

Response to comments of editor:

We thank the editor for providing additional specific comments/edits on our paper that helped further improving our manuscript. Please find **the editors comments in bold**, our point-by-point answers without formatting, and *changes to the initial manuscript in Italics*.

Comment about replies

In your reply to P15 L13: “We have shown that the across-model spread in the representation of snow insulation effects”. How did you show that? I think I lost something while reading Sections 3.1 and 3.2

You answered “*This done in Figs.1-5 as well as in Table2.*”

Did you address this comment also by clarifying the text? Please give further details.

The across-model spread in the representation of the snow insulation effects is presented in both Sections 3.1 and 3.2. Therein the detailed discussion is with help of the Figures 1-5 as well as Table 2. This was a misunderstanding and we clarified this.

Comments about revised paper

Page 1, line 15: the author’s contribution should be added at the end of the paper, see also

http://www.the-cryosphere.net/for_authors/manuscript_preparation.html

We add one paragraph for author contribution before the “Acknowledgement”.

“Author contributions: A.R. developed the idea that lead to this paper; W.W. and A.R. performed research; J.C.M. contributed with ideas; A.R., J.C.M., W.W., and D.J. wrote the paper; and all authors discussed the results and commented on the paper at all stages.”

Page 2, line 5: sentence contains three times “bias”, please rephrase

Done. We revised this sentence with “Most of the model biases and across-model differences in simulating permafrost area are due to *inaccurate* atmospheric simulation e.g. of air temperature and precipitation, *deficient* simulation of snow and soil temperature and the coupling between atmosphere and land surface”.

Page 8, line 24: please correct sentence

Done. We revised this sentence to “The modal value of the *station-based* ΔT PDF is 5 K for “shallow” snow and 14 K for “thick” snow...”

Page 8, line 3: “Our analysis is focused on the common winter (DJF) condition, although snow can begin in November“. Snow cover build up can start as early as September. Please state why you choose November.

Yes, the first snow cover can build up in early autumn. The time of the first snow fall in the Arctic differ regionally and has a large inter-annual variability. But, the

intention of our sentence is just to emphasize that a different winter definition does not change our results. Therefore, we investigated the common winter definition (DJF) and also another winter definition (NDJFMA) find in literature. However, to be more precise, we revised this sentence to “Our analysis is focused on the common winter (DJF) condition, although snow can begin in November *or even earlier...*”

Reference list:

Please check the references for format, completeness, consistency; some journal page information is give as pp., some references have not a complete author list. (example McGuire, A.D., et al.: A model-based analysis of the vulnerability of carbon in the permafrost region between 1960 and 2009, Global Biogeochemical Cycles, in press, 2016)

Done.

Response to comments of reviewer:

We thank the reviewer for providing additional specific comments/edits on our paper that helped further improving our manuscript. Please find **the reviewers comments in bold**, our point-by-point answers without formatting, and *changes to the initial manuscript in Italics*.

1.“Many individual model studies have shown the strong impact of snow parameters on soil temperature simulations (e.g.,Langer et al.,2013; Dutra et al., 2012; Gouttevin et al., 2012; Essery et al., 2013; Wang et al., 2013; Jafarov et al., 2014).” Parameter or parametrization or properties?

In the sentence following the cited one we give the details about the investigated “snow parameters”: “Most importantly, these studies showed that the consideration of wet snow metamorphism and snow compaction, improved snow thermal conductivity and multi-layer snow schemes can improve the simulation of snow dynamics and soil temperature. Parameterizations that take into account snow compaction (e.g. related to overburden pressure, thermal metamorphism and liquid water) work better than simpler schemes such as an exponential increase of density with time (Dutra et al., 2010). The influence of snow thermal conductivity on soil temperature...”. We agree that “snow parameterization” describe this better. Accordingly, we have changed “snow parameter” to “snow *parameterization*”.

2.“Other models (LPJ-GUESS, MIROC-ESM, ORCHIDEE) strongly underestimate the increase of the Tsoil vs. Tair regression slope for decreasing snow depth.” Hard to understand.

We change the sentence with “*Other models (LPJ-GUESS, MIROC-ESM, ORCHIDEE) do not reproduce the much greater regression slope between Tsoil vs. Tair for shallow snow than for thick snow as the observations show.*”

3.“5 Permafrost area” This section is quite short compared to other sections. You may simply put to the above section and revise the section head as Discussion.

Thanks for the suggestion, but the structure as it stands seems clearer to us. This section “5 Permafrost area” stands for its own and does clearly not fit into the other sections. Yes, it is smaller than the other sections, but we do not see any problem with this. It is shorter because this part is not the main part of the paper’s analysis which focus on the air-soil temperature-snow relationships, but present the result that there is no clear relationship between the simulated permafrost area and the sophistication of the snow insulation in the model as someone could expect.

4.We accepted all the other specific text edits.

1 **Evaluation of air-soil temperature relationships**
2 **simulated by land surface models during winter across**
3 **the permafrost region**

4

5 **Wenli Wang¹⁺, Annette Rinke^{1,2+}, John C. Moore¹, Duoying Ji^{1*}, Xuefeng Cui³,**
6 **Shushi Peng^{4,17,18}, David M. Lawrence⁵, A. David McGuire⁶, Eleanor J. Burke⁷,**
7 **Xiaodong Chen²¹, Bertrand Decharme⁹, Charles Koven¹⁰, Andrew MacDougall¹¹,**
8 **Kazuyuki Saito^{12,15}, Wenxin Zhang^{13, 19}, Ramdane Alkama^{9,16}, Theodore J. Bohn⁸,**
9 **Philippe Ciais¹⁸, Christine Delire⁹, Isabelle Gouttevin⁴, Tomohiro Hajima¹²,**
10 **Gerhard Krinner^{4,17}, Dennis P. Lettenmaier⁸, Paul A. Miller¹³, Benjamin Smith¹³,**
11 **Tetsuo Sueyoshi¹⁴, Artem B. Sherstiukov²⁰**

12 ¹College of Global Change and Earth System Science, Beijing Normal University, Beijing,
13 China

14 ²Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research (AWI), Potsdam,
15 Germany

16 ³School of System Science, Beijing Normal University, Beijing, China

17 ⁴CNRS, LGGE, Grenoble, France

18 ⁵National Center for Atmospheric Research, Boulder, USA

19 ⁶U.S. Geological Survey, Alaska Cooperative Fish and Wildlife Research Unit, University of
20 Alaska Fairbanks, Fairbanks, AK, USA

21 ⁷Met Office Hadley Centre, Exeter, UK

22 ⁸School of Earth and Space Exploration, Arizona State University, Tempe, AZ, USA

23 ⁹GAME, Unité mixte de recherche CNRS/Meteo-France, Toulouse cedex, France

24 ¹⁰Lawrence Berkeley National Laboratory, Berkeley, CA, USA

25 ¹¹School of Earth and Ocean Sciences, University of Victoria, Victoria, BC, Canada

- 1 ¹²Japan Agency for Marine-Earth Science and Technology, Yokohama, Japan
- 2 ¹³Department of Physical Geography and Ecosystem Science, Lund University, Lund, Sweden
- 3 ¹⁴National Institute of Polar Research, Tachikawa, Japan
- 4 ¹⁵University of Alaska Fairbanks, Fairbanks, AK, USA
- 5 ¹⁶Institute for Environment and Sustainability (IES), Ispra, Italy
- 6 ¹⁷Université Grenoble Alpes, LGGE, Grenoble, France
- 7 ¹⁸LSCE, CEA/CNRS/UVSQ, Saclay, France
- 8 ¹⁹Center for Permafrost (CENPERM), Department of Geosciences and Natural Resource
9 Management, University of Copenhagen, Copenhagen, Denmark
- 10 ²⁰All-Russian Research Institute of Hydrometeorological Information-World Data Centre,
11 Obninsk, Russia
- 12 ²¹Department of Civil and Environmental Engineering, University of Washington, Seattle, WA,
13 USA
- 14 + *these authors contributed equally to this work*
- 15 * *correspondence to: Duoying Ji (duoyingji@bnu.edu.cn)*
- 16

1 **Abstract.** A realistic simulation of snow cover and its thermal properties are important for
2 accurate modelling of permafrost. We analyze simulated relationships between air and near-
3 surface (20 cm) soil temperatures in the Northern Hemisphere permafrost region during winter,
4 with a particular focus on snow insulation effects in nine land surface models and compare
5 them with observations from 268 Russian stations. There are large across-model differences in
6 the simulated differences between near-surface soil and air temperatures (T) (~~of~~ 3 to 14 $K^{\circ}C$),
7 in the sensitivity of soil to air temperature (0.13 to 0.96 $^{\circ}C/^{\circ}C$), and in the relationship between
8 T and snow depth. The observed relationship between T and snow depth can be used as a
9 metric to evaluate the effects of each model's representation of snow insulation, and hence
10 guide improvements to the model's conceptual structure and process parameterizations. Models
11 with better performance apply multi-layer snow schemes and consider complex snow processes.
12 Some models show poor performance in representing snow insulation due to underestimation
13 of snow depth and/or overestimation of snow conductivity. Generally, models identified as
14 most acceptable with respect to snow insulation simulate reasonable areas of near-surface
15 permafrost (13.19 to 15.77 million km²). However, there is not a simple relationship between
16 the sophistication of the snow insulation in the acceptable models and the simulated area of
17 Northern Hemisphere near-surface permafrost, because several other factors such as soil depth
18 used in the models, the treatment of soil organic matter content, ~~and~~ hydrology, and vegetation
19 cover also affect the ~~provide important controls on~~ simulated permafrost distribution.

1 **1 Introduction**

2 Present-day permafrost simulations by global climate models are limited and future
3 projections contain high, model-induced uncertainty (e.g., Slater and Lawrence, 2013; Koven
4 et al., 2013). Most of the model biases and across-model differences in simulating permafrost
5 area are due to ~~biased~~inaccurate atmospheric simulation e.g. of air temperature and
6 precipitation, ~~deficient~~biased simulation of snow and soil temperature, and the coupling
7 between atmosphere and land-surface. In winter, the snow insulation effect is a key process
8 for the air-soil temperature coupling. Its strength depends on the snow depth, areal coverage,
9 snow density and conductivity (see overview by Zhang, 2005). Many individual model
10 studies have shown the strong impact of snow ~~parameterizations~~parameters on soil
11 temperature simulations (e.g., ~~recently~~, Langer et al., 2013; Dutra et al., 2012; Gouttevin et al.,
12 2012; Essery et al., 2013; Wang et al., 2013; Jafarov et al., 2014). Most importantly, these
13 studies showed that the consideration of wet snow metamorphism and snow compaction,
14 improved snow thermal conductivity and multi-layer snow schemes can improve the
15 simulation of snow dynamics and soil temperature. Parameterizations that take into account
16 snow compaction (e.g. related to overburden pressure, thermal metamorphism and liquid
17 water) work better than simpler schemes such as an exponential increase of density with time
18 (Dutra et al., 2010). The influence of snow thermal conductivity on soil ~~regime~~temperature
19 has been demonstrated by many model studies (e.g., Bartlett et al., 2006; Saha et al., 2006;
20 Vavrus, 2007; Nicolsky et al., 2007; Dankers et al., 2011; Gouttevin et al., 2012). Winter soil
21 temperature can change by up to 20 K simply by varying the snow thermal conductivity by
22 0.1-0.5 W m⁻¹ K⁻¹ (Cook et al., 2008). The snow insulation effect also plays an important role
23 for the Arctic soil temperature response to climate change and therefore for future near-
24 surface permafrost thawing and soil carbon vulnerability (e.g., Schuur et al., 2008). Shallower
25 snow can reduce soil warming while shorter snow season can enhance soil warming
26 (Lawrence and Slater, 2010). The model skill in atmosphere-soil coupling with the
27 concomitant snow cover in the Arctic is an important factor in the assessment of limitations
28 and uncertainty of carbon mobility estimates (Schaefer et al., 2011).

29

30 The Snow Model Intercomparison Project (Snow MIP) (Essery et al., 2009) and the Project
31 for Intercomparison of Land-Surface Parameterization Schemes (PILPS) Phase 2e (Slater et
32 al., 2001) examined the snow simulations of an ensemble of land-surface schemes for the
33 mid-latitudes. However, until now there has been no attempt to evaluate the air-soil

1 temperature relationship in the Northern Hemisphere permafrost region and the detailed role
2 of snow depth therein across an ensemble of models. In such an investigation, a first suitable
3 approach is the evaluation of stand-alone (off-line) land surface models (LSMs). The
4 retrospective (1960-2009) simulations from the model integration group of the Permafrost
5 Carbon Network ("PCN"; <http://www.permafrostcarbon.org>) provide an opportunity to
6 evaluate an ensemble of nine state-of-the-art LSMs. Here, the LSMs are run with observation-
7 based atmospheric forcing, meaning that snow depth is not influenced by biases in the
8 atmospheric forcing in a coupled model set-up. The evaluation of the offline modeled air
9 temperature - snow depth - near-surface soil temperature relationship in winter is therefore
10 important for revealing a model's skill in representing the effects of snow insulation.

11

12 Most of the LSMs participating in PCN are the land-surface modules of Earth System Models
13 (ESMs) participating in the Coupled Model Intercomparison Project (CMIP5; [http://cmip-
14 pcmdi.llnl.gov/cmip5/](http://cmip-pcmdi.llnl.gov/cmip5/)) although in some cases different versions were used for PCN and
15 CMIP5 simulations. Thus, the results we present can guide the corresponding evaluation of
16 these ESMs, though analysis of coupled model results requires consideration of couplings
17 between model components and is necessarily more complex.

18

19 The scope of the present study is to analyze the extent to which the ensemble of PCN models
20 can reproduce the observed relationship between air and near-surface soil temperatures in the
21 Northern Hemisphere permafrost region during winter, with a particular focus on the snow
22 insulation effect. For the latter we analyze the impact of snow depth on the difference
23 between near-surface soil and air temperatures. Our related key questions are: How well do
24 the models represent the observed spatial pattern of the air-soil temperature difference in
25 winter and its control by the snow depth? What is the range of the simulated air-soil
26 temperature relationship across the model ensemble? To the extent possible, we try to relate
27 the performance ~~of the models~~ level to their ~~model's~~ snow schemes. With this aim in mind, a
28 simultaneous analysis of simulated air and near-surface soil temperatures, and snow depth is
29 presented and compared with those from a novel data set of Russian station observations. We
30 ~~used this data set focus here on a comprehensive Russian station data set~~ because it this has
31 been compiled within PCN, and it is hard to find other station data sets which provide
32 simultaneous observations of both air and soil temperatures as well as snow depth over a long
33 period.

34

1 In Sect. 2, we describe the model simulations, the station observations used for evaluation,
2 and the analysis methods. In Sect. 3, we present a detailed analysis of near-surface air
3 temperature - snow depth - soil temperature relationships in winter. In Sect. 4, we discuss the
4 roles of atmospheric forcing and model processes. In Sect. 5, we investigate the relation of
5 simulated snow insulation and permafrost area. We summarize our findings and present
6 conclusions in Sect. 6.

7 **2 Data and Analysis**

8 **2.1 Models**

9 We use data from nine LSMs participating in the PCN, including CLM4.5, CoLM, ISBA,
10 JULES, LPJ-GUESS, MIROC-ESM, ORCHIDEE, UVic, and UW-VIC. For detailed
11 information about the models and simulations we refer to Rawlins et al. (2015), Peng et al.
12 (2015), and Mc-Guire et al. (2016). The total soil depth for soil thermal calculations ranges
13 from 3 m (divided into 8 layers) in LPJ-GUESS to 250 m (divided into 14 layers) in UVic.
14 The soil physical properties differ among the models as well, and four of them (CLM4.5,
15 ISBA, UVic, UW-VIC) include organic horizons. Three models (ISBA, LPJ-GUESS, UW-
16 VIC) ~~did~~ do not archive soil sub-grid results and provide only area-weighted ground
17 temperature (i.e. averaged over wetlands and vegetated areas, and in some cases lake
18 fractions).

19

20 Table 1 lists relevant snow model details. One model (UVic) uses an implicit snow scheme
21 which replaces the upper soil column with snow-like properties, i.e. the near-surface soil layer
22 takes the temperature of the air-snow interface. The other models use separate snow layers on
23 top of the ground, either a single bucket (LPJ-GUESS, UW-VIC) or multi-layer snow
24 schemes (CLM4.5, CoLM, ISBA, JULES, MIROC-ESM, ORCHIDEE). Snow insulation is
25 explicitly considered in all models; increasing snow depth increases the insulation effect.
26 Many-Most models consider the effect of varying snow density on insulation (Table 1). This
27 is parameterized by a snow conductivity-density relationship ~~that describes how, as snow~~
28 ~~density increases, thermal conductivity increases, thereby reducing the snow insulation~~. Some
29 of the models (LPJ-GUESS, MIROC-ESM, ORCHIDEE, UVic) use a fixed snow density,
30 consider only dry snow and no compaction effects, while others represent liquid water in
31 snow and different processes for snow densification such as mechanical compaction, and
32 thermal and destructive metamorphism (Table 1).

33

1 The simulations were generally run for the period 1960-2009, although some simulations
2 were stopped a few years earlier. Each model team was free to choose appropriate driving
3 data sets for weather and climate, atmospheric CO₂, nitrogen deposition, disturbance, land
4 cover, soil texture, etc. However, the climate forcing data (surface pressure, surface incident
5 longwave and shortwave radiation, near-surface air temperature, wind and specific humidity,
6 rain and snowfall rates) are from gridded observational datasets (e.g. CRUNCEP, WATCH)
7 (SI Table 1). The exception is MIROC-ESM, which was run as a fully-coupled model, forced
8 by its own simulated climate. Mean annual air temperature ~~of the simulated by~~ MIROC-ESM
9 ~~simulations~~ for the permafrost region was within the range (-7.2 to 2.2°C) of the other
10 forcing data sets used in this study and the trend in near-surface air temperature (+0.03°C yr⁻¹)
11 was the same for all forcing data sets. However, MIROC-ESM had both the highest annual
12 precipitation (range 433 to 686 mm) and the highest trend in annual precipitation (range -2.1
13 to +0.8 mm yr⁻¹) among the forcing data sets.

14

15 The spatial domain of interest is the Northern Hemisphere permafrost land regions. Our
16 analysis is based on the 0.5° × 0.5° resolution gridded driving and modeled data for winter
17 (DJF) 1980-2000.

18 2.2 Observations

19 A quality-checked data set of monthly near-surface air temperature, 20 cm soil temperatures
20 and snow depth from Russian meteorological stations have been provided by the All-Russian
21 Research Institute of Hydrometeorological Information-World Data Centre (RIHMI-WDC;
22 <http://meteo.ru/>). 579 stations report snow depth and 268 stations provide simultaneous data
23 of all three variables. Ground surface temperature data are not available. A detailed
24 description of dataset preparation is provided in Sherstiukov (2012a). Observing conditions at
25 the Russian stations in all meteorological elements correspond with WMO standards. The
26 observations presented have been included in other data sets, such as the Global Summary of
27 the Day (GSOD) data set, HadSRUT4 etc., and are widely used in climate researches (e.g.
28 Anisimov and Sherstiukov, 2016; Decharme et al. 2016; Park et al., 2014; Brun et al., 2013;
29 Pavlov and Malkova, 2009; PaiMazumder et al., 2008). The soil temperature dataset was run
30 through four independent methods of quality control (Sherstiukov, 2012b). However, some
31 soil temperature observations could be disturbed by grass cutting during the warm season and
32 the removal of organic materials, mainly at agricultural sites, which may affect the trend in
33 warm season (Park et al., 2014), but this does not affect our results about the air - upper soil
34 temperature relationship in winter.

1
 2 Precipitation station data have been compiled from ~~the Global Summary of the Day (GSOD)~~
 3 data set produced by the National Climatic Data Center (NCDC; <http://www.ncdc.noaa.gov>)
 4 for all of the stations that are included in the RIHMI-WDC data set. In addition to the
 5 station's ground snow depth observations we use gridded snow water equivalent (SWE) data
 6 from the GlobSnow-2 product (<http://www.globsnow.info/swe/>), which has been produced
 7 using a combination of passive microwave radiometer and ground-based weather station data
 8 (Takala et al., 2011). Orographic complexity, vegetation cover, and snow state (e.g. wet snow)
 9 affect the accuracy of this product. When compared with ground measurements in Eurasia, the
 10 GlobSnow product shows root-mean-square error (RMSE) values of 30 to 40 mm for SWE
 11 values below 150 mm, with retrieval uncertainty increases when SWE is above this threshold
 12 (e.g., Takala et al., 2011; Muskett, 2012; Klehmet et al., 2013). To compare with station data,
 13 snow depth was then calculated from SWE using a snow density of 250 kg m^{-3} , which is a
 14 median observed value in winter. Zhong et al. (2013) report snow density values of 180-250
 15 kg m^{-3} for tundra/taiga and 156-193 kg m^{-3} for alpine snow classes. Woo et al. (1983) report
 16 snow density values of 250-400 kg m^{-3} for various terrain types. Choice of density does not
 17 materially affect the results.

18
 19 All these data have been compiled for winter (DJF) and the same time period of 1980-2000.
 20 This period was chosen because soil temperature data are sparse before 1980 and the JULES
 21 simulation stopped in the year 2000. Comparison of the simulations with the station data was
 22 done using a weighted bilinear interpolation from the 4 surrounding model grid points onto
 23 the station locations.

24 **2.3 Analysis Methods**

25 Our analysis is focused on the common winter (DJF) condition, although snow can begin in
 26 November or even earlier and end at the beginning of May, but we checked that a different
 27 winter definition (NDJFMA) does not qualitatively change any of the inter-variables
 28 relationships found. The focus in our study is on the evaluation of the simulated air-soil
 29 temperature relationships, modulated by snow depth. For this, we analyze the winter mean as
 30 well as the interannual variability (expressed as the standard deviation) of 4-four key variables:
 31 near-surface air temperature (T_{air}), near-surface soil temperature (soil temperature at 20 cm
 32 depth; T_{soil}), snow depth (d_{snow}), and the difference between T_{soil} and T_{air} . This difference T
 33 ($T = T_{soil} - T_{air}$) is called the air-soil temperature difference. By limiting our analysis to the
 34 winter only, we are able to attribute the across-model and model-to-observation differences in

1 T primarily to snow insulation effects. In winter, the effects of other factors (e.g. soil
 2 moisture, texture) on T are much smaller than that of snow. Ground surface temperatures
 3 were not recorded in the Russian data set, but 20 cm soil depth temperatures were. To test
 4 how sensitive are results using 20 cm temperatures instead of ground surface, we also
 5 analyzed model simulated temperature differences between ground surface and T_{air} , and found
 6 no qualitative differences, hence justifying use of 20 cm observations.

7

8 We use the Pearson product-moment correlation coefficient and its significance (von Storch
 9 and Zwiers, 1999) to investigate the co-variability between T and d_{snow} as well as between
 10 T_{soil} and its two forcing factors (T_{air} and d_{snow}). Before we compute the correlations we
 11 detrended the data by removing a least squares regression line. The calculated correlation
 12 maps (i.e. spatial distributions of correlation coefficients) based on model and observation
 13 data, allow the comparison of the spatial patterns of these relationships.

14

15 To further examine the functional behavior between the key variables, we present relation
 16 diagrams between pairs of variables (e.g. variation of T with change of d_{snow}). To evaluate
 17 the performance of the individual LSMs we calculate the RMSE between the observed and
 18 modeled relationships. We illustrate the dependence of T vs. d_{snow} and T_{soil} vs. d_{snow} relations
 19 for three T_{air} ranges. To distinguish dry snow pack regimes from those where sporadic melt
 20 may occur even in winter, we split T_{air} into ~~3~~three regimes: the coldest conditions (T_{air} -
 21 25°C ,) representing 24% of observations), the intermediate temperature conditions ($-25^{\circ}\text{C} <$
 22 T_{air} -15°C ,) representing 42% of the observations), and the warmest conditions ($-15^{\circ}\text{C} < T_{air}$
 23 -5°C ,) representing 34% of observations). Hence it is an indirect separation of temperature-
 24 gradient metamorphosis regimes and density-gradient metamorphosis snow pack regimes.
 25 Additionally, we present conditional probability density functions (PDFs) of T for different
 26 snow depth and air temperature regimes and compare the simulated PDFs with those obtained
 27 from station observations.

28 **3 Results**

29 **3.1 Relationship between air – soil temperature difference and snow depth**

30 The air-soil temperature difference (T) - snow depth (d_{snow}) relationship in winter (Fig. 1)
 31 shows in the Russian station observations an increase of T with increasing d_{snow} . The data
 32 exhibit a linear relation between T and d_{snow} at relatively shallow snow depths with a trend
 33 towards asymptotic behavior at thicker snow, which is in agreement with earlier findings

1 (Zhang, 2005; Ge and Gong, 2010; Morse et al., 2011). There is also significant scatter in the
 2 observation-based relationship indicated by the inter-quartile range in T of 1.5-8.5 K°C at
 3 specific snow depth and air temperature regimes, likely resulting from complicating factors
 4 such as snow pack density and moisture content variability over the winter, as well as
 5 observational errors.

6
 7 All models reproduce the observed relationship, i.e. increasing T with increasing d_{snow} .
 8 However, Fig. 1 also shows a wide across-model spread in the simulated relationships, and
 9 that some of the models are not consistent with the behavior in the observations. Only three
 10 models (CLM4.5, CoLM, JULES) reproduce reasonably well the observed T vs. d_{snow}
 11 relationship using a benchmark of RMSE $< 5 \text{K}^\circ\text{C}$ for all temperature regimes. In particular
 12 LPJ-GUESS, ORCHIDEE, UVic, UW-VIC, MIROC-ESM show large RMSE for cold air
 13 conditions. ISBA stands out overall, with a RMSE of 7-18 K°C in all temperature ranges. We
 14 conclude that these models do not adequately represent the features of the observed T vs.
 15 d_{snow} relationship. The scatter in the modeled relationships, indicated by the inter-quartile
 16 range, is of the same order as in the observations, except for ISBA and MIROC-ESM which
 17 produce noticeably smaller variations.

18
 19 Figure 2a views the T vs. d_{snow} relationship in a complementary form using the PDFs of T
 20 for different snow depth regimes. This analysis allows a detailed evaluation of the snow
 21 regime-dependent T separation by quantifying and comparing the modal value and width of
 22 the different conditional PDFs. Since the Russian snow depths are clearly non-Normal in
 23 distribution (SI Fig. 1, with a median d_{snow} of 30 cm), we divide the data into "shallow" (d_{snow}
 24 < 20 cm) and "thick" ($d_{snow} > 45$ cm) regimes to separate two snow depth regimes. The modal
 25 value of the ~~station-based station data~~ T PDF is 5 K°C for "shallow" snow and 14 K°C for
 26 "thick" snow - that is thick snow is a better insulator than thin snow. Based on the T PDFs,
 27 five models (CoLM, CLM4.5, JULES, ORCHIDEE, MIROC-ESM) successfully separate the
 28 T regimes under different snow depth conditions. Their simulated T PDFs have a smaller
 29 modal value for thin snow than for thick snow, like in the observations. The other models
 30 clearly fail in separating the T PDFs for the two different snow depth regimes. However,
 31 even for the five successful models, both the shapes and the modal values of the simulated
 32 PDFs differ from the observed PDF.

33

1 Both Figs. 1 and 2b further indicate that T are related to T_{air} conditions. This is expected due
 2 to the effects of T_{air} on snow pack properties, particularly its density and moisture content,
 3 that affect the thermal conductivity of the snow. For example, the density of fresh fallen
 4 snow tends to be much lower under cold T_{air} than warm (Anderson, 1976), leading to
 5 increased insulation (larger T). Snow densification is also a function of T_{air} , for example,
 6 depth hoar metamorphosis of the snow pack, which produces more insulation (loosely packed
 7 depth-hoar crystals have very low thermal conductivity), is promoted by strong thermal
 8 gradients in the snow pack, and is typical of continental climates (e.g., Zhang et al., 1996).
 9 Therefore, we can expect that the same thickness of snow in colder climates will provide
 10 greater insulation than it would in warmer climates.

11

12 Our analysis of observations (Figs. 1 and 2b) confirms i) a larger T for colder T_{air} than for
 13 warmer T_{air} (for a given snow depth), ii) a greater sensitivity of T to changes in d_{snow} in
 14 colder T_{air} (Fig. 1), and iii) larger modal value of the T PDF for colder T_{air} than for warmer
 15 T_{air} (21 K°C for $T_{air} < -25^{\circ}\text{C}$ and 9 K°C for $-15^{\circ}\text{C} < T_{air} < -5^{\circ}\text{C}$; Fig. 2b). These effects are
 16 consistent with colder climates having lower density snow packs, and the differences are in
 17 line with measurements of snow density variability (Zhong et al., 2013). Additionally, both
 18 the inter-quartile range in Fig. 1 and the width of the PDFs in Fig. 2b become larger as T_{air}
 19 cool. This may be related to the formation of depth hoar, which is a very good insulator and
 20 its varying presence in the snow pack decouples T from d_{snow} . Cold, thin snow packs tend to
 21 contain much more low density depth hoar than warmer snow packs (e.g., Zhang et al., 1996;
 22 Singh et al., 2011). Continental regions have large annual temperature cycles, with greater
 23 interannual variability and thinner snow packs, than maritime ones. This variability leads to
 24 greater scatter and greater sensitivity of the T vs. d_{snow} relationship in the cold winter regions.
 25 An additional cause of scatter is that the density of fresh-fallen snow decreases with the
 26 decrease of falling-temperature. Accordingly, we find in the cold T_{air} regime ($T_{air} < -25^{\circ}\text{C}$) a
 27 larger T in early winter (November-December) when the snow pack is composed of thin,
 28 low density fresh snow (and depth hoar) than in late winter (January-February) (SI Fig. 2).
 29 Under warm conditions ($-15^{\circ}\text{C} < T_{air} < -5^{\circ}\text{C}$) such a separation is not observed.

30

31 If we evaluate the models with respect to this observed impact of T_{air} ~~to~~ on the T vs. d_{snow}
 32 relationship, we demonstrate that some models (CLM4.5, CoLM, JULES) are better able to
 33 replicate the effect than others (LPJ-GUESS, MIROC-ESM, ORCHIDEE, UW-VIC) (Fig. 1).
 34 The latter do not fully replicate the larger T under cold T_{air} conditions. CLM4.5, CoLM and

1 JULES capture a larger T for colder T_{air} for a given d_{snow} in agreement with the observations.
 2 However, for shallow snow JULES simulates an increase of T with increasing d_{snow} for all
 3 temperature ranges that is twice as large as observations. Two models (ISBA, UVic) clearly
 4 fail in this evaluation. Poor model performance in reflecting T_{air} influence on the T vs. d_{snow}
 5 also manifests itself in regime separation of the PDFs (Fig. 2b). Some models do not separate
 6 the T regimes under different T_{air} conditions well or at all (ISBA, LPJ-GUESS, MIROC-
 7 ESM, UVic), while others cannot capture the observed cold temperature regime features (i.e.,
 8 too broad PDFs and shifts towards smaller modal values; ORCHIDEE, UW-VIC). The three
 9 models with reasonable inter-variable relations (CLM4.5, CoLM, JULES) also capture the
 10 regime separation in the PDFs. These three models as well as LPJ-GUESS and ORCHIDEE
 11 also represent the observed greater insulation of early winter snow packs under cold
 12 conditions (SI Fig. 2).

13

14 The maps of the T vs. d_{snow} correlations in winter (Fig. 3) demonstrates a pronounced spatial
 15 variability in the T vs. d_{snow} relationship. Highest positive correlation occurs in the region of
 16 the East Siberian Plain and Siberian High ~~lands~~Lands. In other regions, namely in
 17 Scandinavia, West Russian Arctic, West and Central Siberian Plains, the correlation is much
 18 weaker and often not statistically significant. These ~~are the~~ regions ~~have of large winter~~ snow
 19 ~~depth~~ (Sect. 4.1.2) ~~which are~~ influenced by North Atlantic cyclonic activity which brings
 20 relatively warm moist air and heavy precipitation in winter (and a positive correlation
 21 between d_{snow} and T_{air}), leading to relatively small mean T .

22

23 Some models (CLM4.5, CoLM, ORCHIDEE, UW-VIC) show a reasonable spatial pattern of
 24 correlation coefficient ($r \approx 0.4$) comparing to that of the ~~with~~ observations, while the others do
 25 not (Fig. 3). Obvious outliers are the LPJ-GUESS and UVic models, which do not reproduce
 26 the observed pattern of correlation. UVic calculates a reverse spatial pattern comparing to that
 27 of the observations (e.g. significant positive correlation in West Siberian Plain and Central
 28 Siberian Highlands). LPJ-GUESS produces very few statistically significant correlations.

29 3.2 Variability of soil temperature with air temperature and snow depth

30 Next we assess whether or not the models can correctly reproduce the interannual near-
 31 surface soil temperature (T_{soil}) variability in relation to snow depth (d_{snow}) and near-surface air
 32 temperature (T_{air}) variability. Previous studies have noted that the strength of relationship
 33 between T_{soil} and T_{air} is modulated by d_{snow} and the snow insulation effect increases only up to
 34 a limiting depth beyond which extra snow makes little difference to soil temperatures (Smith

1 and Riseborough, 2002; Sokratov and Barry, 2002; Zhang, 2005; Lawrence and Slater, 2010).
 2 Zhang (2005) reported that the limiting snow depth is approximately 40 cm.

3
 4 To inspect the difference of the insulation ~~capacity for shallow and thick snow effects on both~~
 5 ~~sides of such a limiting snow depth~~, we investigate the T_{soil} vs. T_{air} relationship under shallow
 6 ($d_{snow} = 20$ cm) and thick ($d_{snow} = 45$ cm) snow conditions. Our Russian observation analysis
 7 (Fig. 4, Table 2) indicate a three times higher regression slope between T_{soil} and T_{air}
 8 ($0.62^{\circ}\text{C}/^{\circ}\text{C}$, $R^2=0.8$) under shallow snow pack than thicker snow conditions ($0.21^{\circ}\text{C}/^{\circ}\text{C}$,
 9 $R^2=0.4$). This is consistent with observations that the mean freezing n-factor (the ratio of
 10 freezing degree days at the ground surface to air freezing degree days) is high at sites where
 11 the snow cover is thin or absent, and low at sites where the snow cover is thick (e.g., for
 12 Yukon Territory in Canada; Karunaratne and Burn, 2003).

13
 14 Figure 4 clearly shows that some models (CoLM, CLM45, JULES) can well capture this
 15 ~~difference modification of the T_{soil} vs. T_{air} relation by snow depth regime well~~. Their
 16 regression slopes for thick and thin snow are well separated and in agreement with those
 17 from the observed relationship (Table 2). The RMSE of their modeled T_{soil} vs. T_{air}
 18 relationships from observations is smaller than 4°C . These models better reproduce the
 19 observed T vs. d_{snow} relationship. Other models (LPJ-GUESS, MIROC-ESM, ORCHIDEE)
 20 ~~strongly underestimate the increase of the T_{soil} vs. T_{air} regression slope for decreasing snow~~
 21 ~~depth do not reproduce the much greater regression slope between T_{soil} vs. T_{air} for shallow~~
 22 ~~snow than for thick snow as the observations show~~. They also produce a regression slope for
 23 thick snow more than twice as large as observations. Two models (ISBA, UVic) ~~fail here and~~
 24 do not show any sensitivity in the T_{soil} vs. T_{air} relation to snow conditions (Fig.4, Table 2).
 25 Another measure quantitatively confirms the same models behavior: The observed average
 26 d_{snow} in the shallow snow regime is 13.7 cm and that for the thick snow regime is 58.5 cm, so
 27 we would expect, if near-surface T_{air} and conductivities were equal in both snow depth
 28 classes, a ratio between the slopes for shallow and thick snow ~~of would be~~ 4.3. CLM4.5,
 29 CoLM, and JULES reproduce this observed variation in the T_{soil} vs. T_{air} relation better than
 30 others (Table 2). JULES and CoLM indicate a factor of 4 change, while CLM4.5 indicates a
 31 factor of 2 change. Other models (LPJ-GUESS, MIROC-ESM, ORCHIDEE) underestimate
 32 the increase of the regression slope for decreasing snow depth; they simulate only a factor
 33 change of about 1.5. The two models with unrealistic T vs. d_{snow} relationships (ISBA, UVic)
 34 also fail in this evaluation of their T_{soil} vs. T_{air} relationship. They simulate a too strong

1 sensitivity of T_{soil} to T_{air} (regression slopes larger than $0.9^{\circ}\text{C}/^{\circ}\text{C}$, $R^2 > 0.7$; Table 2) that are
 2 almost completely independent of the snow depth regimes, particularly in ISBA, which is not
 3 consistent with observations. These models' spatial correlation patterns between T_{soil} and T_{air}
 4 also differ greatly from the observations and the other models (SI Fig. 3) and show very high
 5 positive correlation ($r > 0.8$) in most regions, as may be expected from the large regression
 6 slope shown in Fig. 4. The RMSE of their modeled T_{soil} vs. T_{air} relationships from
 7 observations reaches ca. 10°C .

8

9 The T_{soil} vs. d_{snow} relationship (Fig. 5) displays the variation of T_{soil} with changing snow depth
 10 and emphasizes the reduced sensitivity of T_{soil} to snow depth under thick snow conditions.
 11 With increasing d_{snow} , T_{soil} asymptotically converges towards a value of around 0°C . Overall,
 12 the Russian observations indicate that snow depth above about 80-90 cm has very little
 13 additional insulation effect on T_{soil} . Most models show consistent results with regard to this
 14 aspect, although the inter-quartile range of T_{soil} for specific snow depths is quite large in some
 15 models (ISBA, ORCHIDEE, UVic, UW-VIC) (Fig. 5). The figure further points to the air
 16 temperature dependency of the relation. On average, for a given d_{snow} , a colder T_{soil} is
 17 observed for colder near-surface air temperatures, compared with warmer air temperatures.
 18 Most models can replicate this effect of air temperatures on the T_{soil} vs. d_{snow} relationship,
 19 though with differing accuracy. The RMSE between the observed and modeled relationships
 20 can reach ca. 10°C ~~and~~ or more (in ISBA, UVic, UW-VIC), particularly under cold conditions.

21

22 The spatial patterns of the correlation coefficients between T_{soil} and T_{air} (SI Fig. 3) and
 23 between T_{soil} and d_{snow} (SI Fig. 4) show a relatively large across-model scatter in many regions.
 24 Obvious outliers in the T_{soil} vs. T_{air} correlation maps are ISBA and UVic which strongly
 25 overestimate the correlation ($r > 0.9$) over most of the Arctic. This indicates an
 26 underestimated snow insulation effect, and confirms the weak insulation in both models,
 27 which we already discussed based on their underestimated T (Fig. 1) and weak correlation
 28 between T and d_{snow} (Fig. 3). Other models (LPJ-GUESS, ORCHIDEE, UW-VIC) also
 29 overestimate the correlation in some regions (e.g. western Russian Arctic, $r > 0.7$). Most of
 30 the simulated maps of T_{soil} vs. d_{snow} correlation agree with the observations on a strong
 31 positive correlation in East Siberia. This is a region of relatively shallow snow (10-40 cm; Fig.
 32 6) and there T_{soil} is very sensitive to variations in snow depth (e.g., Romanovsky et al., 2007).
 33 Comparing both simulated correlation maps, it is obvious that in this region, T_{soil} correlates

1 more strongly with d_{snow} than with T_{air} , in agreement with the Russian data and earlier studies
 2 (Romanovsky et al., 2007; Sherstyukov, 2008).

3 **4 Roles of atmospheric forcing and model processes**

4 The across-model differences in the snow insulation effect, presented by the air temperature -
 5 snow depth - soil temperature relationships described above, are partially due to [the](#)
 6 differences in the atmospheric forcing data and also due to differences in the snow and soil
 7 physics used in the LSMs. However, because the climate forcing data sets utilized with each
 8 model are observation-based (except for MIROC-ESM), obvious outliers in individual model
 9 performance likely ~~mainly~~ indicate poor or deficient physical descriptions of the air/snow/soil
 10 relations in that specific LSM.

11 **4.1 Atmospheric forcing and snow depth**

12 **4.1.1 Air temperature and precipitation**

13 Both near-surface air temperature (T_{air}) and precipitation are given by the climate forcing data
 14 sets (SI Table 1) for all models, except for MIROC-ESM which simulates both. The across-
 15 model differences in forcing T_{air} used are relatively small and the simulated spatial patterns of
 16 temperature are very similar (SI Fig. 5). All forcing datasets are somewhat colder than
 17 Russian station data in their grid cells. The biases of winter mean T_{air} ranges from $-0.8 \text{ K}^\circ\text{C}$ to
 18 $-4.7 \text{ K}^\circ\text{C}$ (SI Table 2), reflecting biases in the climate forcing data used by the models. In
 19 contrast, MIROC-ESM has a positive (mean) T_{air} bias of $+2.7 \text{ K}^\circ\text{C}$.

20
 21 The large-scale patterns of precipitation are similar across the models, but regional differences
 22 can be large (SI Fig. 6). The individual differences in winter precipitation range from -0.2
 23 mm/day to $+0.5 \text{ mm/day}$ (SI Table 2) relative to the average of the Russian station data.
 24 Unfortunately, snowfall was archived in only a few models, however large-scale spatial
 25 patterns are similar across these models (SI Fig. 7).

26 **4.1.2 Snow depth**

27 The broad-scale spatial snow depth (d_{snow}) patterns are similar across the models and show
 28 general agreement with the observed patterns (Fig. 6). The well-pronounced areas of
 29 maximum winter d_{snow} (50-100 cm) are in Scandinavia, the Urals, the West Siberian Plain,
 30 Central Siberian Highlands, the Far East, Alaskan Rocky mountains, and Labrador Peninsula
 31 and isle of Newfoundland. However, large regional across-model variability is obvious. Some

1 models (JULES, LPJ-GUESS, ORCHIDEE, UVic) underestimate d_{snow} , while others
2 (CLM4.5, CoLM, ISBA, UW-VIC) overestimate it (Fig. 6; Table 3). The model biases are
3 quite similar with respect to station observations and GlobSnow data. It should be noted, that
4 the models do not account for snowdrift. However, redistribution of snow due to wind is an
5 important aspect, which makes comparison between in-situ measured and modeled snow
6 depths difficult (e.g., Vionnet et al., 2013; Sturm and Stuefer, 2013; Gisnas et al., 2014).

7
8 Precipitation/snowfall across-model differences cannot be the primary explanation of these
9 d_{snow} differences since some models (JULES, MIROC-ESM, ORCHIDEE) have positive bias
10 in precipitation (> 0.2 mm/d, SI Table 2) but simulate much lower d_{snow} compared to other
11 models (Fig. 6, SI Figs. 6, 7, Table 3). Across-model differences in the interannual variability
12 of winter precipitation do not translate simply to corresponding differences in the interannual
13 d_{snow} variability (not shown). For example, UVic calculates the (unrealistically) largest
14 interannual d_{snow} variability in the boreal Europe permafrost region which is not reflected in
15 the precipitation variability. These results indicate that the simulated snow depth is a function
16 of both the prescribed winter precipitation, and the model's snow energy and water balance.

17 **4.2 Model processes**

18 We have shown that the across-model spread in the representation of snow insulation effects
19 (Sects. 3.1, 3.2) can not predominantly be explained by differences in the forcing data (Sect.
20 4.1), but to a large extent is due to the representation of snow processes in the models. By
21 considering the relationship plots (Figs. 1, 4 and 5), and the conditional PDFs (Fig. 2) we
22 were able to ~~categorize~~classify the models in terms of their snow insulation performance. In
23 this section we discuss the influence of the different snow parameterizations in the models.

24
25 Models with better performance (CLM4.5, CoLM, JULES) apply multi-layer snow schemes.
26 This allows them to simulate more realistic (stronger) insulation because they consider the
27 snowpack's vertical structure and variability. They calculate the energy and mass balance in
28 each snow layer, are able to capture nonlinear profiles of snow temperature, and can also
29 account for thermal insulation within the snowpack such as when the upper layer thermally
30 insulates the lower layers (e.g., Dutra et al., 2012). These models also incorporate storage and
31 refreezing of liquid water within the snow, parameterize wet snow metamorphism, snow
32 compaction, and snow thermal conductivity (Table 1), which have been found to be among
33 the most important processes for good snow depth and surface soil temperature simulation
34 (e.g., Wang et al., 2013).

1

2 An underestimated snow depth directly leads to insulation that is too weak in JULES, LPJ-
3 GUESS, ORCHIDEE, and UVic (Fig. 6, Table 3). However only in ORCHIDEE and UVic
4 does this lead to a significant underestimation of T (Table 3, SI Fig. 8) indicating bias
5 compensation in the two other models. Thus, compensating error effects occur due to snow
6 density and conductivity (SI Fig. 9, Table 1), which impact snow thermal insulation.

7

8 Our analysis showed that two models (ISBA, UVic) have T_{soil} vs. T_{air} correlation that are too
9 high indicating that they do not represent the modulation of the T_{soil} vs. T_{air} relationship by
10 snow depth (Fig. 4). This is consistent with their underestimation of T (Figs. 1 and 2, SI Fig.
11 8, Table 3). In UVic, the snowpack is treated not as a separate layer but as an extension of the
12 top soil layer and a combined surface-to-soil thermal conductivity is calculated (Table 1).
13 Such a scheme largely negates or reduces the insulating capacity of snow (Slater et al., 2001).
14 Koven et al. (2013) noted that such a scheme simulates very little warming of soil, and
15 sometimes even cooling. The slightly underestimated snow depth (Table 3, Fig. 6) contributes
16 (but not as the primary factor) to reduced snow insulation, as reported for UVic (Avis, 2012).

17

18 ISBA strongly underestimates T , while strongly overestimating d_{snow} , compared with
19 observations (Table 3, Fig. 6). However, ISBA uses the same atmospheric forcing data as
20 JULES (accordingly the air temperature and precipitation are quite similar; SI Table 2). Also,
21 the model's snow density (150-250 kg m⁻³) is similar to other models (CLM45, CoLM,
22 JULES) (SI Fig. 9) and in agreement with Zhong et al. (2013) who report snow density values
23 of on 180-250 kg m⁻³ for tundra/taiga and 156-193 kg m⁻³ for alpine snow classes in winter.
24 This apparent contradiction comes from the parameterization of snow cover fraction within
25 each grid cell (SCF). The version of ISBA used here calculates a unique superficial soil
26 temperature whether or not the soil is covered by snow and all the energy and radiative fluxes
27 are area-weighted by SCF (equations 7 and 20 in *Douville et al.*, 1995). In order to get
28 reasonable albedos in snow-covered forests, as is necessary when ISBA is coupled to the
29 CNRM-CM climate model, the parameterization gives very low SCF in the boreal forest
30 (between 0.2 and 0.5). Hence, snow insulates only 20% to 50% of the grid cell, despite fairly
31 high snow depths. The heat fluxes from the snow-covered fraction are averaged with the
32 fluxes from the snow-free surface, strongly concealing the actual insulating effect of snow
33 and underestimating it over the grid cell. Using the detailed snow model Crocus (Brun et al.,
34 1992; Vionnet et al., 2012) with a SCF equal to 100% leads to an almost perfect simulation of

1 near-surface soil temperature over Northern Eurasia (Brun et al., 2013). A similar experiment
 2 with ISBA and a SCF equal to 100% (Decharme et al., 2016) leads to good performances
 3 showing that the low T in ISBA despite high snow depth in the present study is mostly due
 4 to this sub-grid snow fraction. Decharme et al. (2016) still showed that the ISBA results are
 5 further improved by updating the snow albedo and snow densification parameterization.

6
 7 Interestingly, the ORCHIDEE performance in simulating snow depth and T is similar to
 8 UVic (underestimation of d_{snow} and T ; Table 3). However, ORCHIDEE can better represent
 9 the observed T_{soil} vs. T_{air} relationship and its modulation due to snow pack. ORCHIDEE
 10 employs, similarly to UVic, a fixed snow density and thermal conductivity. However, in
 11 contrast with UVic, ORCHIDEE applies a multi-layer scheme and simulates heat diffusion in
 12 the snowpack in up to 7 discrete layers (Table 1; Koven et al., 2009). This helps resolving the
 13 snow thermal gradients between the top and the base of the snow cover, and might explain
 14 how some of the snow insulation effects are reasonably represented in ORCHIDEE, despite
 15 the simpler treatment of temperature diffusion.

16 5 Permafrost area

17 Snow cover plays an important role in modulating the variations of soil thermodynamics, and
 18 hence near-surface permafrost extent (e.g., Park et al., 2015). Here we evaluate if there is a
 19 simple relationship between the simulated Northern hemisphere permafrost area and the
 20 sophistication and ability of the snow insulation component in the LSM to match observed
 21 snow packs. The simulated near-surface permafrost area varies greatly across the nine models
 22 in the ~~hindcast hind cast~~ simulation (1960-2009; Table 4). Some of the better performing snow
 23 insulation effect models (CLM4.5, JULES) simulate a near-surface permafrost area of 13.19
 24 to 15.77 million km², which is comparable with the IPA map estimate (16.2 million km²)
 25 (Brown et al., 1997; Slater and Lawrence, 2013). CoLM and ORCHIDEE, identified as
 26 reasonable models with respect to snow insulation, simulate much lower (7.62 million km²)
 27 and higher (20.01 million km²) areas, respectively. The main deficiency of CoLM is its too
 28 ~~small-shallow~~ soil depth (3.4 m) compared with CLM4.5 (45.1 m) despite having very similar
 29 snow modules (Table 1). However, ISBA, one of the two models that showed rather limited
 30 skill in representing snow insulation effects, also ~~significantly over-estimate simulates the~~
 31 ~~highest~~ permafrost area (20.86 million km²). This is inconsistent with previous studies (e.g.,
 32 Vavrus, 2007; Koven et al., 2013) which concluded that the first-order control on modelled
 33 near-surface permafrost distribution is the representation of the air-to-surface soil temperature

1 difference. Table 4 shows that the situation is more complex and that snow insulation
 2 simulation is not the dominant factor in a good permafrost extent simulation. When the land
 3 surface models having poor snow models are eliminated, the remaining models' simulated
 4 permafrost area show little or no relationship with the performance of the snow insulation
 5 component, because several other factors such as differences in the treatment of soil organic
 6 matter, soil hydrology, surface energy calculations, model soil depth, and vegetation also
 7 provide important controls on the simulated permafrost distribution (e.g., Marchenko and
 8 Etzelmüller, 2013).

9 **6 Summary and conclusions**

10 The aim of this work was to evaluate how state-of-the-art LSMs capture the observed
 11 relationship between winter near-surface soil and air temperatures (T_{soil} , T_{air}) and their
 12 modulation by snow depth (d_{snow}) and climate regime. We presented some benchmarks to
 13 evaluate model performance. The presented relation diagrams of T_{soil} and the difference of
 14 $T_{soil}-T_{air}$ to snow depth allow a much better assessment to reveal structural issues of the
 15 models than a direct point-by-point comparison with station observations. The results are
 16 based on the comparison of LSMs with a comprehensive Russian station data set.

17

18 We see large differences across the models in their mean air-soil temperature difference (T)
 19 of 3 to 14 K°C , in the sensitivity of near-surface soil temperature to ~~and~~ air temperature (T_{soil}
 20 vs. T_{air}) (0.49 to 0.96°C/°C for shallow snow, 0.13 to 0.93°C/°C for thick snow), and in the
 21 increase of T with increasing snow depth (modal value of T PDF: 0 to 10 K°C for shallow
 22 snow, 5 to 21 K°C for thick snow). Most of the nine models compare to the observations
 23 reasonably well (observations: $T = 12 \text{ K}^{\circ}\text{C}$, modal T values of 5 K°C for shallow snow and
 24 of 14 K°C for thick snow, T_{soil} vs. $T_{air} = 0.62^{\circ}\text{C}/^{\circ}\text{C}$ for shallow snow, T_{soil} vs. $T_{air} = 0.21^{\circ}\text{C}/^{\circ}\text{C}$
 25 for thick snow). Several models also capture the modulation by air temperature condition
 26 (larger increase in T with increasing d_{snow} under colder conditions) and display the control of
 27 snow depth on T_{soil} (weaker T_{soil} vs. T_{air} relationship under thicker snow).- However, while
 28 they generally capture these observed relationships, their strength can differ in the individual
 29 models. Two models (ISBA, UVic) show the largest deficits in snow insulation effects and
 30 cannot separate the T regimes neither for different snow depths nor for different air
 31 temperature conditions.

32

1 This study uses the ensemble of models to document model performance with respect to T_{soil}
2 versus T_{air} relationships, and to identify those with better performance, rather than to quantify
3 the best model. We were able to attribute performance strength/weakness to snow model
4 features and complexity. Models with better performance apply multi-layer snow schemes
5 and consider complex snow processes (e.g. storage and refreezing of liquid water within the
6 snow, wet snow metamorphism, snow compaction). Those models which show limited skill in
7 snow insulation representation (underestimated T , very weak dependency of T on d_{snow} ,
8 almost unity ratio of T_{soil} vs. T_{air}) have some deficiencies or over simplification in the
9 simulation of heat transfer in snow and soil layer, particularly in the representation of snow
10 depth and density (conductivity). We also emphasize that compensating errors in snow depth
11 and conductivity can occur. For example, an excessive correlation between T_{soil} and T_{air} can
12 be attributed to excessively high thermal conductivity even when the snow depth is correctly
13 (or over) simulated. This finding underscores the need for detailed model evaluations using
14 multiple, independent performance metrics to establish that the models get the right
15 functionality for the right reason. It should be noted that the treatment of ground properties,
16 particularly soil organic matter and soil moisture/ice content, also affect the simulated winter
17 ground temperatures. The specific evaluation of these individual processes is more robustly
18 investigated with experiments conducted for individual models (e.g. recently, Wang et al.,
19 2013; Gubler et al., 2013; Decharme et al., 2015).

20

21 Snow and its insulation effects are critical for accurately simulating soil temperature and
22 permafrost in high latitudes. The simulated near-surface permafrost area varies greatly across
23 the nine models (from 7.62 to 20.86 million km²). However, it is hard to find a clear
24 relationship between the performance of the snow insulation in the models and the simulated
25 area of permafrost, because several other factors e.g. related to soil depth and properties and
26 vegetation cover also provide important controls on the simulated permafrost distribution.

27

28 *Author contributions:* A.R. developed the idea that lead to this paper; W.W. and A.R.
29 performed research; J.C.M. contributed with ideas; A.R., J.C.M., W.W., and D.J. wrote the
30 paper; and all authors discussed the results and commented on the paper at all stages.

31

32 *Acknowledgments.* The data will be made available through the National Snow and Ice Data
33 Center (NSIDC; <http://nsidc.org>); the contact person is Kevin Schaefer
34 (kevin.schaefer@nsidc.org). This study was supported by the Permafrost Carbon

1 Vulnerability Research Coordination Network, which is funded by the U.S. National Science
2 Foundation (NSF). Any use of trade, firm, or product names is for descriptive purposes only
3 and does not imply endorsement by the U.S. Government. E.J.B. was supported by the Joint
4 UK DECC / Defra Met Office Hadley Centre Climate Program (GA01101). E.J.B., S.P., P.C.
5 and G.K. were supported by the European Union Seventh Framework Program (FP7/2007-
6 2013) under grant agreement n°282700. T.J.B. was supported by grant 1216037 from the NSF
7 Science, Engineering and Education for Sustainability (SEES) Post-Doctoral Fellowship
8 program. B.D., R.A. and C.D. were supported by the French Agence Nationale de la
9 Recherche under agreement ANR-10-CEPL-012-03. This research was sponsored by the
10 Integrated approaches and impacts, China Global Change Program (973 Project), National
11 Basic Research Program of China Grant 2015CB953602 and the National Natural Science
12 Foundation of China Grant 40905047. We also thank the editor and reviewers for their
13 comments to improve the manuscript.

14 **References**

- 15 Anderson, E.A.: A point energy and mass balance model of a snow cover, Office of
16 Hydrology, National Weather Service, Silver Spring, Maryland, NOAA Technical Report
17 NWS 19, 1976.
- 18 Andreadis, K., Storck, P. and Lettenmaier, D.P.: Modeling snow accumulation and ablation
19 processes in forested environments, *Water Resour. Res.*, 45, W05429,
20 doi:10.1029/2008WR007042, 2009.
- 21 Anisimov, O.A., and Sherstiukov A.B.: Evaluating the effect of environmental factors on
22 permafrost factors in Russia, *Earth's Cryosphere*, XX(2), 90-99, 2016.
- 23 Avis, C.A.: Simulating the present-day and future distribution of permafrost in the UVic Earth
24 System Climate Model, Dissertation, University of Victoria, Canada, 274 pp., 2012.
- 25 Bartlett, P.A., MacKay, M.D., and Verseghy, D.L.: Modified snow algorithms in the
26 Canadian Land Surface Scheme: model runs and sensitivity analysis at three boreal forest
27 stands, *Atmosphere-Ocean*, 44, 207–222, 2006.
- 28 Best, M.J., and 16 co-authors: The Joint UK Land Environment Simulator (JULES), model
29 description-Part 1: energy and water fluxes, *Geosci. Model. Dev.*, 4, 677–699,
30 doi:10.5194/gmd-4-677-2011, 2011.
- 31 Boone, A., and Etchevers, P.: An intercomparison of three snow schemes of varying
32 complexity coupled to the same land-surface model: Local scale evaluation at an Alpine
33 site, *J. Hydrometeor.*, 2, 374-394, 2001.

- 1 Brown, J., Ferrians, O.J., Heginbottom, J.A. and Melnikov, S.E.: International Permafrost
2 Association Circum-Arctic Map of Permafrost and Ground Ice Conditions, scale
3 1:10,000,000, Circum-Pacific Map Series, USGS Circum-Pacific Map Series, ~~Map CP-45~~,
4 1997.
- 5 Brun, E., David, P., Sudul, M. and Brunot, G.: A numerical model to simulate snow cover
6 stratigraphy for operational avalanche forecasting, *J. Glaciol.*, 38, 13-22, 1992.
- 7 Brun, E., Vionnet, V., Boone, A., Decharme, B., Peings, Y., Valette, R., Karbou, F. and
8 Morin, S.: Simulation of northern Eurasian local snow depth, mass and density using a
9 detailed snowpack model and meteorological reanalysis, *J. Hydrometeorol.*, 14, 203-214,
10 doi:10.1175/jhm-d-12-012.1, 2013.
- 11 Cook, B.I., Bonan, G.B., Levis, S. and Epstein, H.E.: The thermo-insulation effect of snow
12 cover within a climate model, *Clim. Dyn.*, 31, 107-124, doi:10.1007/s00382-007-0341-y,
13 2008.
- 14 Dai, Y., Zeng, X., Dickinson, R.E., Baker, I., Bonan, G.B., Bosilovich, M.G., Denning, A.S.,
15 Dirmeyer, P.A., Houser, P.R., Niu, G., Oleson, K.W., Schlosser, C.A. and Yang, Z.: The
16 Common Land Model (CLM), *Bull. Am. Meteorol. Soc.*, 84, 1013-1023,
17 doi:10.1175/BAMS-84-8-1013, 2003.
- 18 Dankers, R., Burke, E.J., and Price, J.: Simulation of permafrost and seasonal thaw depth in
19 the JULES land surface scheme, *The Cryosphere*, 5, 773-790, doi:10.5194/tc-5-773-2011,
20 2011.
- 21 Decharme, B., Brun, E., Boone, A., Delire, C., Le Moigne, P. and Morin, S.: Impacts of snow
22 and organic soils parameterization on North-Eurasian soil temperature profiles simulated
23 by the ISBA land surface model, *The Cryosphere*, 10, 853-877, doi: 10.5194/tc-10-853-
24 2016, 2016.
- 25 Douville, H., Royer, J.F. and Mahfouf, J.F.: A new snow parameterization for the Meteo-
26 France climate model. Part 1: Validation in stand-alone experiments, *Clim. Dyn.*, 12, 21-
27 35, 1995.
- 28 Dutra, E., Viterbo, P., Miranda, P.M.A. and Balsamo, G.: Complexity of snow schemes in a
29 climate model and its impact on surface energy and hydrology, *J. Hydrometeorol.*, 13,
30 521-538, doi:10.1175/jhm-d-11-072.1, 2012.
- 31 Dutra, E., Balsamo, G., Viterbo, P., Miranda, P.M.A., Beljaars, A., Schär, C. and Elder, K.:
32 An improved snow scheme for the ECMWF land surface model: description and offline
33 validation, *J. Hydrometeorol.*, 11, 899-916, 2010.

- 1 Essery, R., Morin, S., Lejeune, Y. and Ménard, C.B.: A comparison of 1701 snow models
2 using observations from an alpine site, *Adv. Water Resour.*, 55, 131–148,
3 doi:10.1016/j.advwatres.2012.07.013, 2013.
- 4 Essery, R.L.H, Rutter, N., Pomeroy, J., Baxter, R., Staehli, M., Gustafsson, D., Barr, A.,
5 Bartlett, P. and Elder, K.: SnowMIP2: An evaluation of forest snow process simulations,
6 *Bull. Am. Meteorol. Soc.*, 90, 1120-1135, doi:10.1175/2009BAMS2629.1, 2009.
- 7 Ge, Y. and Gong, G.: Land surface insulation response to snow depth variability, *J. Geophys.*
8 *Res.*, 115, D08107, doi:10.1029/2009JD012798, 2010.
- 9 Gerten, D., Schaphoff, S., Haberlandt, U., Lucht, W. and Sitch, S.: Terrestrial vegetation and
10 water balance: Hydrological evaluation of a dynamic global vegetation model, *J. Hydrol.*,
11 286, 249–270, 2004.
- 12 Gisnas, K., Westermann, S., Schuler, T., Litherland, T., Isaksen, K., Boike, J. and Etzelmüller,
13 B.: A statistical approach to represent small-scale variability of permafrost temperatures
14 due to snow cover, *The Cryosphere*, -8, 2063-2074, doi: 10.5194/tc-8-2063-2014, 2014.
- 15 Gouttevin, I., Menegoz, M., Domine, F., Krinner, G., Koven, C.D., Ciais, P., Tarnocai, C. and
16 Boike, J.: How the insulating properties of snow affect soil carbon distribution in the
17 continental pan-Arctic area, *J. Geophys. Res.*, 117, G02020, doi:10.1029/2011JG001916,
18 2012.
- 19 Gubler, S., Endrizzi, S., Gruber, S. and Purves, R.S.: Sensitivities and uncertainties of
20 modeled ground temperatures in mountain environments, *Geosci. Model Dev.*, 6, 1319-
21 1336, doi:10.5194/gmd-6-1319-2013, 2013.
- 22 Jafarov, E.E., Nicolsky, D.J., Romanovsky, V.E., Walsh, J.E., Panda, S.K. and Serreze, M.C.:
23 The effect of snow: How to better model ground surface temperatures, *Cold Regions Sci.*
24 *Technol.*, 102, 63-77, doi:10.1016/j.coldregions.2014.02.007, 2014.
- 25 Ji, D., Wang, L., Feng, J., Wu, Q., Cheng, H., et al.: Description and basic evaluation of
26 Beijing Normal University Earth System Model (BNU-ESM) version 1, *Geosci. Model*
27 *Dev.*, 7, 2039-2064, 2014.
- 28 Jordan, R.: A one-dimensional temperature model for a snow cover, technical documentation
29 for SNTHERM.89, ~~Special Report 91-16~~, U.S. Army Cold Regions Research and
30 Engineering Laboratory, Hanover, N.H, 49 pp., 1991.
- 31 Karunaratne, K.C., and Burn, C.R.: Freezing n-factors in discontinuous permafrost terrain,
32 Takhini River, Yukon Territory, Canada, *Proc. 8th Int. Conf. on Permafrost, Zurich*, ~~Eds. M.~~
33 ~~Phillips, S.M. Springman and L.U. Arenson~~, 519–524, 2003.

- 1 Klehmet, K., Geyer, B., and Rockel, B.: A regional climate model hindcast for Siberia:
 2 analysis of snow water equivalent, *The Cryosphere*, 7, 1017-1034, doi:10.5194/tc-7-1017-
 3 2013, 2013.
- 4 Koven, C.D., Riley, W.J. and Stern, A.: Analysis of Permafrost Thermal Dynamics and
 5 Response to Climate Change in the CMIP5 Earth System Models, *J. Clim.*, 26, 1877--
 6 1900. doi:10.1175/JCLI-D-12-00228.1, 2013.
- 7 Koven, C., Friedlingstein, P., Ciais, P., Khvorostyanov, D., Krinner, G. and Tarnocai, C.: On
 8 the formation of high-latitude soil carbon stocks: Effects of cryoturbation and insulation by
 9 organic matter in a land surface model, *Geophys. Res. Lett.*, 36, L21501,
 10 doi:10.1029/2009GL040150, 2009.
- 11 Langer, M., Westermann, S., Heikenfeld, M., Dorn, W., Boike, J.: Satellite-based modeling of
 12 permafrost temperatures in a tundra lowland landscape, *Remote Sensing of Environment*,
 13 135, pp-12-24, doi:10.1016/j.rse.2013.03.011, 2013.
- 14 Lawrence, D.M., and Slater, A.G.: The contribution of snow condition trends to future ground
 15 climate, *Clim. Dyn.*, 34, 969-981, doi:10.1007/s00382-009-0537-4, 2010.
- 16 Ling, F., and Zhang, T.: Sensitivity study of tundra snow density on surface energy fluxes and
 17 ground thermal regime in Northernmost Alaska, *Cold Regions Sci. Technol.*, 44, 121--130,
 18 2006.
- 19 Marchenko, S. and Etzelmüller, B.: Permafrost: Formation and Distribution, Thermal and
 20 Mechanical Properties--~~In: John F. Shroder (ed.) *Treatise on Geomorphology*, Volume 8,~~
 21 ~~pp. 202-222. San Diego: Academic Press, San Diego, 202-222,~~ 2013.
- 22 [McGuire, A.D., C. Koven, D.M. Lawrence, J.S. Clein, J. Xia, C. Beer, E. Burke, G. Chen,](#)
 23 [X.Chen, C. Delire, E. Jafarov, A.H. MacDougall, S. Marchenko, D. Nicolsky, S. Peng, A.](#)
 24 [Rinke, K. Saito, W. Zhang, R. Alkama, T.J. Bohn, P. Ciais, B. Decharme, A. Ekici, I.](#)
 25 [Gouttevin, T. Hajima, D.J. Hayes, D. Ji, G. Krinner, D.P. Lettenmaier, Y. Luo, P.A. Miller,](#)
 26 [J.C. Moore, V. Romanovsky, C. Schädel, K. Schaefer, E.A.G. Schuur, B. Smith, T.](#)
 27 [Sueyoshi, and Q. Zhuang: A model-based analysis of the vulnerability of carbon in the](#)
 28 [permafrost region between 1960 and 2009, *Global Biogeochemical Cycles*,](#)
 29 [doi:10.1002/2016GB005405, 2016](#)
- 30 Meissner, K.J., Weaver, A.J., Matthews, H.D. and Cox, P.M.: The role of land-surface
 31 dynamics in glacial inception: A study with the UVic earth system model, *Clim. Dyn.*, 21,
 32 515-537, 2003.

- 1 Morse, P.D., Burn, C.R., and Kokelj, S.V.: Influence of snow on near-surface ground
2 temperatures in upland and alluvial environments of the outer Mackenzie Delta, Northwest
3 Territories, *Can. J. Earth Sci.*, 49, 895–913, doi:10.1139/E2012-012, 2012.
- 4 Muskett, R.: Remote sensing, model-derived and ground measurements of snow water
5 equivalent and snow density in Alaska, *Int. J. Geosci.*, 3, 1127-1136,
6 doi:10.4236/ijg.2012.35114, 2012.
- 7 Nicolsky, D.J., Romanovsky, V.E., Alexeev, V.A. and Lawrence, D.M.: Improved modelling
8 of permafrost dynamics in a GCM land-surface scheme, *Geophys. Res. Lett.*, 34, L08501,
9 doi:10.1029/2007GL029525, 2007.
- 10 Oleson, K.W., Lawrence, D.M., Bonan, G.B., Drewniak, B., Huang, M., Koven, C.D., Levis,
11 S., Li, F., Riley, W.J., Subin, Z.M., Swenson, S.C., Thornton, P.E., Bozbiyik, A., Fisher,
12 R., Kluzek, E., Lamarque, J.-F., Lawrence, P.J., Leung, L.R., Lipscomb, W., Muszala, S.,
13 Ricciuto, D.M., Sacks, W., Sun, Y., Tang, J., Yang, Z.-L.: Technical description of version
14 4.5 of the Community Land Model (CLM). NCAR Technical Note NCAR/TN-503+STR,
15 doi: 10.5065/D6RR1W7M, 2013.
- 16 PaiMazumder, D., Miller, J., Li, Z., Walsh, J. E., Etringer, A., McCreight, J., Zhang, T.,
17 Mölders, N. Evaluation of Community Climate System Model soil temperatures using
18 observations from Russia, *Theoretical and Applied Climatology* , 94(3),187-213, 2008.
- 19 Paquin, J.P., and Sushama, L.: On the Arctic near-surface permafrost and climate sensitivities
20 to soil and snow model formulations in climate models, *Clim. Dyn.*, 44, 203-228,
21 doi:10.1007/s00382-014-2185-6, 2015.
- 22 Park, H., Fedorov, A.N., Zheleznyak, M.N., Konstantinov, P.Y. and Walsh, J.E.: Effect of
23 snow cover on pan-Arctic permafrost thermal regimes, *Clim. Dyn.*, 44, 2873-2895,
24 doi:10.1007/s00382-014-2356-5, 2015.
- 25 Park, H., A.B. Sherstiukov, A.N. Fedorov, I.V. Polyakov, and J.E Walsh: An observation-
26 based assessment of the influences of air temperature and snow depth on soil temperature
27 in Russia, *Environ. Res. Lett.*, 9, 064026, doi:10.1088/1748-9326/9/6/064026, 2014.
- 28 Pavlov, A.V., Malkova, G.V.: Small-scale mapping of trends of the contemporary ground
29 temperature changes in the Russian North, *Earth's Cryosphere*, XIII(4), 32-39, 2009.
- 30 Peng, S., Ciais, P., Krinner, G., Wang, T., Gouttevin, I., McGuire, A., Lawrence, D., Burke,
31 E., Chen, X., Decharme, B., Koven, C., MacDougall, A., Rinke, A., Saito, K., Zhang, W.,
32 Alkama, R., Bohn, T.J., Delire, C., Hajima, T., Ji, D., Lettenmaier, D.P., Miller, P.A.,
33 Moore, J.C., Smith, B. and Sueyoshi, T.: Simulated high-latitude soil thermal dynamics
34 during the past four decades, *The Cryosphere*, 10, 1–14, doi:10.5194/tc-10-1-2016, 2015.

- 1 Rawlins, M., McGuire, A., Kimball, J., Dass, P., Lawrence, D., Burke, E., Chen, X., Delire,
2 C., Koven, C., MacDougall, A., Peng, S., Rinke, A., Saito, K., Zhang, W., Alkama, R.,
3 Bohn, T.J., Ciais, P., Decharme, B., Gouttevin, I., Hajima, T., Ji, D., Krinner, G.,
4 Lettenmaier, D.P., Miller, P.A., Moore, J.C., Smith, B. and Sueyoshi, T.: Assessment of
5 model estimates of land-atmosphere CO₂ exchange across Northern Eurasia,
6 *Biogeosciences*, 12, 4385-4405, doi:10.5194/bg-12-4385-2015, 2015.
- 7 Riseborough, D.W.: An analytical model of the ground surface temperature under snow cover
8 with soil freezing, 58th Eastern Snow Conference, Ottawa, Ontario, Canada, 2001
- 9 Romanovsky, V.E., Sazonova, T.S., Balobaev, V.T., Shender, N.I. and Sergueev, D.O.: Past
10 and recent changes in air and permafrost temperatures in eastern Siberia, *Glob. Planet.*
11 *Change*, 56, 399-413, doi:10.1016/j.gloplacha.2006.07.022, 2007.
- 12 Saha, S., Rinke, A., Dethloff, K. and Kuhry, P.: Influence of complex land surface scheme on
13 Arctic climate simulations, *J. Geophys. Res.*, 111, D22104, doi:10.1029/2006JD007188,
14 2006.
- 15 Schaefer, K., Zhang, T., Bruhwiler, L. and Barrett, A.P.: Amount and timing of permafrost
16 carbon release in response to climate warming, *Tellus B*, 63, 165–180,
17 doi:10.1111/j.1600-0889.2011.00527.x, 2011.
- 18 Schuur, E.A.G., and 18 coauthors: Vulnerability of permafrost carbon to climate change:
19 Implications for the global carbon cycle, *Bioscience*, 58, 701-714. doi:10.1641/b580807,
20 2008.
- 21 Sherstiukov, A. Dataset of daily soil temperature up to 320 cm depth based on meteorological
22 stations of Russian Federation, RIHMI-WDC, 176, 224-232, 2012a.
- 23 Sherstiukov, A. Statistical quality control of soil temperature dataset, RIHMI-WDC, 176,
24 224-232, 2012b.
- 25 Singh, V.P., Singh, P. and Haritashya, U.: *Encyclopedia of snow, ice and glaciers*, Springer,
26 1240 pp., 2011.
- 27 Slater, A.G., Schlosser, C.A. and Desborough, C.E.: The representation of snow in land-
28 surface schemes: Results from PILPS 2(d), *J. Hydrometeorol.*, 2, 7-25, 2001.
- 29 Slater, A., and Lawrence, D.: Diagnosing present and future permafrost from climate models,
30 *J. Clim.*, [26, 5608-5623](#), doi:10.1175/JCLI-D-12-00341.1, 2013.
- 31 Smith, M.W., and Riseborough, D.W.: Climate and the limits of permafrost: a zonal analysis,
32 *Permafrost Periglac. Process.*, 13, 1-15, doi: 10.1002/ppp.410, 2002.

- 1 Sokratov, S.A., and Barry, R.G.: Intraseasonal variation in the thermoinsulation effect of
2 snow cover on soil temperatures and energy balance, *J. Geophys. Res.*, 107, [D10](#),
3 doi:10.1029/2001JD000489, 2002.
- 4 Sturm, M., Holmgren, J., König, M. and Morris, K.: The thermal conductivity of seasonal
5 snow, *J. Glaciol.*, 43-~~(143)~~, 26–41, 1997.
- 6 Sturm, M., and Stuefer, S.: Wind-blown flux rates derived from drifts at arctic snow fences, *J.*
7 *Glaciol.*, 59-~~(213)~~, 21-34, 2013.
- 8 Swenson, S.C., and Lawrence, D.M.: A new fractional snow-covered area parameterization
9 for the Community Land Model and its effect on the surface energy balance, *J. Geophys.*
10 *Res.*, 117, D21107, doi:10.1029/2012JD018178, 2012.
- 11 Takala, M., Luojus, K., Pulliainen, J., Derksen, C., Lemmetyinen, J., Karna, J.P., Koskinen, J.
12 and Bojkov, B.: Estimating Northern Hemisphere snow water equivalent for climate
13 research through assimilation of space-borne radiometer data and ground-based
14 measurements, *Remote Sens. Environ.*, 115, 3517-3529, 2011.
- 15 Takata, K., Emori, S. and Watanabe, T.: Development of the minimal advanced treatments of
16 surface interaction and runoff, *Glob. Planet. Change*, 38, 209–222, 2003.
- 17 Vavrus, S.J.: The role of terrestrial snow cover in the climate system, *Clim. Dyn.*, 20, 73-88,
18 doi:10.1007/s00382-007-0226-0, 2007.
- 19 Vionnet, V., Brun, E., Morin, S., Boone, A., Faroux, S., Moigne, P.L., Martin, E. and
20 Willemet, J.-M.: The detailed snowpack scheme Crocus and its implementation in
21 SURFEX v7.2, *Geosci. Model Dev.*, 5, 773–791, 2012.
- 22 Vionnet, V., Guyomarch, G., Martin, E., Durand, Y., Bellot, H., Bel, C. and Pugliese, P.:
23 Occurrence of blowing snow events at an alpine site over a 10-year period: observations
24 and modelling, *Adv. Water Resour.*, 55, 53-63, 2013.
- 25 Von Storch, H. and Zwiers, F.W.: *Statistical Analysis in Climate Research*, Cambridge
26 University Press, Cambridge, 484pp., 1999.
- 27 Wang, T., Otle, C., Boone, A., Ciais, P., Brun, E., Morin, S., Krinner, G., Piao, S. and Peng,
28 S.: Evaluation of an improved intermediate complexity snow scheme in the ORCHIDEE
29 land surface model, *J. Geophys. Res.*, 118, 6064-6079, doi:10.1002/jgrd.50395, 2013.
- 30 Wania, R., Ross, I. and Prentice, I.C.: Integrating peatlands and permafrost into a dynamic
31 global vegetation model: 2. Evaluation and sensitivity of vegetation and carbon cycle
32 processes, *Global Biogeochem. Cycles*, 23, GB3015, doi:10.1029/2008GB003413, 2009.

- 1 Woo, M., Heron, R., Marsh, P. and Steer, P.: Comparison of weather station snowfall with
2 winter snow accumulation in high arctic basins, *Atmosph-Ocean*, 21-(3), 312-325,
3 doi:10.1080/07055900.1983.9649171, 1983.
- 4 Zhang, T.: Influence of the seasonal snow cover on the ground thermal regime: An overview,
5 *Rev. Geophys.*, 43, RG4002, doi:10.1029/2004RG000157, 2005.
- 6 Zhang, T., Osterkamp, T.E. and Stamnes, K.: Influence of the depth hoar layer of the seasonal
7 snow cover on the ground thermal regime, *Water Resour. Res.*, 32, 2075–2086,
8 doi:10.1029/96WR00996, 1996.
- 9 Zhong, X., Zhang, T. and Wang, K.: Snow density climatology across the former USSR, *The*
10 *Cryosphere*, 8, 785-799, doi:10.5194/tc-8-785-2014, 2013.
- 11
12
13
14
15

1 **Tables**

2

3 **Table 1.** PCN snow model details.

Model Reference for snow scheme	Snow scheme ¹	Snow layers	Water phases	Liquid water treatment ²	Snow density ³	Snow thermal conductivity ⁴
CLM4.5 Swenson and Lawrence, 2012 Oleson et al., 2013	ML	Dynamic (max. 5)	Liquid, Ice	Bucket-type prognostic in each layer	depends on snow depth; compaction ^{3) a,b,c}	quadratic equation on
CoLM Dai et al., 2003 Ji et al. 2014	ML	Dynamic (max. 5)	Liquid, Ice	Bucket-type prognostic in each layer	depends on snow depth; compaction ^{3) a,b,c}	quadratic equation on
ISBA Boone and Etchevers, 2001	ML	Static (3)	Liquid, Ice, Vapor	Diagnosed from snow temperature, mass, density	compaction ^{3) a,b}	quadratic equation on , contribution due to vapor transfer
JULES Best et al., 2011	ML	Dynamic (max. 3)	Liquid, Ice, Vapor	Bucket-type prognostic in each layer	compaction ^{3) a}	power equation on
LPJ-GUESS Gerten et al., 2004 Wania et al., 2009	BL	Static (1)	Ice	Not represented	fixed 362 kg m ⁻³	fixed 0.196 Wm ⁻¹ K ⁻¹
MIROC-ESM Takata et al., 2003	ML	Dynamic (max. 3)	Ice	Not represented	fixed 300 kg m ⁻³	fixed 0.3 Wm ⁻¹ K ⁻¹
ORCHIDEE Gouttevin et al.,2012	ML	Dynamic (max. 7)	Ice	Not represented	fixed 330 kg m ⁻³	fixed 0.25 Wm ⁻¹ K ⁻¹ for tundra, 0.042 Wm ⁻¹ K ⁻¹ for taiga
UVic Meissner et al., 2003 Avis, 2012	I	Static (1)	Ice	Not represented	fixed 330 kg m ⁻³	bulk conductivity
UW-VIC Andreadis et al., 2009	BL	Dynamic (max. 2)	Liquid, Ice, Vapor	Constant liquid water holding capacity	compaction ^{3) a,b}	fixed 0.7 Wm ⁻¹ K ⁻¹

4 ¹ ML: Multi-layer, BL: Bulk-layer, I: Implicit; according to Slater et al. (2001)5 ² Not represented means dry snow6 ³ Processes for densification of the snow: a) mechanical compaction (due to the weight of the overburden), b) thermal metamorphosis (via the melting–refreezing process), c) destructive metamorphosis (crystal breakdown due to wind, thermodynamic stress); Anderson (1976), Jordan (1991), Kojima (1967)7 ⁴ quadratic equation on according to Jordan (1991), Anderson (1976); contribution due to vapor transfer according to Sun et al.(1999)

8

9

10

1 **Table 2.** Sensitivity of near-surface soil temperature (T_{soil}) to air temperature (T_{air}) in winter
 2 (DJF) calculated by the slopes of the linear regression between T_{soil} ($^{\circ}\text{C}$) and T_{air} ($^{\circ}\text{C}$) for
 3 different regimes of snow depth (d_{snow}), using data from all Russian station grid points and 21
 4 individual winter 1980-2000. All relationships are statistically significant at $p < 0.01$.

	Snow depth regimes			
	Shallow		Thick	
	d_{snow} 20 cm		d_{snow} 45 cm	
	T_{soil} vs. T_{air} ($^{\circ}\text{C}/^{\circ}\text{C}$)	R^2	T_{soil} vs. T_{air} ($^{\circ}\text{C}/^{\circ}\text{C}$)	R^2
Observation	0.62	0.79	0.21	0.41
CLM4.5	0.69	0.89	0.33	0.56
CoLM	0.49	0.73	0.13	0.44
ISBA	0.93	0.98	0.93	0.94
JULES	0.68	0.77	0.19	0.46
LPJ-GUESS	0.73	0.89	0.52	0.75
MIROC-ESM	0.78	0.98	0.49	0.67
ORCHIDEE	0.86	0.83	0.56	0.64
UVic	0.96	0.97	0.81	0.68
UW-VIC	0.54	0.74	0.76	0.65

6

7

1 **Table 3.** Russian-station-location averaged error statistics for snow depth (cm) and
 2 temperature difference between 20 cm soil and air temperature (T; K°C) for winter 1980-
 3 2000. For each variable, the maximum available number of observations (n) is used. Mean^{St,GS}
 4 and stdev^{St,GS} are the observed mean and interannual variability (standard deviation), while
 5 stdev is the standard deviations of each model. Bias is the mean error ‘simulation minus
 6 observation’ and rmse is the root-mean-square error. The statistics for snow depth is given
 7 based on both station observation (St) and GlobSnow (GS) data.

	Snow depth (n=579)					T (n=268)		
	mean St = 26.4 cm, mean ^{GS} =23.4 cm					mean St = 11.9 K°C		
	stdev St = 9.0 cm, stdev ^{GS} = 6.5 cm					stdev St = 2.3 K°C		
	bias St	rmse St	bias ^{GS}	rmse ^{GS}	stdev	bias St	rmse St	stdev
CLM4.5	11.5	18.1	14.3	18.1	5.8	2.3	4.1	2.2
CoLM	15.6	21.4	17.8	22.1	9.8	2.7	3.7	2.4
ISBA	13.0	18.8	15.7	19.8	9.5	-8.4	9.1	0.9
JULES	-4.1	14.1	-1.3	12.8	7.7	-0.8	4.2	3.2
LPJ-GUESS	-5.3	17.3	-2.5	16.0	5.0	-0.7	3.7	1.7
MIROC-ESM	-0.4	17.9	1.9	14.0	6.3	-4.9	6.7	2.0
ORCHIDEE	-8.7	16.5	-5.3	15.3	6.9	-5.2	6.0	1.9
UVic	-3.7	18.9	-0.5	16.8	9.4	-5.1	6.5	1.4
UW-VIC	12.5	19.8	15.0	20.0	10.4	-1.3	4.8	2.1

9

10

11

1 **Table 4.** Permafrost area, defined as maximum seasonal active layer thickness < 3 m in 1960
 2 (Mc Guire et al., 2016). The IPA map estimate is 16 million km² (Brown et al., 1997; Slater
 3 and Lawrence, 2013).

4

5

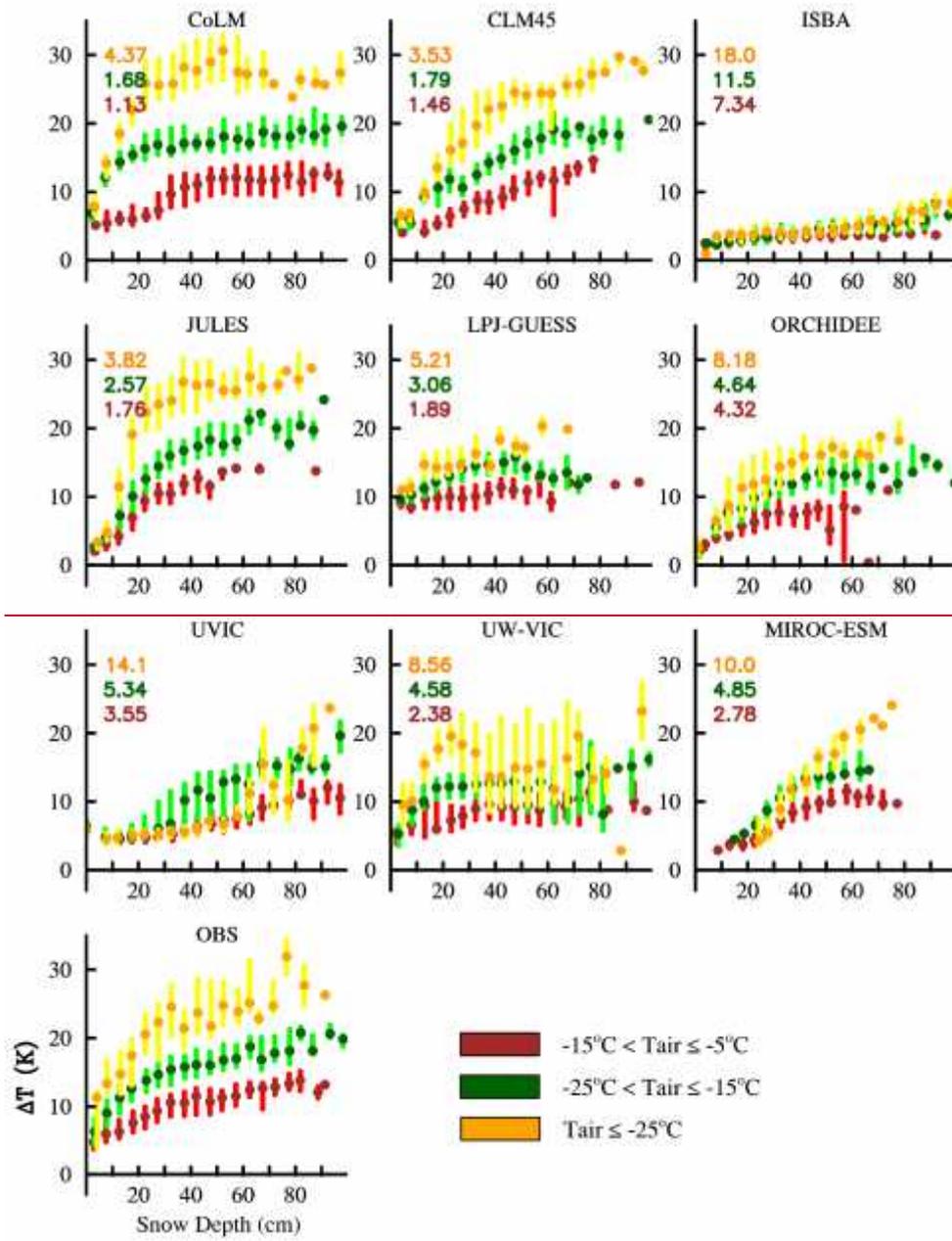
Land Surface Model	Snow Insulation skill	Permafrost Area (10 ⁶ km ²)
CLM4.5	High	15.77
CoLM	High	7.62
ISBA	Low	20.86
JULES	High	13.19
LPJ-GUESS	Medium	17.41
MIROC-ESM	Medium	13.02
ORCHIDEE	Medium	20.01
UVic	Low	16.47
UW-VIC	Medium	17.56

6

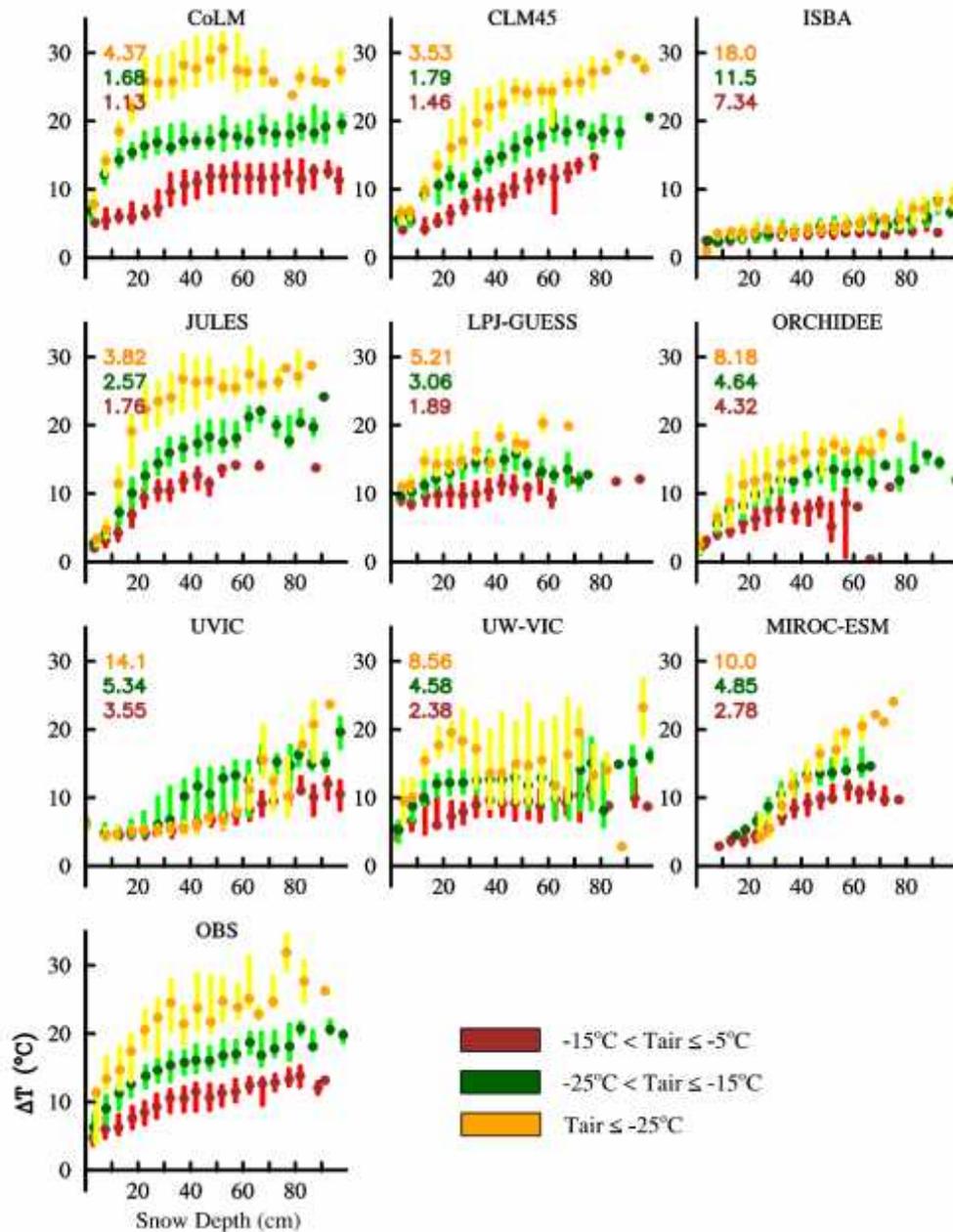
7

8

9



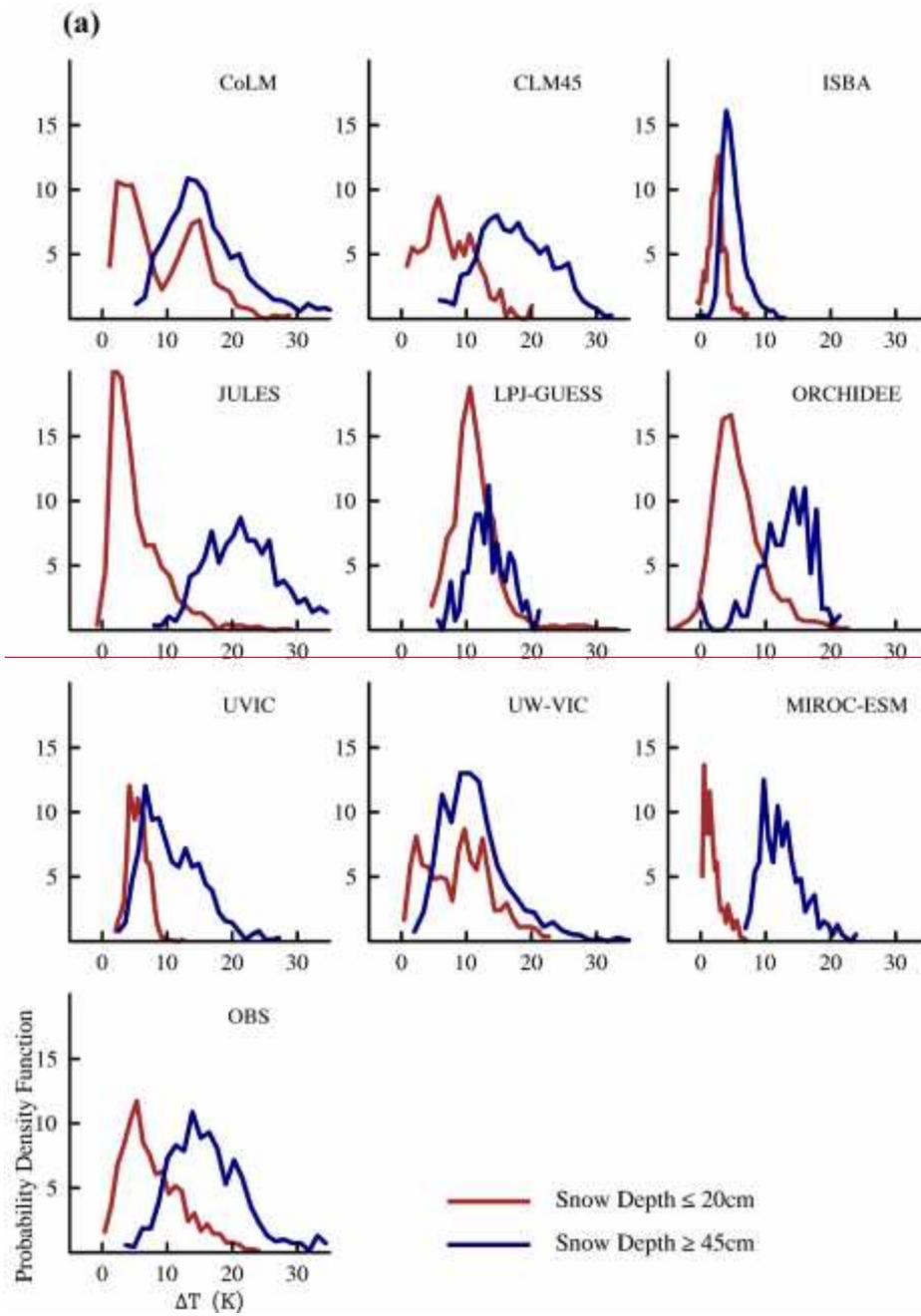
1



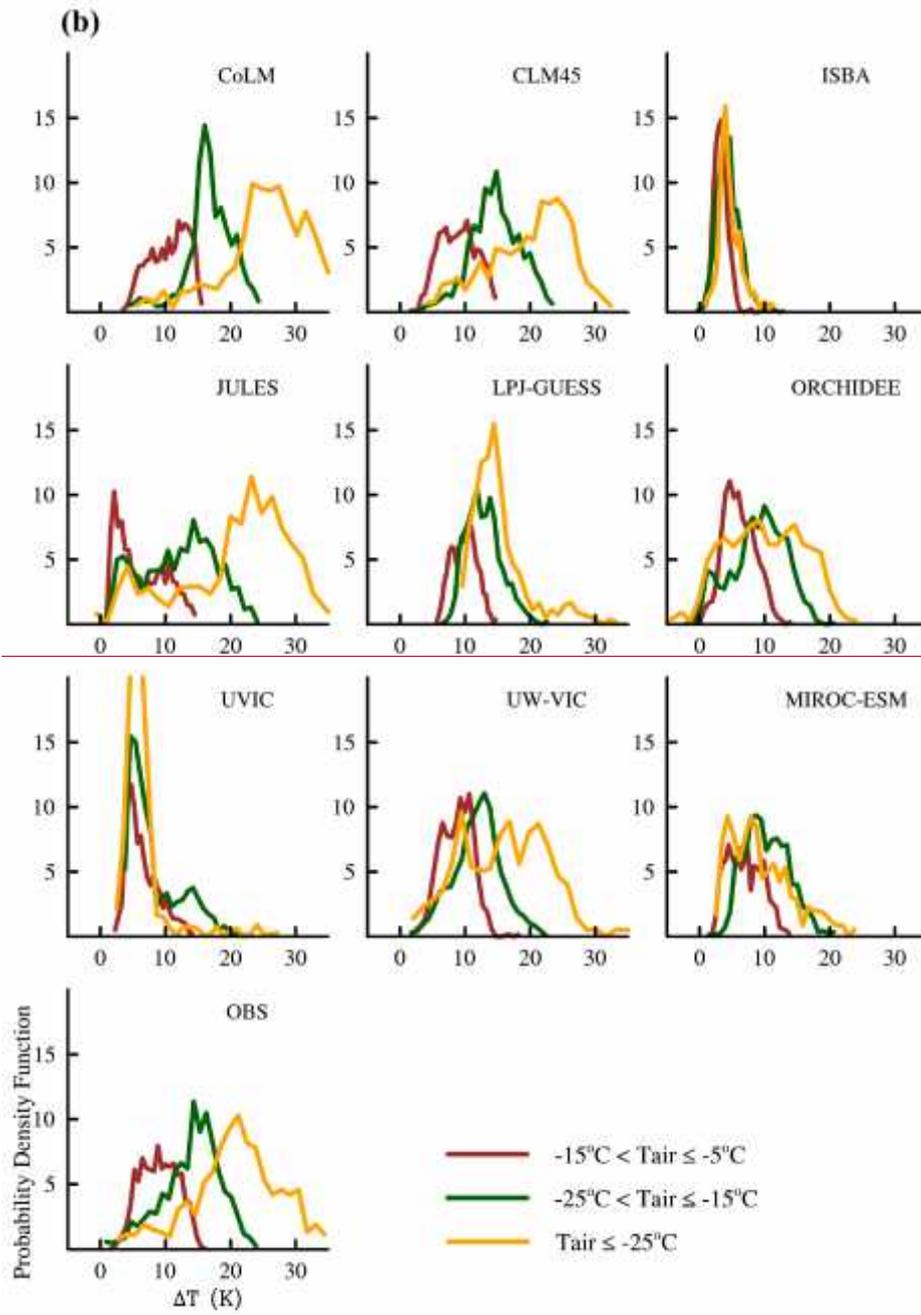
1

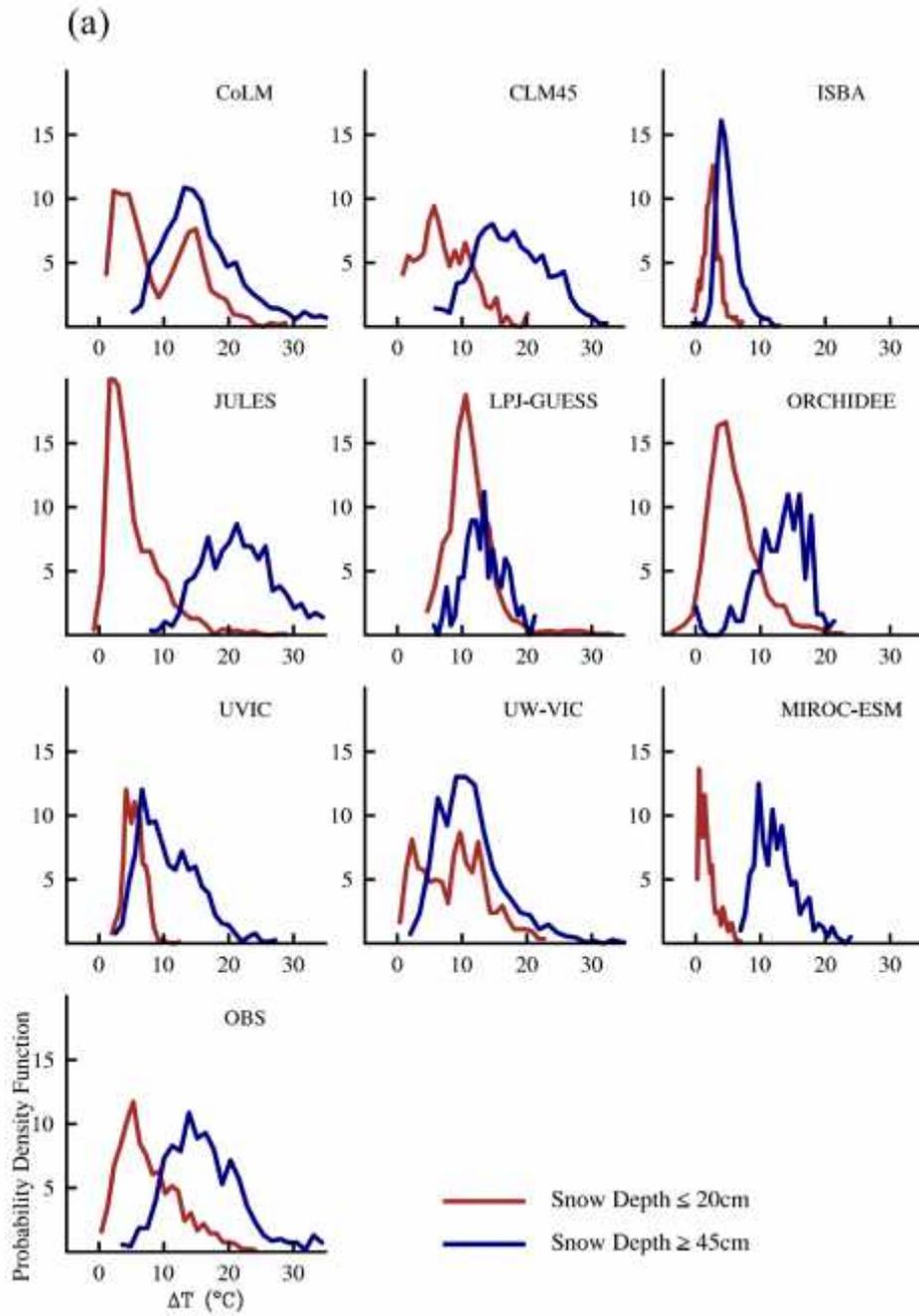
2 **Figure 1.** Variation of ΔT (K°C), the difference between soil temperature at 20 cm depth and
3 air temperature) with snow depth (cm) for winter 1980-2000. The dots represent the medians
4 of 5 cm snow depth bins and the upper and lower bars indicate the 25th and 75th percentiles,
5 calculated from all Russian station grid points ($n=268$) and 21 individual winters. The
6 numbers in each model panel indicate the RMSE between the observed and modeled
7 relationship. Color represents different air temperature regimes.

8

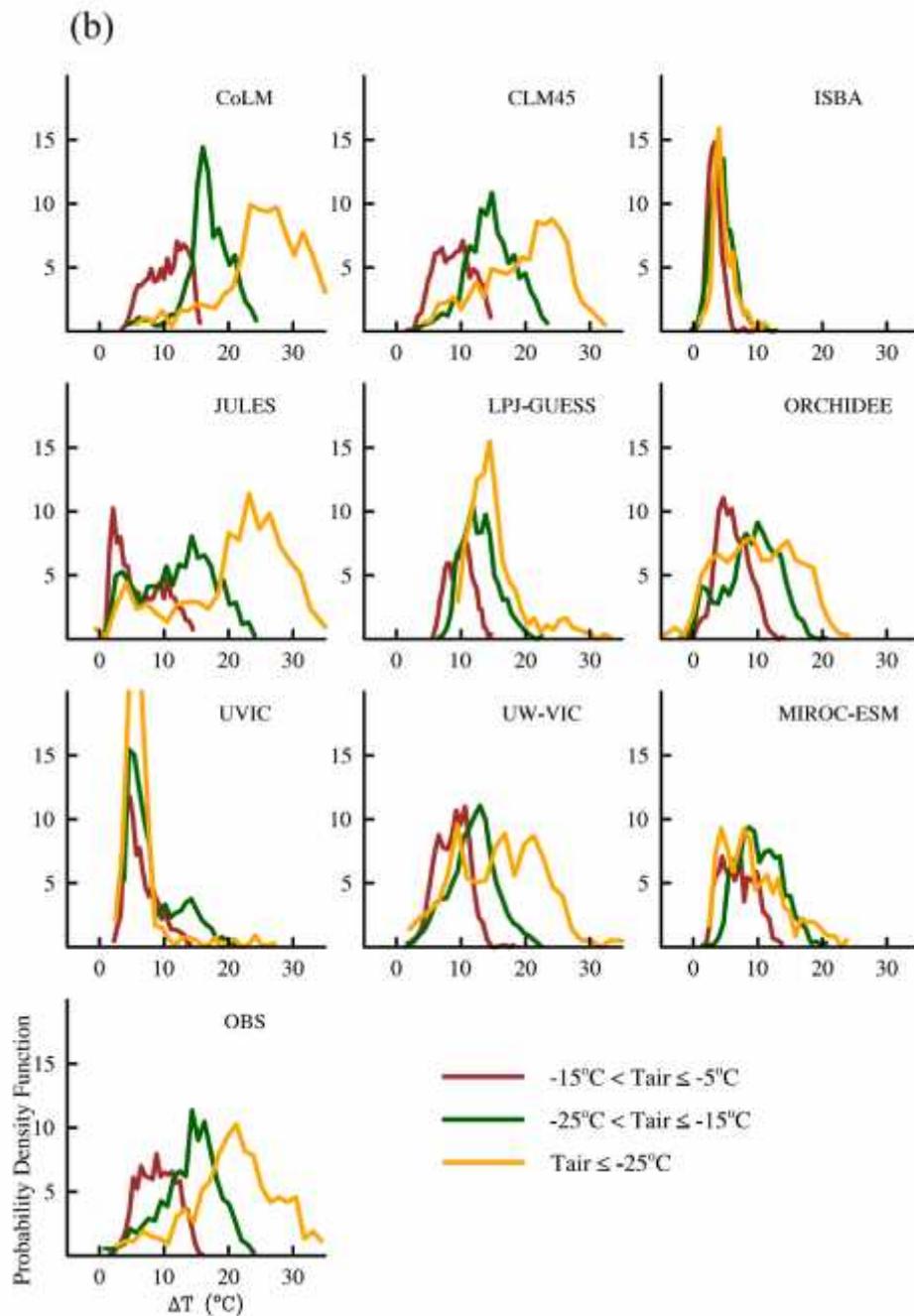


1
2



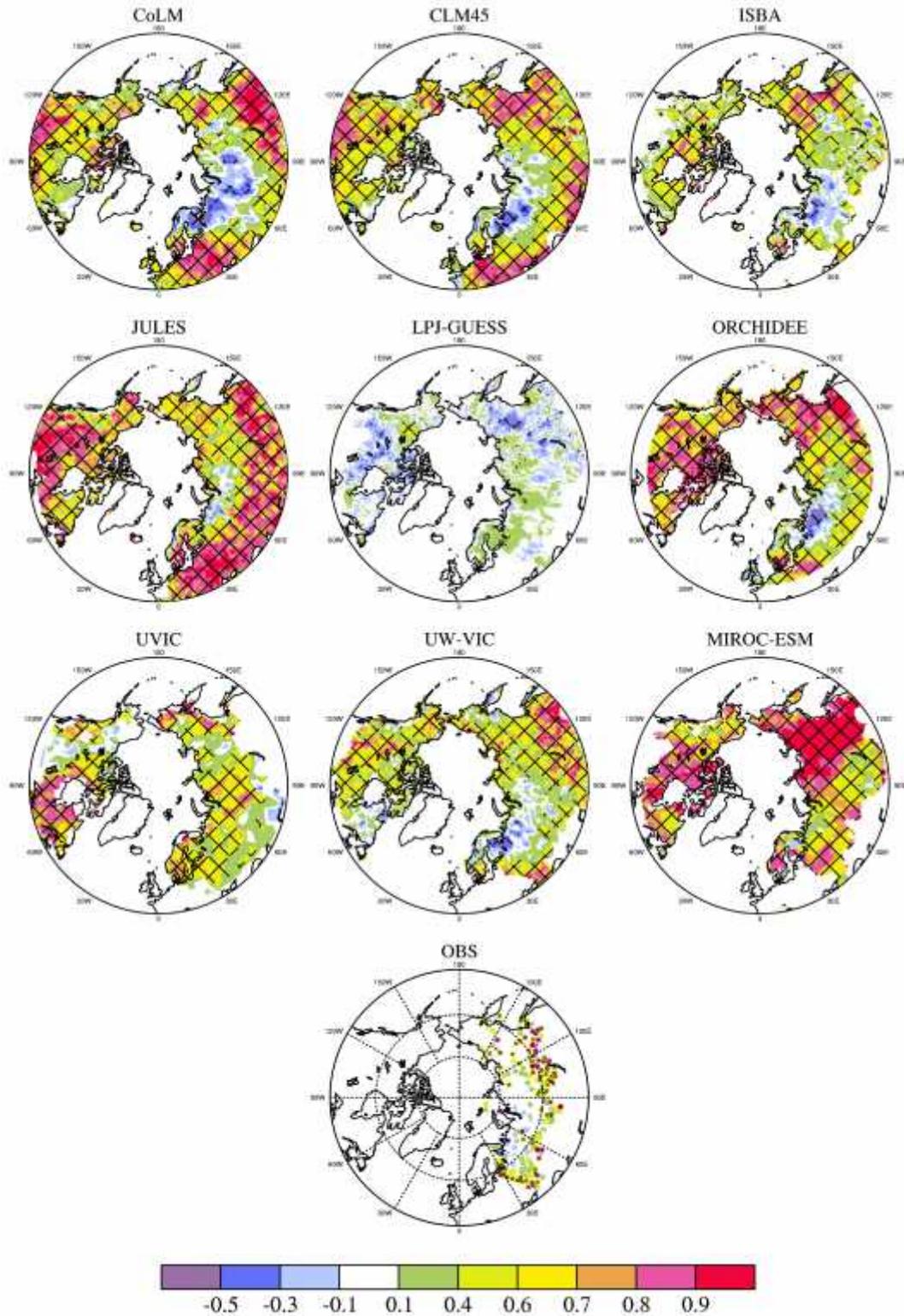


1



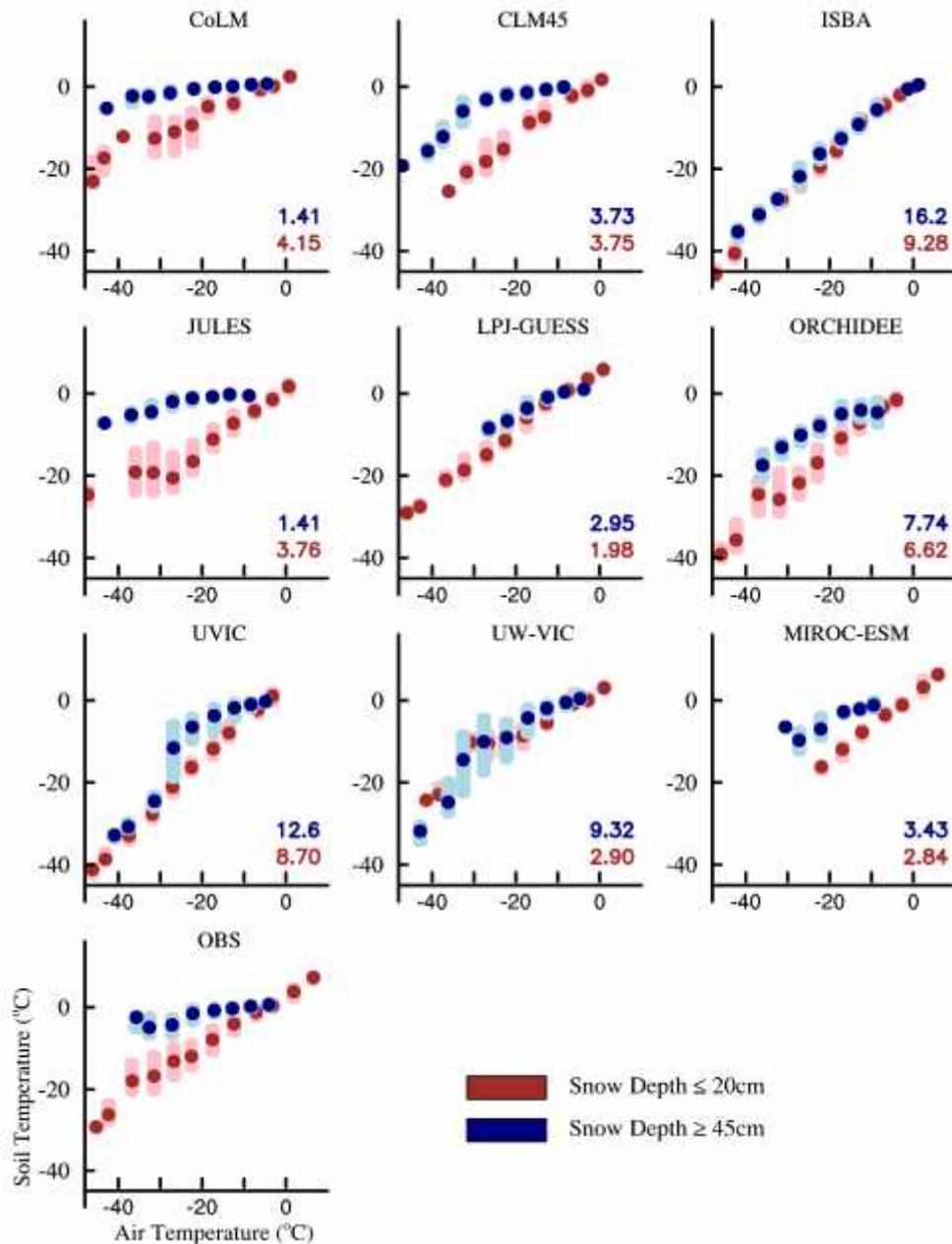
1

2 **Figure 2.** Conditional probability density functions (PDFs) of ΔT ($^{\circ}\text{C}$), the difference
 3 between soil temperature at 20 cm depth and air temperature for (a) different snow depth
 4 classes and (b) air temperature regimes, for winter 1980-2000.



1
2
3
4
5
6

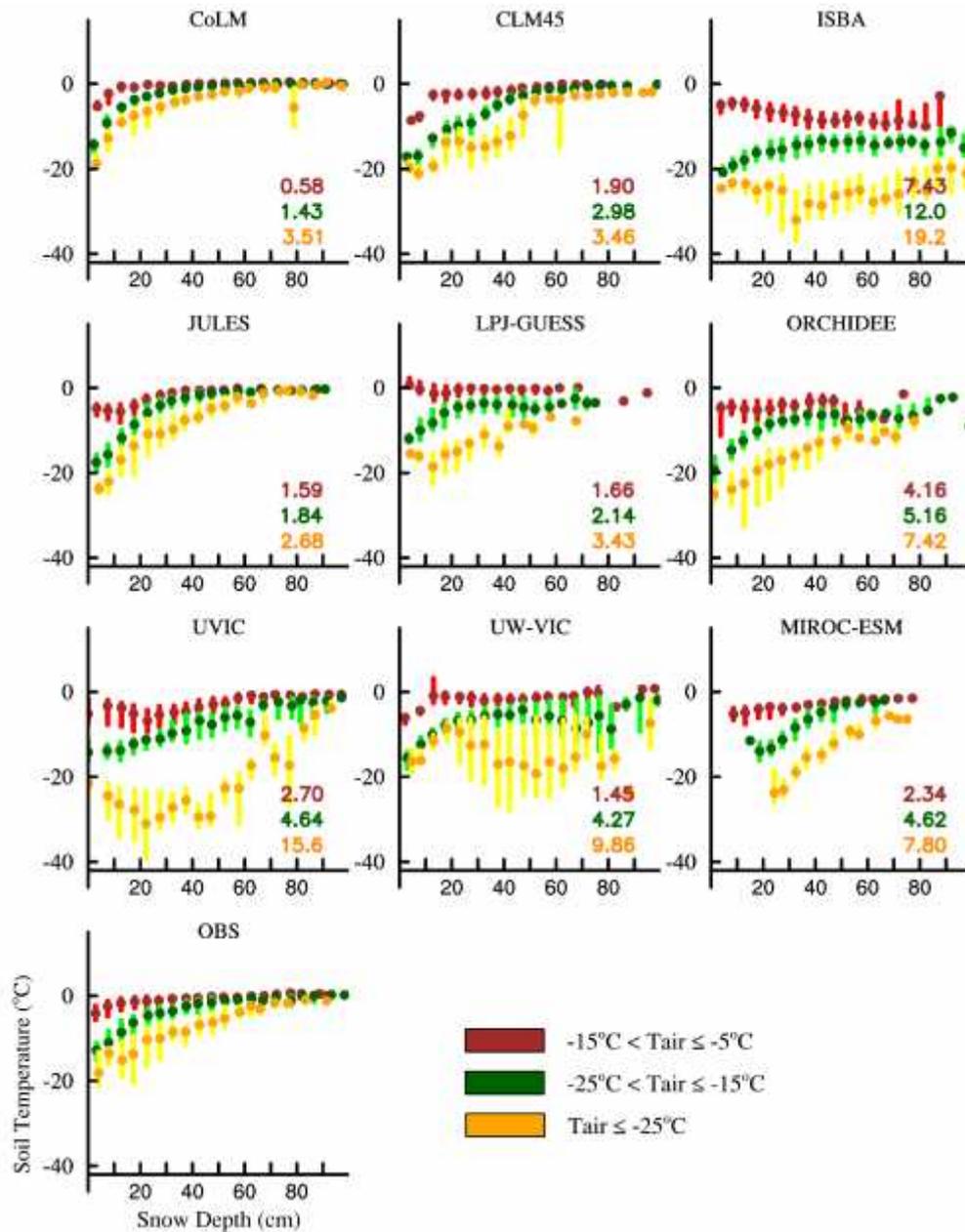
Figure 3. Spatial maps of the correlation coefficients between snow depth and T , the difference between soil temperature at 20 cm depth and air temperature for winter 1980-2000. Regions with greater than 95% significance are hashed.



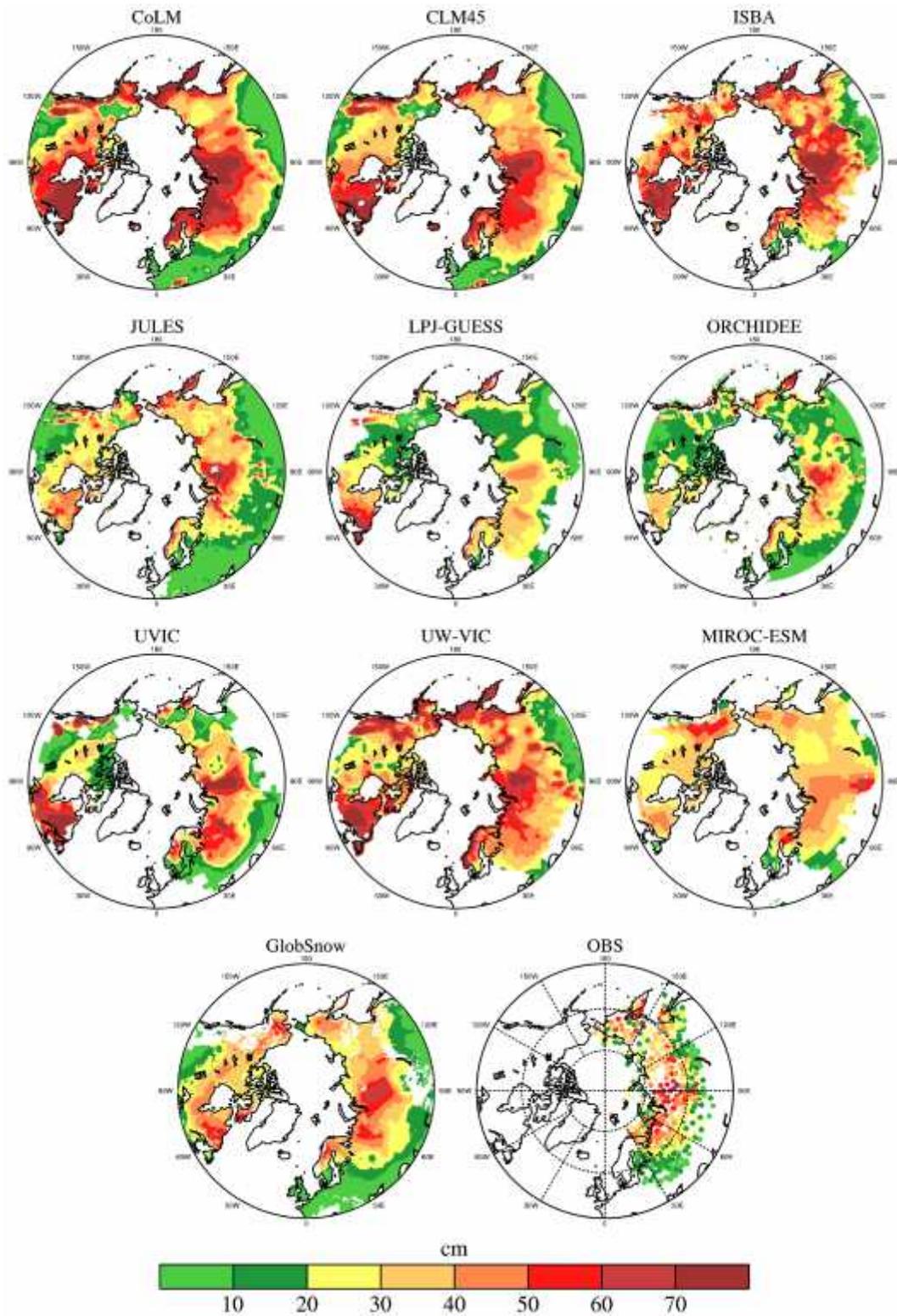
1

2 **Figure 4.** Variation of soil temperature at 20 cm depth (°C) with air temperature (°C) for
 3 winter 1980-2000. The dots represent the medians of 5°C air temperature bins and the upper
 4 and lower bars indicate the 25th and 75th percentiles, calculated from all Russian station grid
 5 points (n=268) and 21 individual winters. The numbers in each model panel indicate the
 6 RMSE between the observed and modeled relationship. Color represents different snow depth
 7 regimes.

8



1
 2 **Figure 5.** Variation of soil temperature at 20 cm depth ($^{\circ}\text{C}$; y axis) with snow depth (cm) for
 3 winter 1980-2000. The dots represent the medians of 5 cm snow depth bins and the upper and
 4 lower bars indicate the 25th and 75th percentiles, calculated from all Russian station grid points
 5 ($n=268$) and 21 individual winters. The numbers in each model panel indicate the RMSE
 6 between the observed and modeled relationship. Color represents different air temperature
 7 regimes.



1

2

3 **Figure 6.** Spatial maps of snow depth (cm) for winter 1980-2000.

1 **Supplementary Tables**

2 **SI Table 1.** Climate data sets used to drive each model.

Model	Climate forcing data
CLM4.5	CRUNCEP4 ¹
CoLM	Princeton ²
ISBA	WATCH (1901-2010) ³
JULES	WATCH (1901-2001) ³
LPJ-GUESS	CRU TS 3.1 ⁴
MIROC-ESM	CMIP5 Drivers ⁵ , WATCH (1901-1978) ³
ORCHIDEE	WFDEI (1978-2009) ⁶
UVic	CRUNCEP4 ¹ , CRU ⁷ , UDeI ⁸
UW-VIC	NCEP-NCAR ⁹

3

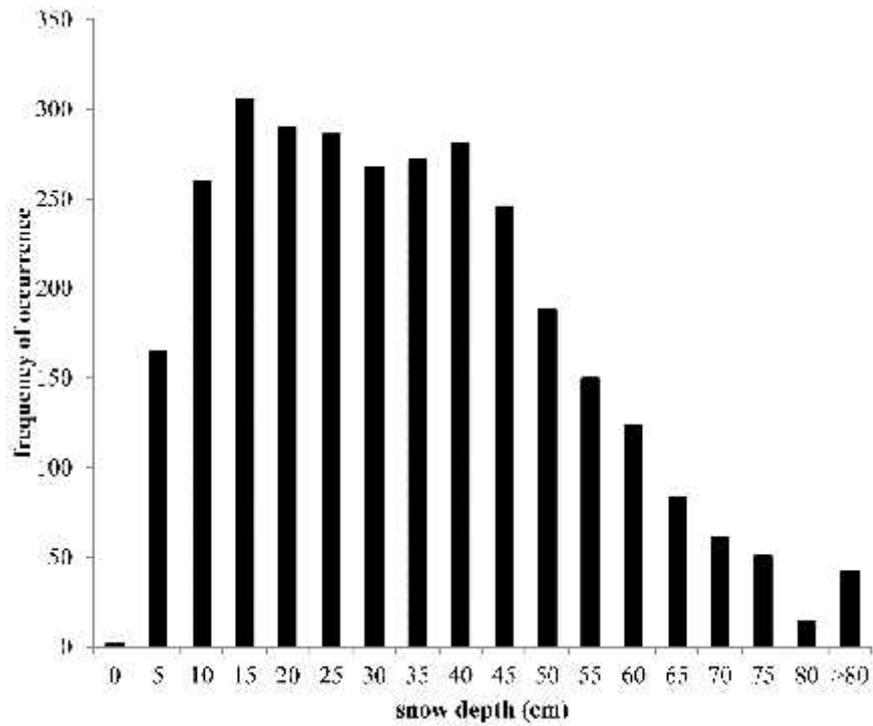
4 ¹Viovy and Ciais (2011) (<http://dods.extra.cea.fr/>)5 ²Sheffield et al. (2006) (<http://hydrology.princeton.edu/data/pgf.php>)6 ³Weedon et al. (2011) (<http://www.waterandclimatechange.eu/about/watch-forcing-data-20th-century>)7 ⁴Harris et al. (2013), University of East Anglia Climate Research Unit8 ⁵Watanabe et al. (2011)9 ⁶http://www.eu-watch.org/gfx_content/documents/README-WFDEI.pdf10 ⁷Mitchell and Jones (2005) for temperature11 ⁸Willmott and Matura (2001) for precipitation12 ⁹Kalnay et al. (2006)

13

14 Harris, I., Jones, P.D., Osborn, T.J., and Lister, D.H.: Updated high-resolution grids of
15 monthly climatic observations. *Int. J. Clim.*, doi: 10.1002/joc.3711, 2013.16 Kalnay, E. et al.: The NCEP Climate Forecast System. *J. Clim.*, 19, 3483-3517, 2006.17 Mitchell, T.D., and Jones, P.D.: An improved method of constructing a database of monthly
18 climate observations and associated high-resolution grids, *Int. J. Clim.*, 25(6), 693-712, doi:
19 10.1002/joc.1181, 2005.20 Sheffield, J., Goteti, G., and Wood, E.F.: Development of a 50-yr high-resolution global
21 dataset of meteorological forcings for land surface modeling, *J. Clim.*, 19, 3088-3111,
22 2006.23 Viovy, N. and Ciais, P.: CRUNCEP data set for 1901–2008, Tech. Rep. Version 4, Labora-
24 toire des Sciences du Climat et de l'Environnement, 4078, 4122, 2011.25 Watanabe, S. et al.: MIROC-ESM 2010: model description and basic results of
26 CMIP5-20c3m experiments. *Geosci. Model Dev.*, 4, 845–872,
27 doi:10.5194/gmd-4-845-2011, 2011.

-
- 1 Weedon, G.P., Gomes, S., Viterbo, P., Shuttleworth, W.J., Blyth, E., Österle, H., Adam, J.C.,
2 Bellouin, N., Boucher, O., and Best, M.: Creation of the WATCH Forcing data and its use
3 to assess global and regional reference crop evaporation over land during the twentieth
4 century. *J. Hydromet.*, 12, 823-848, doi: 10.1175/2011JHM1369.1, 2011.
- 5 Willmott, C.J., and Matsuura, K.: Terrestrial air temperature and precipitation: monthly and
6 annual time series (1950-1999) (version 1.02), Center for Climate Research, University of
7 Delaware, Newark, DE, 2001.
- 8

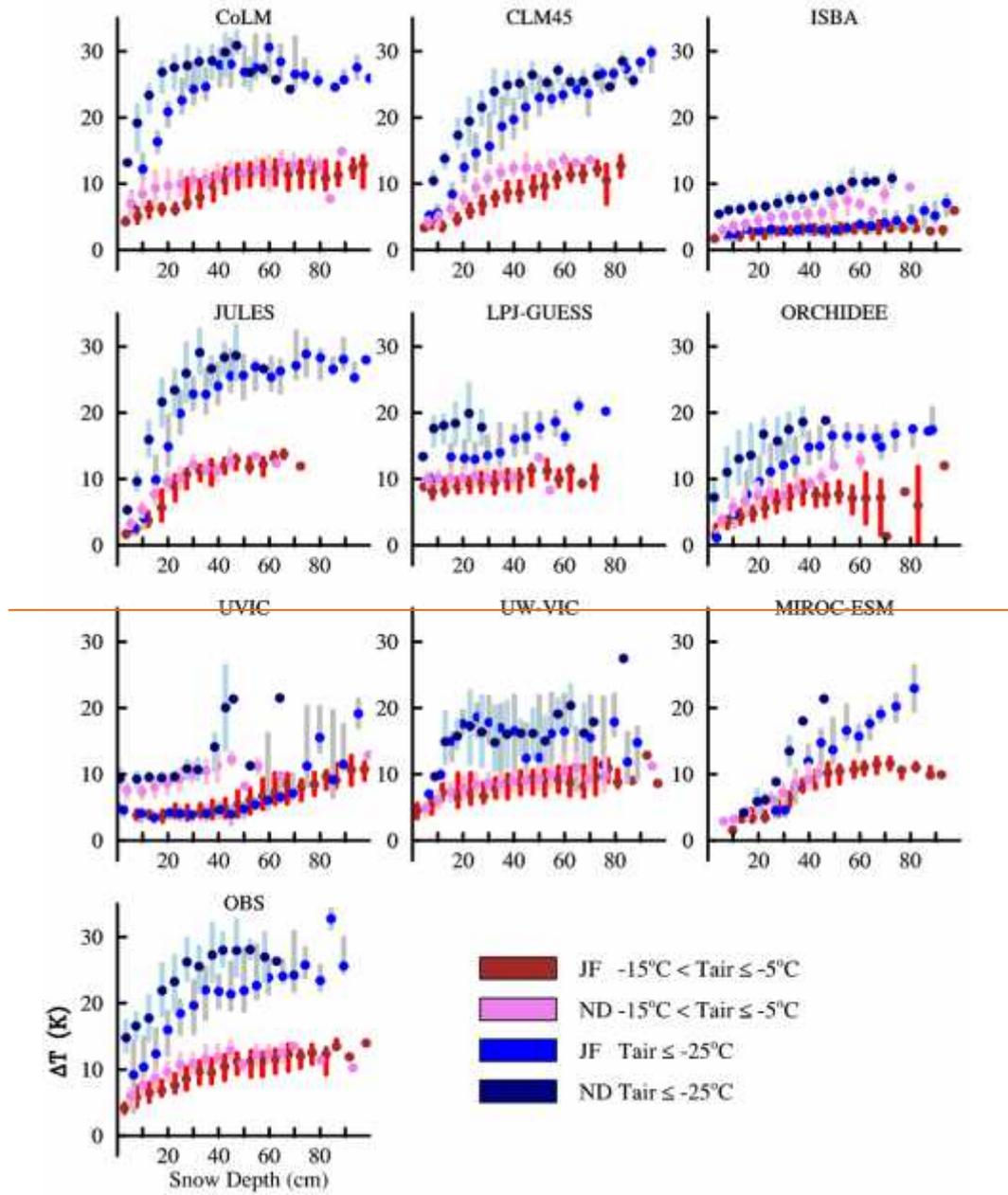
1 **Supplementary Figures**



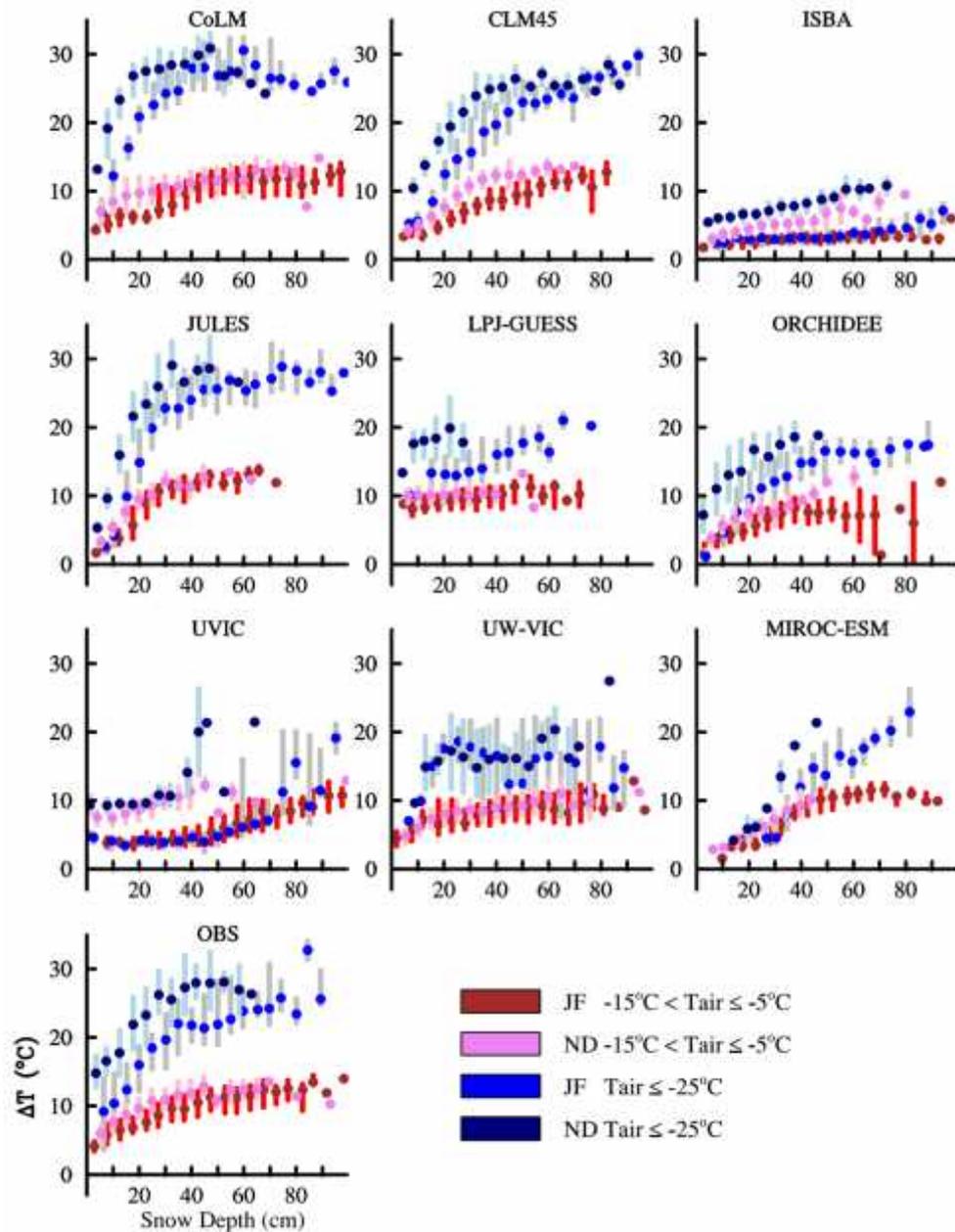
2

3 **SI Figure 1.** Histogram of seasonal winter mean snow depth from 268 Russian stations

4 between 1980-2000.

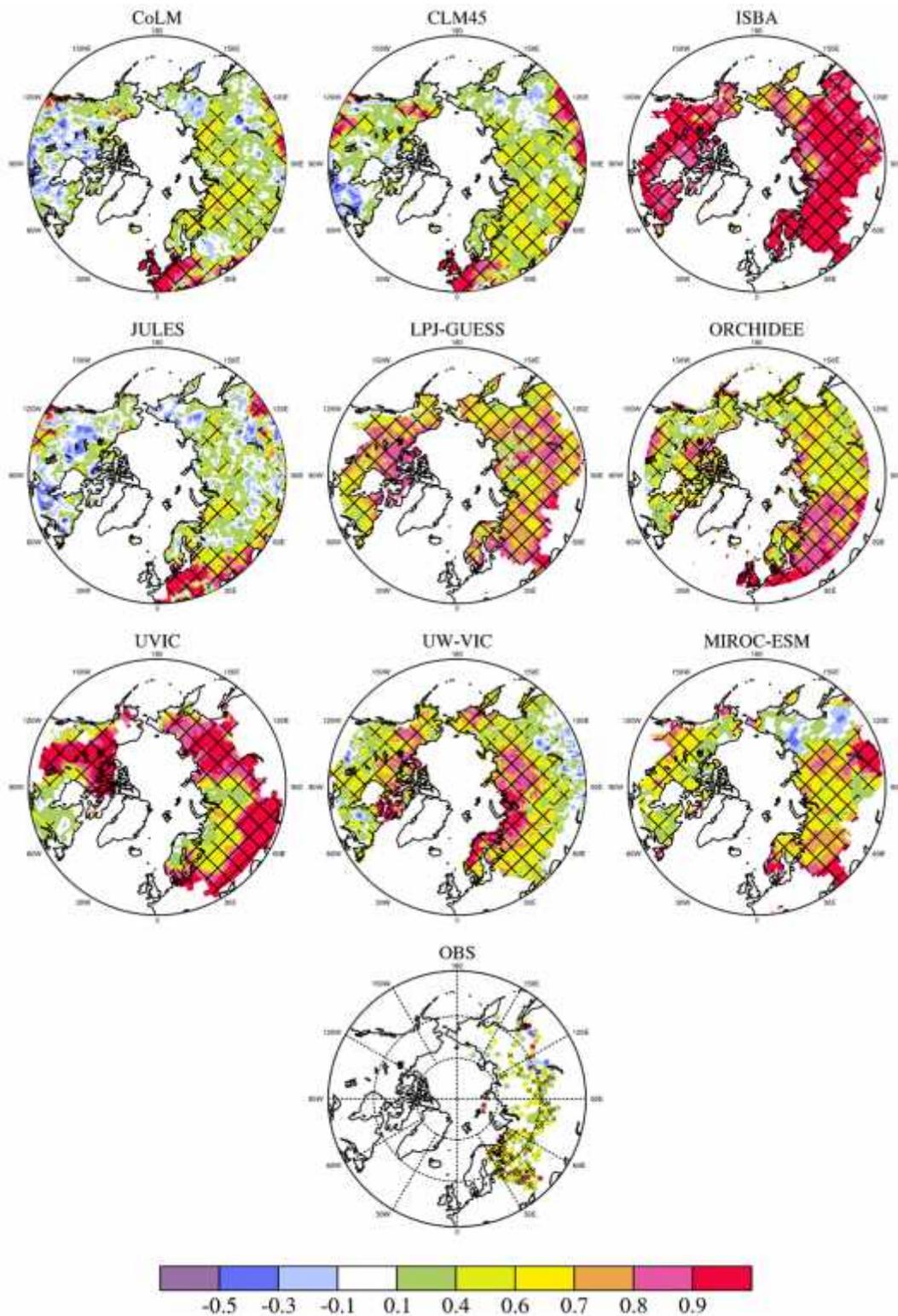


1



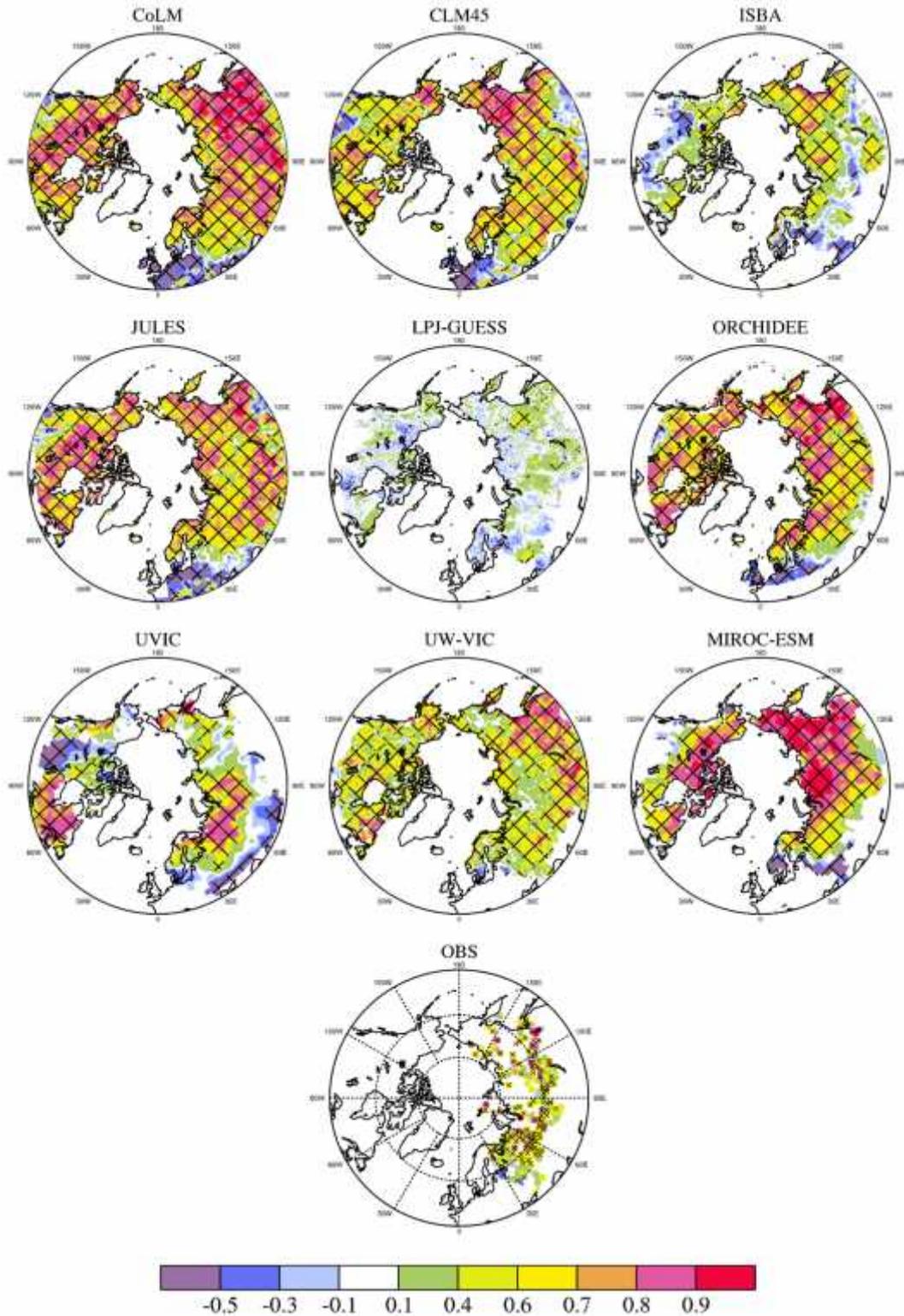
1

2 **SI Figure 2.** Variation of ΔT ($^{\circ}\text{C}$) (the difference between soil temperature at 20 cm depth
3 and air temperature) with snow depth (cm) for winter 1980-2000. The dots represent the
4 medians of 5 cm snow depth bins and the upper and lower bars indicate the 25th and 75th
5 percentiles, calculated from all Russian station grid points ($n=268$) and 21 individual winters.
6 Color represents two different air temperature regimes (reddish: $-15^{\circ}\text{C} < \text{AirT} \leq -5^{\circ}\text{C}$, blueish:
7 $\text{AirT} \leq -25^{\circ}\text{C}$) for early (Nov.-Dec.; ND) and late (Jan.-Feb.; JF) winter.



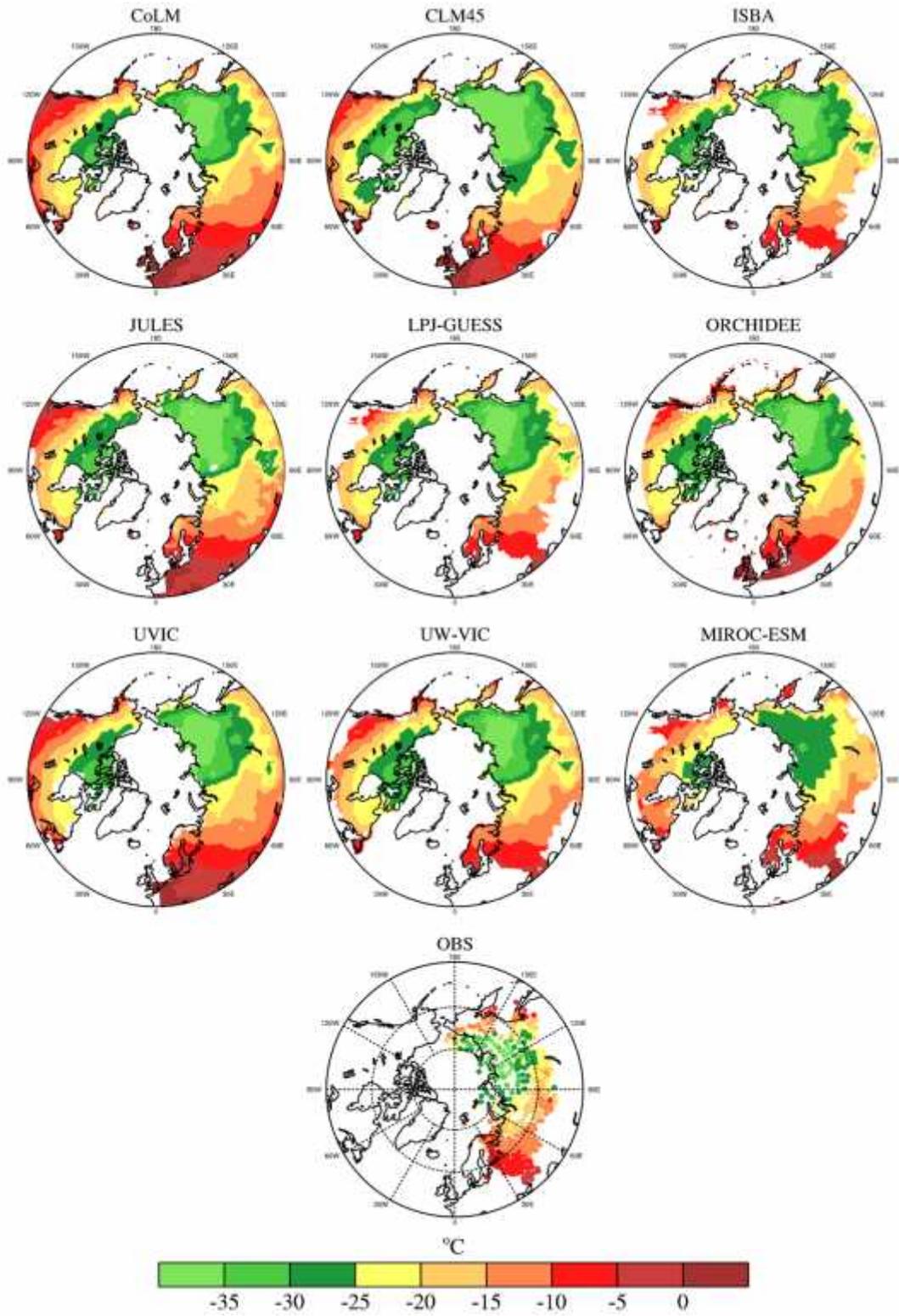
1

2 **SI Figure 3.** Spatial maps of the correlation coefficients between soil temperature at 20 cm
 3 depth and air temperature for winter 1980-2000. Regions with greater than 95% significance
 4 are hashed.



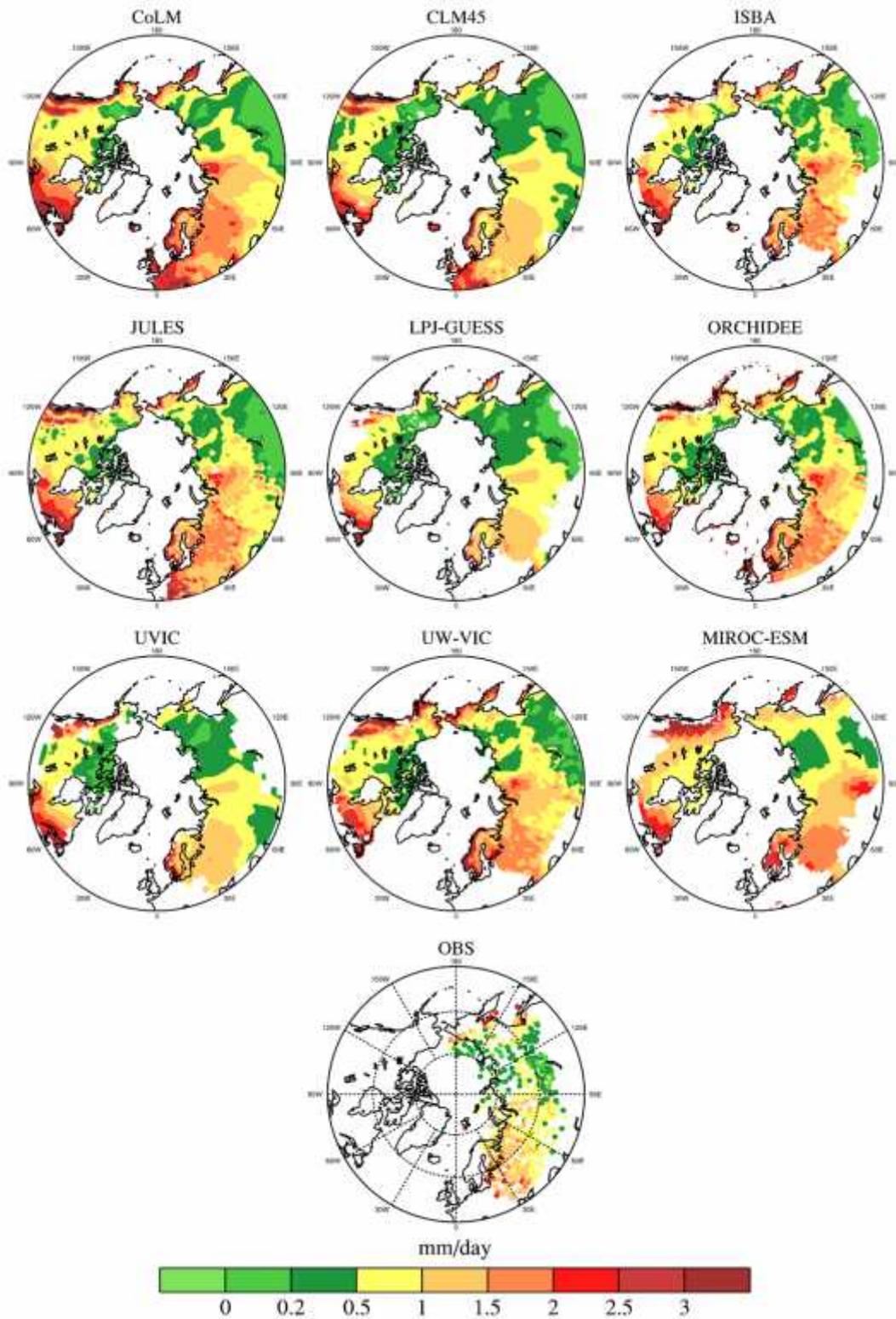
1

2 **SI Figure 4.** Spatial maps of the correlation coefficients between soil temperature at 20 cm
 3 depth and snow depth for winter 1980-2000. Regions with greater than 95% significance are
 4 hashed.

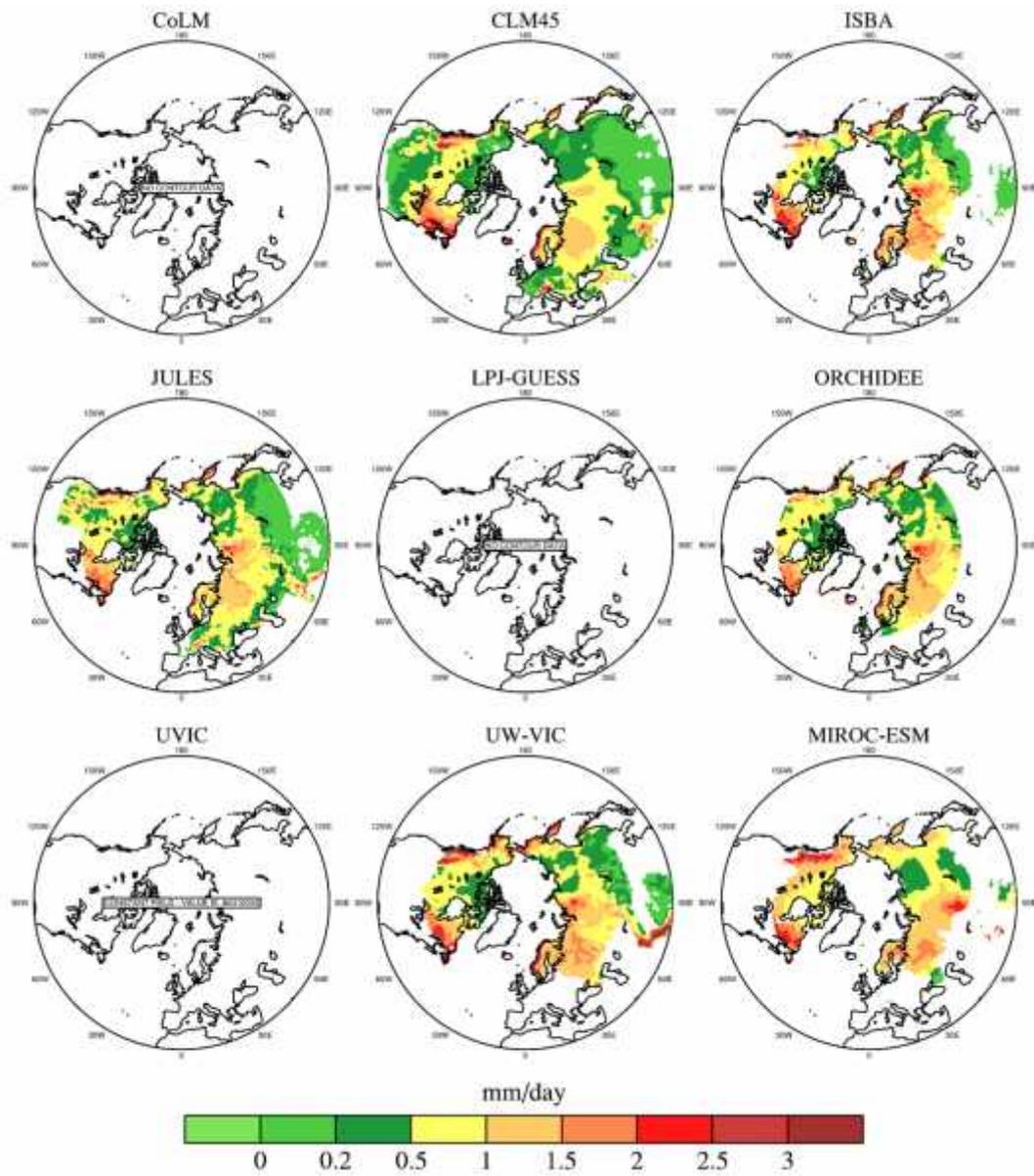


1

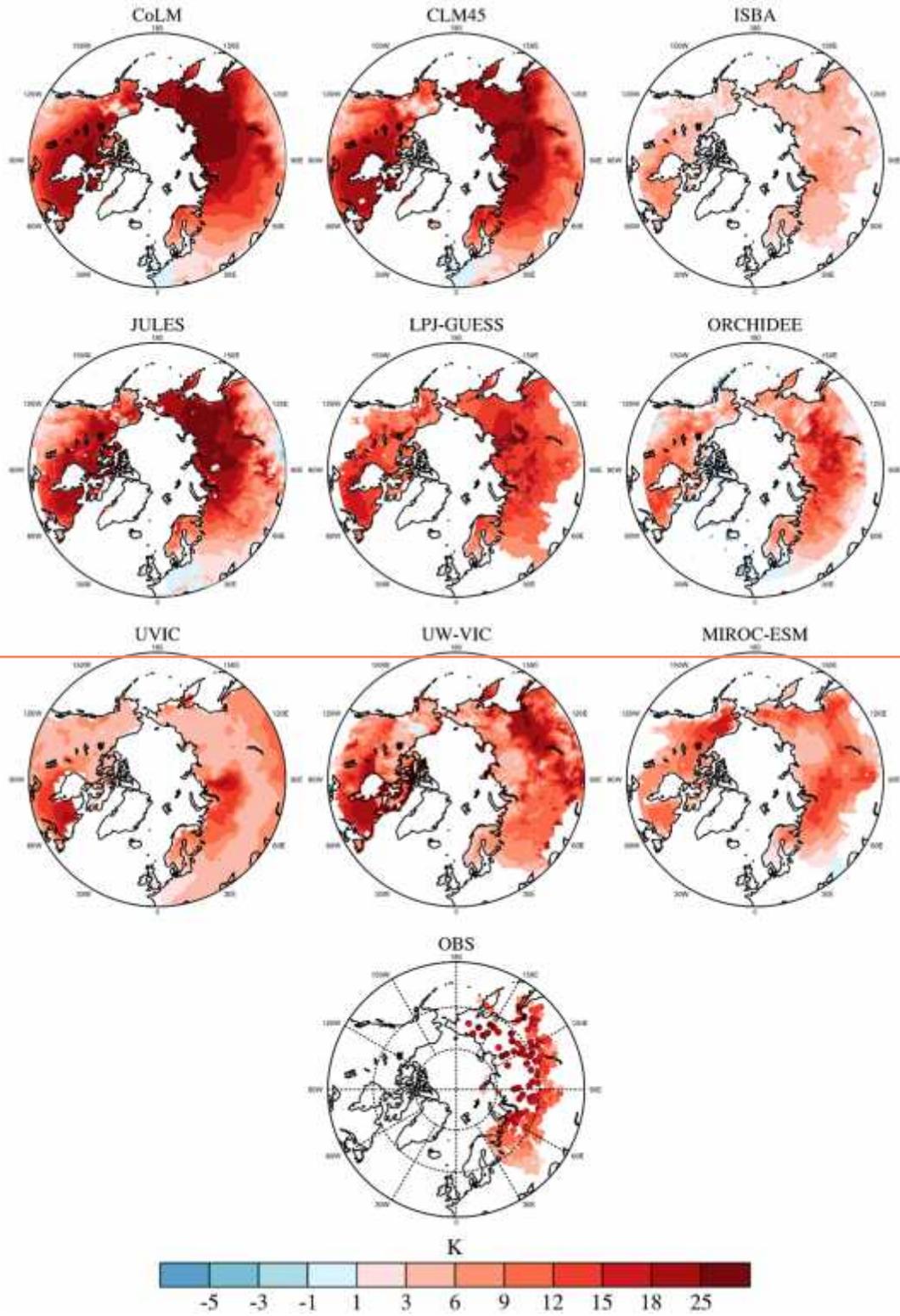
2 **SI Figure 5.** Spatial maps of mean air temperature (°C) for winter 1980-2000.



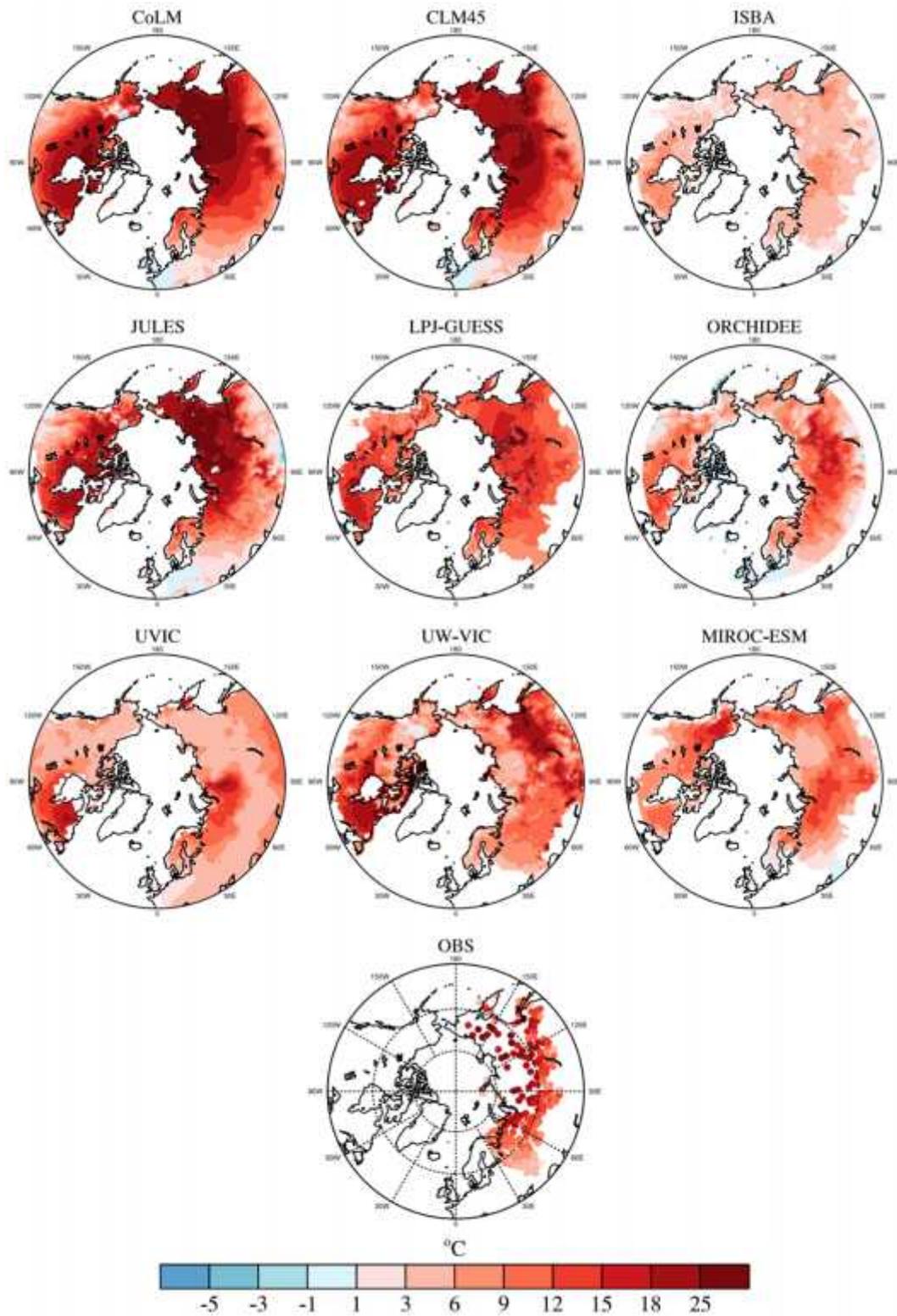
SI Figure 6. Spatial maps of mean precipitation (mm/d) for winter 1980-2000.



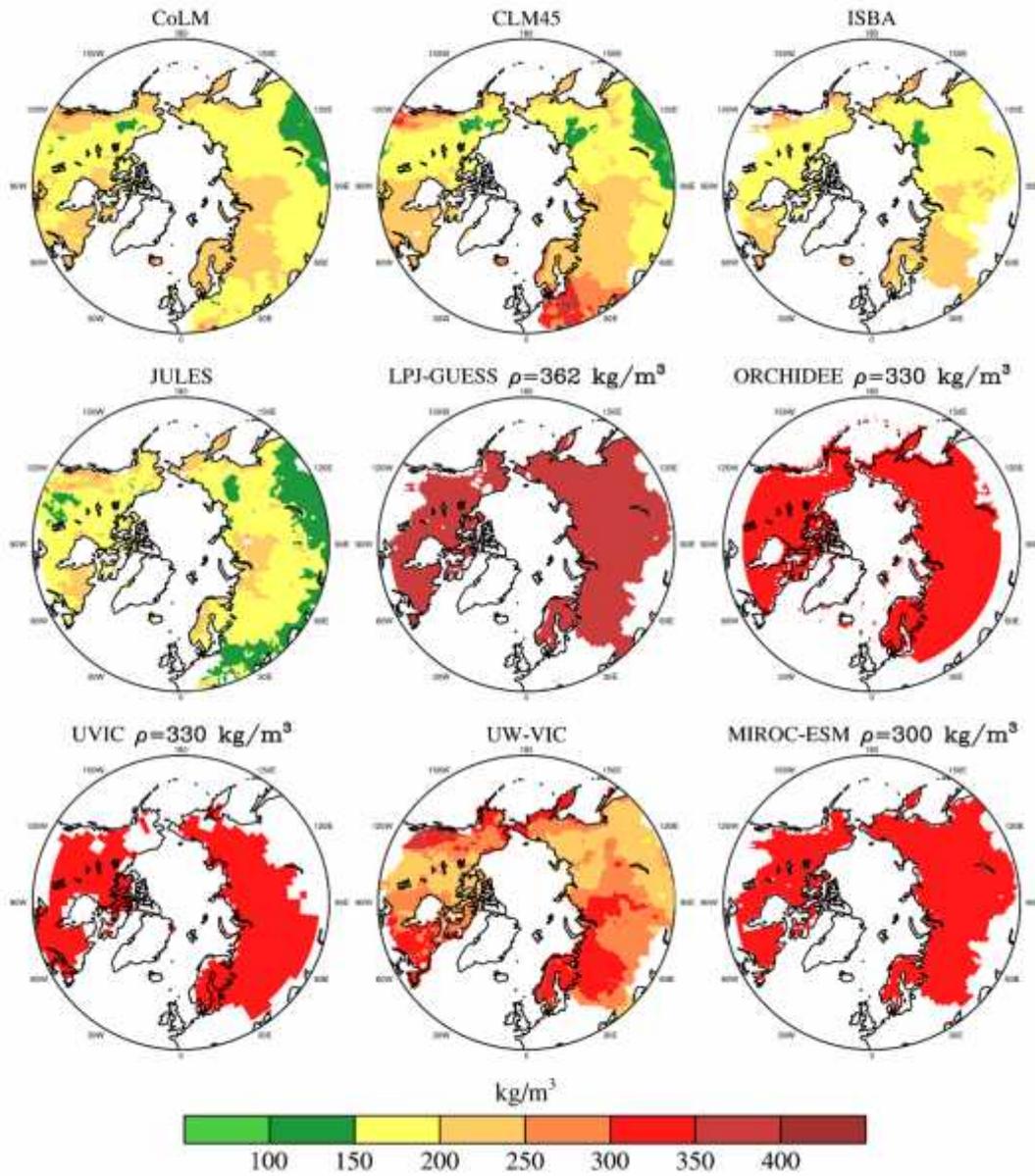
SI Figure 7. Spatial maps of snow fall (mm/d) for winter 1980-2000.



1



1
 2 **SI Figure 8.** Spatial maps of T (K°C) (difference between soil temperature at 20 cm depth
 3 and air temperature) for winter 1980-2000.



SI Figure 9. Spatial maps of snow density (kg m^{-3}) (calculated by the quotient of snow water equivalent and snow depth, if not directly output) for winter 1980-2000.