, 2014.

WMO (Ed.): Guide to instruments and methods of observation: Volume II - Measurement of Cryospheric Variables, 2018th ed., World Meteorological Organization, Geneva, WMO-No. 8, 52 pp., 2018.

Yang, Z.-L., Dickinson, R. E., Robock, A., and Vinnikov, K. Y.: Validation of the Snow Submodel of the Biosphere–Atmosphere Transfer Scheme with Russian Snow Cover and Meteorological Observational Data, J. Climate, 10, 353–373, https://doi.org/10.1175/1520-0442(1997)010<0353:VOTSSO>2.0.CO;2, 1997.

Yao, T., Thompson, L. G., Mosbrugger, V., Zhang, F., Ma, Y., Luo, T., Xu, B., Yang, X., Joswiak, D. R., Wang, W., Joswiak, M. E., Devkota, L. P., Tayal, S., Jilani, R., and Fayziev, R.: Third Pole

1289   Environment (TPE), Environmental Development, 3, 52–64,
1290   https://doi.org/10.1016/j.envdev.2012.04.002, 2012.

1291

1292

1293

1294

1295

1296

1297

1298

1299

1300

1301

1302

1303

1304

1305

1306

1307

1308

1309

1310

1311

1312

1313

1314

1315

1316

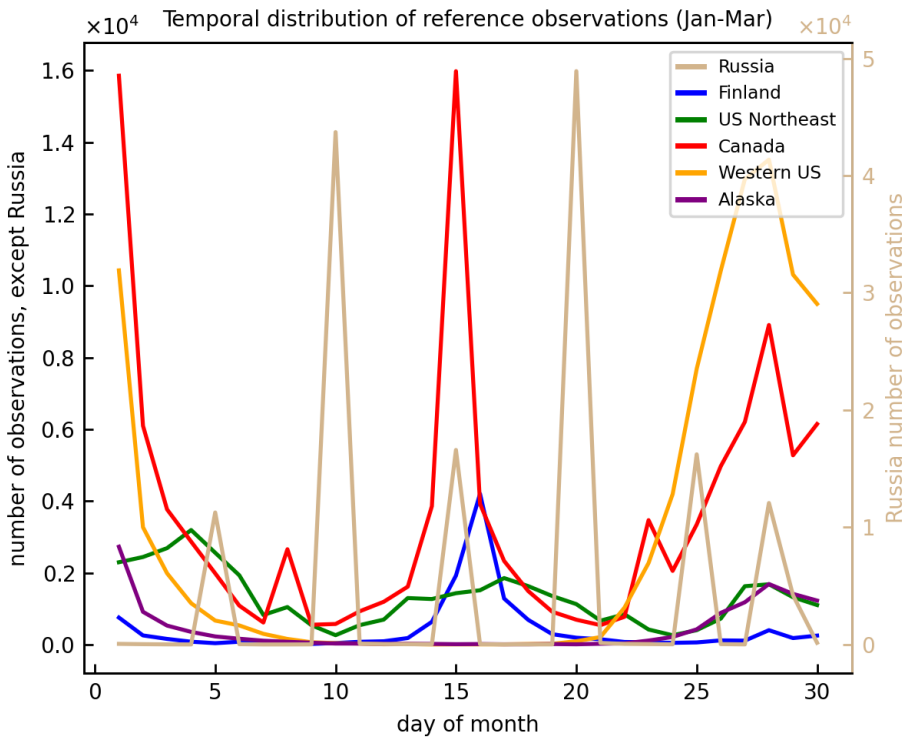1317

1318

Appendix A: Supplemental Figures

1320

1321

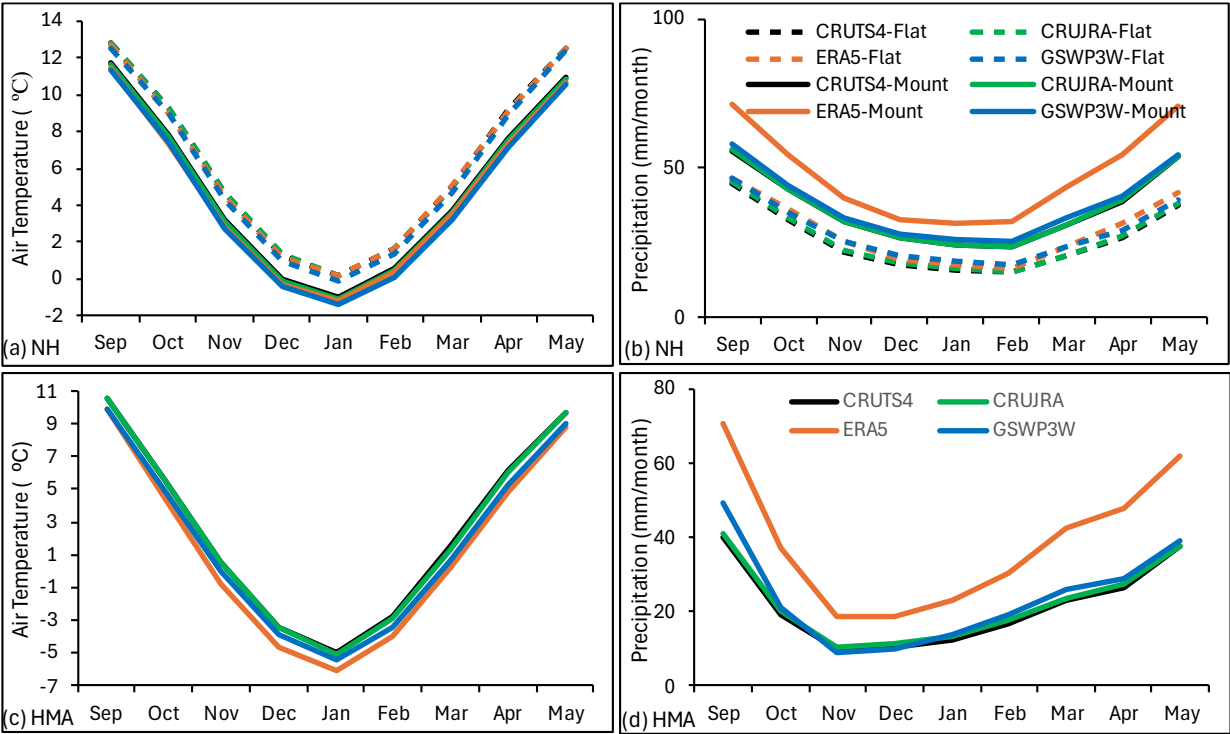**Figure A1**. Number of reference observations by network and day of the month during 1980-2014 for the Jan-March period.

1324

1325

1326

1327

1328

1329

1330

1331

1332

1333

1334

1335

1336

1337

Figure A2. Monthly mean air temperature (a and c) and precipitation (b and d) in the NH mountainous (solid line) and flat (dashed line) regions (a and b) and the HMA mountainous regions (c and d) over the 1980-2014 period.
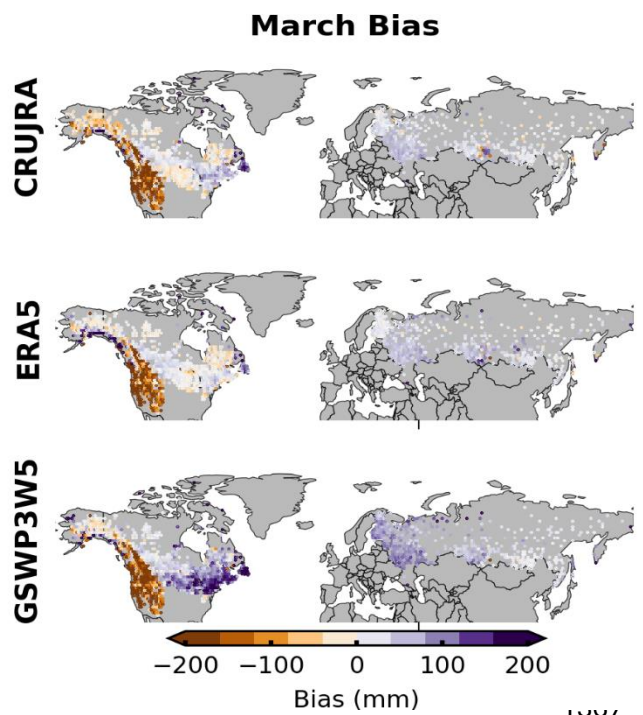
1338
1339
1340

1341

1342

1343

1344

1345

1346

1347

1348

1349

1350

1351

1352

1353

1354

March Bias

−200 −100 0 100 200
Bias (mm)

Figure A31. March SWE bias relative to in-situ measurements over the 1980-2014 period from model runs forced by each of the three meteorological forcings.
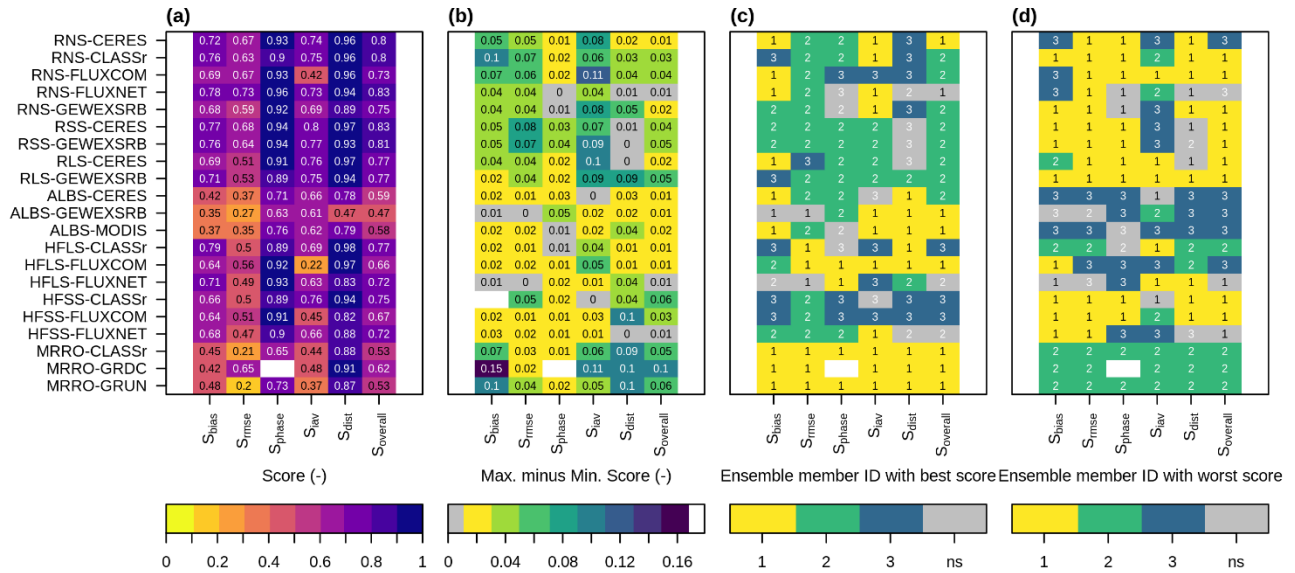
Figure A42. AMBER results for surface radiation, albedo, heat fluxes, and runoff from three model runs using the CTL parameterization, (a) mean ensemble score, (b) maximum score difference among ensemble members, (c) ensemble member with the highest score, and (d) ensemble member with the lowest score. Ensemble member IDs represent the following model runs: 1: CRUJRA-CTL, 2: ERA5-CTL, 3: GSWP3W5-CTL.
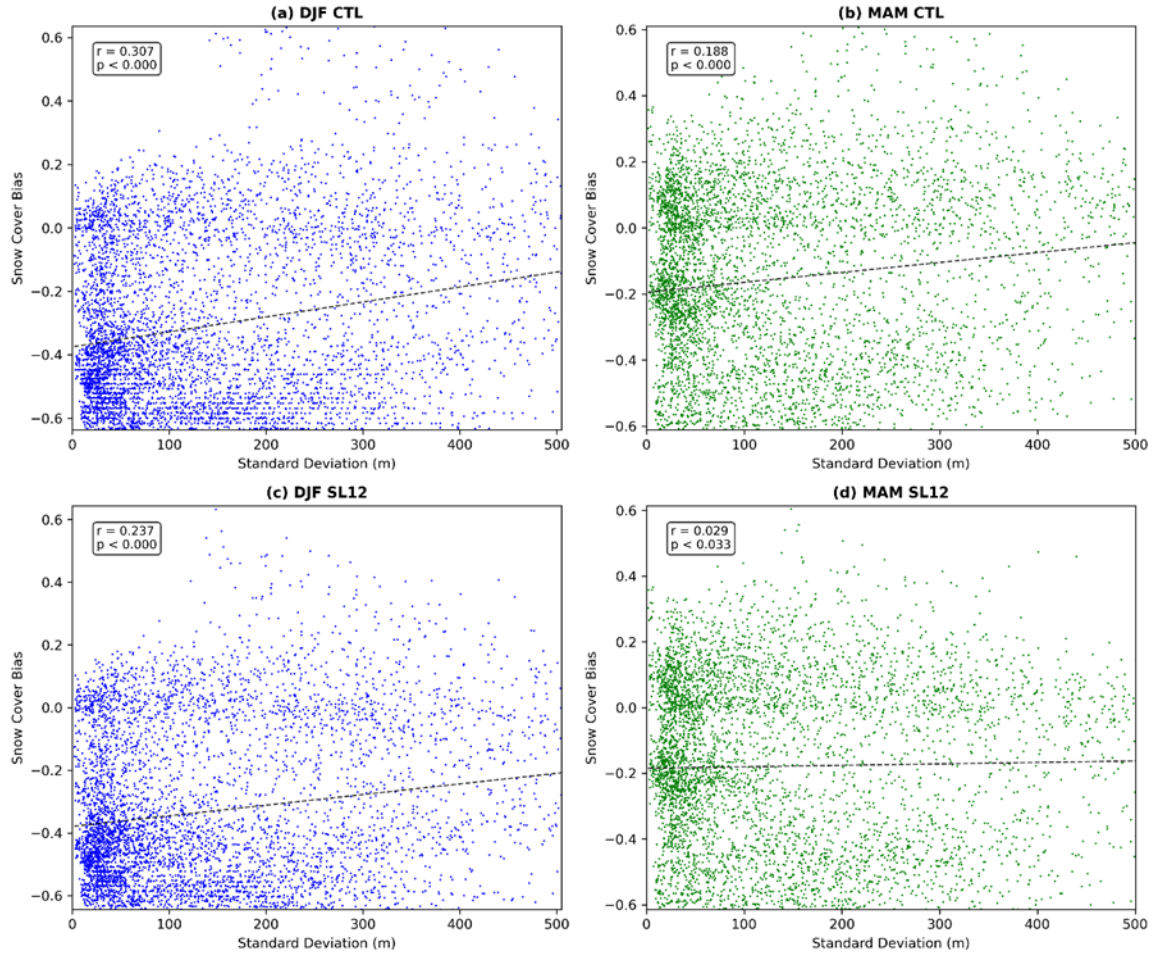
Figure A5. Scatter plots showing the relationship between SCF bias and the standard deviation of sub-grid topography during the winter (left) and spring (right) seasons for model runs using the CTL (top) and SL12 (Bottom) schemes forced by CRUJRA. The correlation coefficient (r) and p-value (using a two-tailed t-test) are provided in the upper-left corner of each plot.
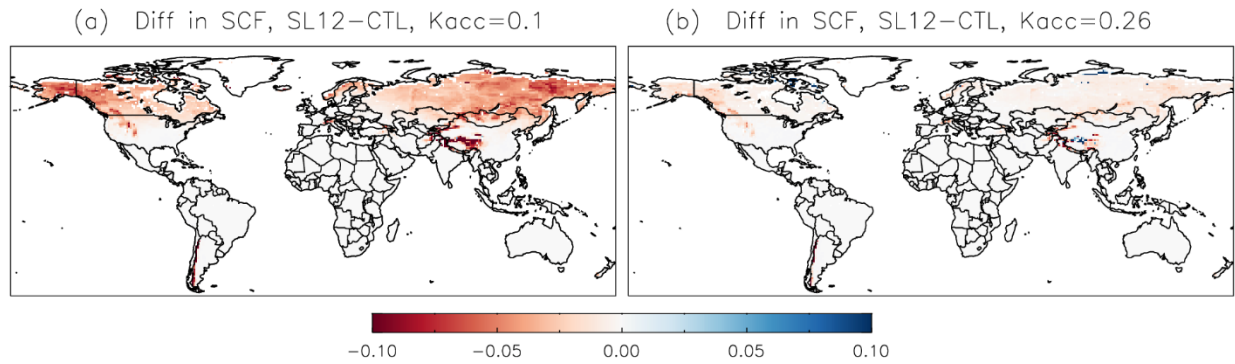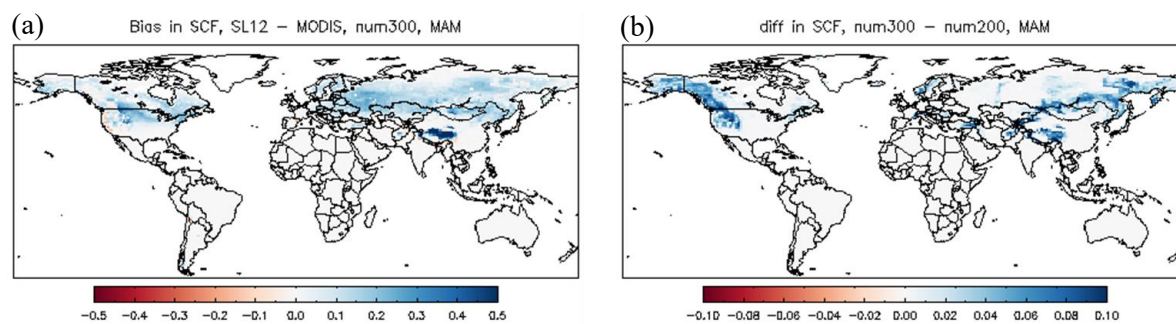
(a) Diff in SCF, SL12-CTL, Kacc=0.1   (b) Diff in SCF, SL12-CTL, Kacc=0.26

−0.10   −0.05   0.00   0.05   0.10

Figure A63. The difference in SCF between the SL12 and Control parameterizations during the fall (SON) in model runs using (a) k_{acc}=0.1, and (b) k_{acc}=0.26 for the SL12 parameterization.

1443

1444

Figure A74. (a) Spring (MAM) SCF bias relative to MODIS using an adjusted N_{melt} parameter (numerator=300 in Eq. 3), and (b) the difference in spring SCF in model runs using the adjusted (numerator=300) and default (numerator=200) N_{melt} parameter.