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

Wenli Wang[1+], Annette Rinke[1,2+], John C. Moore[1], Duoying Ji[1*], Xuefeng Cui[3], Shushi Peng[4,17,18], David M. Lawrence[5], A. David McGuire[6], Eleanor J. Burke[7], Xiaodong Chen[21], Bertrand Decharme[9], Charles Koven[10], Andrew MacDougall[11], Kazuyuki Saito[12,15], Wenxin Zhang[13, 19], Ramdane Alkama[9,16], Theodore J. Bohn[8], Philippe Ciais[18], Christine Delire[9], Isabelle Gouttevin[4], Tomohiro Hajima[12], Gerhard Krinner[4,17], Dennis P. Lettenmaier[8], Paul A. Miller[13], Benjamin Smith[13], Tetsuo Sueyoshi[14], Artem B. Sherstiukov[20]

[1]College of Global Change and Earth System Science, Beijing Normal University, Beijing, China

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

[3]School of System Science, Beijing Normal University, Beijing, China

[4]CNRS, LGGE, Grenoble, France

[5]National Center for Atmospheric Research, Boulder, USA

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

[7]Met Office Hadley Centre, Exeter, UK

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

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

[10]Lawrence Berkeley National Laboratory, Berkeley, CA, USA

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

[12]Japan Agency for Marine-Earth Science and Technology, Yokohama, Japan

[13]Department of Physical Geography and Ecosystem Science, Lund University, Lund, Sweden

[14]National Institute of Polar Research, Tachikawa, Japan

[15]University of Alaska Fairbanks, Fairbanks, AK, USA

[16]l'Institute for Environment and Sustainability (IES), Ispra, Italy

[17]Université Grenoble Alpes, LGGE, Grenoble, France

[18]LSCE, CEA/CNRS/UVSQ, Saclay, France

[19]Center for Permafrost (CENPERM), Department of Geosciences and Natural Resource Management, University of Copenhagen, Copenhagen, Denmark

[20]All-Russian Research Institute of Hydrometeorological Information-World Data Centre, Obninsk, Russia

[21]Department of Civil and Environmental Engineering, University of Washington, Seattle, WA, USA

+ *these authors contributed equally to this work*

* *correspondence to*: Duoying Ji (duoyingji@bnu.edu.cn)

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

1 Introduction

Present-day permafrost simulations by global climate models are limited and future projections contain high, model-induced uncertainty (e.g., Slater and Lawrence, 2013; Koven et al., 2013). Most of the model biases and across-model differences in simulating permafrost area are due to biased atmospheric simulation e.g. of air temperature and precipitation, biased simulation of snow and soil temperature, and the coupling between atmosphere and land-surface. In winter, the snow insulation effect is a key process for the air-soil temperature coupling. Its strength depends on the snow depth, areal coverage, snow density and conductivity (see overview by Zhang, 2005). Many individual model studies have shown the strong impact of snow parameters on soil temperature simulations (e.g., ~~recently,~~ Langer et al., 2013; Dutra et al., 2012; Gouttevin et al., 2012; Essery et al., 2013; Wang et al., 2013; Jafarov et al., 2014). 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 tempearture~~regime~~ has been demonstrated by many model studies (e.g., Bartlett et al., 2006; Saha et al., 2006; Vavrus, 2007; Nicolsky et al., 2007; Dankers et al., 2011; Gouttevin et al., 2012). Winter soil temperature can change by up to 20 K simply by varying the snow thermal conductivity by 0.1-0.5 W m^{-1} K^{-1} (Cook et al., 2008). The snow insulation effect also plays an important role for the Arctic soil temperature response to climate change and therefore for future near-surface permafrost thawing and soil carbon vulnerability (e.g., Schuur et al., 2008). Shallower snow can reduce soil warming while shorter snow season can enhance soil warming (Lawrence and Slater, 2010). The model skill in atmosphere-soil coupling with the concomitant snow cover in the Arctic is an important factor in the assessment of limitations and uncertainty of carbon mobility estimates (Schaefer et al., 2011).

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

Comment [ZY1]: Parametrization or Properties?

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

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

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

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

2 Data and Analysis

2.1 Models

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

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

The simulations were generally run for the period 1960-2009, although some simulations were stopped a few years earlier. Each model team was free to choose appropriate driving data sets for weather and climate, atmospheric CO_2, nitrogen deposition, disturbance, land cover, soil texture, etc. However, the climate forcing data (surface pressure, surface incident longwave and shortwave radiation, near-surface air temperature, wind and specific humidity, rain and snowfall rates) are from gridded observational datasets (e.g. CRUNCEP, WATCH) (SI Table 1). The exception is MIROC-ESM, which was run as a fully-coupled model, forced by its own simulated climate. Mean annual air temperature simulated by ~~of the~~ MIROC-ESM ~~simulations~~ for the permafrost region was~~ere~~ within the range (-7.2 to 2.2°C) of the other forcing data sets used in this study and the trend in near-surface air temperature (+0.03°C yr^{-1}) was the same for all forcing data sets. However, MIROC-ESM had both the highest annual precipitation (range 433 to 686 mm) and the highest trend in annual precipitation (range -2.1 to +0.8 mm yr^{-1}) among the forcing data sets.

The spatial domain of interest is the Northern Hemisphere permafrost land regions. Our analysis is based on the $0.5^o \times 0.5^o$ resolution gridded driving and modeled data for winter (DJF) 1980-2000.

2.2 Observations

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

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

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

2.3 Analysis Methods

Our analysis is focused on the common winter (DJF) condition, although snow can begin in November and end at the beginning of May, but we checked that a different winter definition (NDJFMA) does not qualitatively change any of the inter-variables relationships found. The focus in our study is on the evaluation of the simulated air-soil temperature relationships, modulated by snow depth. For this, we analyze the winter mean as well as the interannual variability (expressed as the standard deviation) of ~~4~~four key variables: near-surface air temperature (T_{air}), near-surface soil temperature (soil temperature at 20 cm depth; T_{soil}), snow depth (d_{snow}), and the difference between T_{soil} and T_{air}. This difference ΔT ($\Delta T = T_{soil} - T_{air}$) is called the air-soil temperature difference. By limiting our analysis to the winter only, we are able to attribute the across-model and model-to-observation differences in ΔT primarily to snow insulation effects. In winter, the effects of other factors (e.g. soil moisture, texture) on ΔT are much smaller than that of snow. Ground surface temperatures were not recorded in the Russian data set, but 20 cm soil depth temperatures were. To test how sensitive are results using 20 cm temperatures instead of ground surface, we also analyzed model simulated temperature differences between ground surface and T_{air}, and found no qualitative differences, hence justifying use of 20 cm observations.

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

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

3 Results

3.1 Relationship between air – soil temperature difference and snow depth

The air-soil temperature difference (ΔT) - snow depth (d_{snow}) relationship in winter (Fig. 1) shows in the Russian station observations an increase of ΔT with increasing d_{snow}. The data exhibit a linear relation between ΔT and d_{snow} at relatively shallow snow depths with a trend towards asymptotic behavior at thicker snow, which is in agreement with earlier findings (Zhang, 2005; Ge and Gong, 2010; Morse et al., 2011). There is also significant scatter in the observation-based relationship indicated by the inter-quartile range in ΔT of 1.5-8.5 K at specific snow depth and air temperature regimes, likely resulting from complicating factors such as snow pack density and moisture content variability over the winter, as well as observational errors.

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

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

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

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

If we evaluate the models with respect to this observed impact of T_{air} ~~to~~on the ΔT vs. d_{snow} relationship, we demonstrate that some models (CLM4.5, CoLM, JULES) are better able to replicate the effect than others (LPJ-GUESS, MIROC-ESM, ORCHIDEE, UW-VIC) (Fig. 1). The latter do not fully replicate the larger ΔT under cold T_{air} conditions. CLM4.5, CoLM and JULES capture a larger ΔT for colder T_{air} for a given d_{snow} in agreement with the observations. However, for shallow snow JULES simulates an increase of ΔT with increasing d_{snow} for all temperature ranges that is twice as large as observations. Two models (ISBA, UVic) clearly fail in this evaluation. Poor model performance in reflecting T_{air} influence on the ΔT vs. d_{snow} also manifests itself in regime separation of the PDFs (Fig. 2b). Some models do not separate the ΔT regimes under different T_{air} conditions well or at all (ISBA, LPJ-GUESS, MIROC-ESM, UVic), while others cannot capture the observed cold temperature regime features (i.e., too broad PDFs and shifts towards smaller modal values; ORCHIDEE, UW-VIC). The three models with reasonable inter-variable relations (CLM4.5, CoLM, JULES) also capture the regime separation in the PDFs. These three models as well as LPJ-GUESS and ORCHIDEE also represent the observed greater insulation of early winter snow packs under cold conditions (SI Fig. 2).

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

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

3. 2 Variability of soil temperature with air temperature and snow depth

Next we assess whether or not the models can correctly reproduce the interannual near-surface soil temperature (T_{soil}) variability in relation to snow depth (d_{snow}) and near-surface air temperature (T_{air}) variability. Previous studies have noted that the strength of relationship between T_{soil} and T_{air} is modulated by d_{snow} and the snow insulation effect increases only up to a limiting depth beyond which extra snow makes little difference to soil temperatures (Smith and Riseborough, 2002; Sokratov and Barry, 2002; Zhang, 2005; Lawrence and Slater, 2010). Zhang (2005) reported that the limiting snow depth is approximately 40 cm.

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

Figure 4 clearly shows that some models (CoLM, CLM45, JULES) can well capture this difference~~modification of the T_{soil} vs. T_{air} relation by snow depth regime well~~. Their regression slopes for thick and thin snow are well separated and in agreement with those from the observed relationship (Table 2). The RMSE of their modeled T_{soil} vs. T_{air} relationships from observations is smaller than 4°C. These models better reproduce the observed ΔT vs. d_{snow} relationship. Other models (LPJ-GUESS, MIROC-ESM, ORCHIDEE) strongly underestimate the increase of the T_{soil} vs. T_{air} regression slope for decreasing snow depth. They also produce a regression slope for thick snow more than twice as large as observations. Two models (ISBA, UVic) ~~fail here and~~ do not show any sensitivity in the T_{soil} vs. T_{air} relation to snow conditions (Fig.4, Table 2). Another measure quantitatively confirms the same models behavior: The observed average d_{snow} in the shallow snow regime is 13.7 cm and that for the thick snow regime is 58.5 cm, so we would expect, if near-surface T_{air} and conductivities were equal in both snow depth classes, a ratio between the slopes for shallow and thick snow would be~~of~~ 4.3. CLM4.5, CoLM, and JULES reproduce this observed variation in the T_{soil} vs. T_{air} relation better than others (Table 2). JULES and CoLM indicate a factor of 4 change, while CLM4.5 indicates a factor of 2 change. Other models (LPJ-GUESS, MIROC-ESM, ORCHIDEE) underestimate the increase of the regression slope for decreasing snow depth; they simulate only a factor change of about 1.5. The two models with unrealistic ΔT vs. d_{snow} relationships (ISBA, UVic) also fail in this evaluation of their T_{soil} vs. T_{air} relationship. They simulate a too strong sensitivity of T_{soil} to T_{air} (regression slopes larger than 0.9°C/°C, R^2>0.7; Table 2) that are almost completely independent of the snow depth regimes, particularly in ISBA, which is not consistent with observations. These models' spatial correlation patterns between T_{soil} and T_{air} also differ greatly from the observations and

Comment [ZY2]: Hard to understand

the other models (SI Fig. 3) and show very high positive correlation (r > 0.8) in most regions, as may be expected from the large regression slope shown in Fig. 4. The RMSE of their modeled T_{soil} vs. T_{air} relationships from observations reaches ca. 10°C .

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

The spatial patterns of the correlation coefficients between T_{soil} and T_{air} (SI Fig. 3) and between T_{soil} and d_{snow} (SI Fig. 4) show a relatively large across-model scatter in many regions. Obvious outliers in the T_{soil} vs. T_{air} correlation maps are ISBA and UVic which strongly overestimate the correlation (r > 0.9) over most of the Arctic. This indicates an underestimated snow insulation effect, and confirms the weak insulation in both models, which we already discussed based on their underestimated ΔT (Fig. 1) and weak correlation between ΔT and d_{snow} (Fig. 3). Other models (LPJ-GUESS, ORCHIDEE, UW-VIC) also overestimate the correlation in some regions (e.g. western Russian Arctic, r > 0.7). Most of the simulated maps of T_{soil} vs. d_{snow} correlation agree with the observations on a strong positive correlation in East Siberia. This is a region of relatively shallow snow (10-40 cm; Fig. 6) and there T_{soil} is very sensitive to variations in snow depth (e.g., Romanovsky et al., 2007). Comparing both simulated correlation maps, it is obvious that in this region, T_{soil} correlates more strongly with d_{snow} than with T_{air}, in agreement with the Russian data and earlier studies (Romanovsky et al., 2007; Sherstyukov, 2008).

4 Roles of atmospheric forcing and model processes

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

4.1 Atmospheric forcing and snow depth

4.1.1 Air temperature and precipitation

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

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

4.1.2 Snow depth

The broad-scale spatial snow depth (d_{snow}) patterns are similar across the models and show general agreement with the observed patterns (Fig. 6). The well-pronounced areas of maximum winter d_{snow} (50-100 cm) are in Scandinavia, the Urals, the West Siberian Plain, Central Siberian Highlands, the Far East, Alaskan Rocky mountains, and Labrador Peninsula and isle of Newfoundland. However, large regional across-model variability is obvious. Some models (JULES, LPJ-GUESS, ORCHIDEE, UVic) underestimate d_{snow}, while others (CLM4.5, CoLM, ISBA, UW-VIC) overestimate it (Fig. 6; Table 3). The model biases are quite similar with respect to station observations and GlobSnow data. It should be noted, that the models do not account for snowdrift. However, redistribution of snow due to wind is an important aspect, which makes comparison between in-situ measured and modeled snow depths difficult (e.g., Vionnet et al., 2013; Sturm and Stuefer, 2013; Gisnas et al., 2014).

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

4.2 Model processes

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

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

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

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

ISBA strongly underestimates ΔT, while strongly overestimating d_{snow}, compared with observations (Table 3, Fig. 6). However, ISBA uses the same atmospheric forcing data as JULES (accordingly the air temperature and precipitation are quite similar; SI Table 2). Also, the model's snow density (150-250 kg m^{-3}) is similar to other models (CLM45, CoLM, JULES) (SI Fig. 9) and in agreement with Zhong et al. (2013) who report snow density values of on 180-250 kg m^{-3} for tundra/taiga and 156-193 kg m^{-3} for alpine snow classes in winter. This apparent contradiction comes from the parameterization of snow cover fraction within each grid cell (SCF). The version of ISBA used here calculates a unique superficial soil temperature whether or not the soil is covered by snow and all the energy and radiative fluxes are area-weighted by SCF (equations 7 and 20 in *Douville et al*., 1995). In order to get reasonable albedos in snow-covered forests, as is necessary when ISBA is coupled to the CNRM-CM climate model, the parameterization gives very low SCF in the boreal forest (between 0.2 and 0.5). Hence, snow insulates only 20% to 50% of the grid cell, despite fairly high snow depths. The heat fluxes from the snow-covered fraction are averaged with the fluxes from the snow-free surface, strongly concealing the actual insulating effect of snow and underestimating it over the grid cell. Using the detailed snow model Crocus (Brun et al., 1992; Vionnet et al., 2012) with a SCF equal to 100% leads to an almost perfect simulation of near-surface soil temperature over Northern Eurasia (Brun et al., 2013). A similar experiment with ISBA and a SCF equal to 100% (Decharme et al., 2016) leads to good performances showing that the low ΔT in ISBA despite high snow depth in the present study is mostly due to this sub-grid snow fraction. Decharme et al. (2016) still showed that the ISBA results are further improved by updating the snow albedo and snow densification parameterization.

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

5 Permafrost area

Comment [ZY3]: This section is quite short compared to other sections. You may simply put to the above section and revise the section head as Discussion

Snow cover plays an important role in modulating the variations of soil thermodynamics, and hence near-surface permafrost extent (e.g., Park et al., 2015). Here we evaluate if there is a simple relationship between the simulated Northern hemisphere permafrost area and the sophistication and ability of the snow insulation component in the LSM to match observed snow packs. The simulated near-surface permafrost area varies greatly across the nine models in the hindcast simulation (1960-2009; Table 4). Some of the better performing snow insulation effect models (CLM4.5, JULES) simulate a near-surface permafrost area of 13.19 to 15.77 million km^2, which is comparable with the IPA map estimate (16.2 million km^2) (Brown et al., 1997; Slater and Lawrence, 2013). CoLM and ORCHIDEE, identified as reasonable models with respect to snow insulation, simulate much lower (7.62 million km^2) and higher (20.01 million km^2) areas, respectively. The main deficiency of CoLM is its too small soil depth (3.4 m) compared with CLM4.5 (45.1 m) despite having very similar snow modules (Table 1). However, ISBA, one of the two models that showed rather limited skill in representing snow insulation effects, also significantly over-estimate ~~simulates the highest~~ permafrost area (20.86 million km^2). This is inconsistent with previous studies (e.g., Vavrus, 2007; Koven et al., 2013) which concluded that the first-order control on modelled near-surface permafrost distribution is the representation of the air-to-surface soil temperature difference. Table 4 shows that the situation is more complex and that snow insulation simulation is not the dominant factor in a good permafrost extent simulation. When the land surface models having poor snow models are eliminated, the remaining models' simulated permafrost area show little or no relationship with the performance of the snow insulation component, because several other factors such as differences in the treatment of soil organic matter, soil hydrology, surface energy calculations, model soil depth, and vegetation also provide important controls on simulated permafrost distribution (e.g., Marchenko and Etzelmüller, 2013).

6 Summary and conclusions

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

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

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

Snow and its insulation effects are critical for accurately simulating soil temperature and permafrost in high latitudes. The simulated near-surface permafrost area varies greatly across the nine models (from 7.62 to 20.86 million km^2). However, it is hard to find a clear relationship between the performance of the snow insulation in the models and the simulated area of permafrost, because several other factors e.g. related to soil depth and properties and vegetation cover also provide important controls on simulated permafrost distribution.

Acknowledgments. The data will be made available through the National Snow and Ice Data Center (NSIDC; http://nsidc.org); the contact person is Kevin Schaefer (kevin.schaefer@nsidc.org). This study was supported by the Permafrost Carbon Vulnerability Research Coordination Network, which is funded by the U.S. National Science Foundation (NSF). Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government. E.J.B. was supported by the Joint UK DECC / Defra Met Office Hadley Centre Climate Program (GA01101). E.J.B., S.P., P.C. and G.K. were supported by the European Union Seventh Framework Program (FP7/2007-2013) under grant agreement n°282700. T.J.B. was supported by grant 1216037 from the NSF Science, Engineering and Education for Sustainability (SEES) Post-Doctoral Fellowship program. B.D., R.A. and C.D. were supported by the French Agence Nationale de la

Recherche under agreement ANR-10-CEPL-012-03. This research was sponsored by the Integrated approaches and impacts, China Global Change Program (973 Project), National Basic Research Program of China Grant 2015CB953602 and the National Natural Science Foundation of China Grant 40905047.

References

Anderson, E.A.: A point energy and mass balance model of a snow cover, Office of Hydrology, National Weather Service, Silver Spring, Maryland, NOAA Technical Report NWS 19, 1976.

Andreadis, K., Storck, P. and Lettenmaier, D.P.: Modeling snow accumulation and ablation processes in forested environments, Water Resour. Res., 45, W05429, doi:10.1029/2008WR007042, 2009.

Anisimov, O.A., Sherstiukov A.B. Evaluating the effect of environmental factors on permafrost factors in Russia, Earth’s Cryosphere, XX(2), 90-99, 2016.

Avis, C.A.: Simulating the present-day and future distribution of permafrost in the UVic Earth System Climate Model, Dissertation, University of Victoria, Canada, 274pp, 2012.

Bartlett, P.A., MacKay, M.D., Verseghy, D.L.: Modified snow algorithms in the Canadian Land Surface Scheme: model runs and sensitivity analysis at three boreal forest stands, Atmosphere-Ocean, 44, 207–222, 2006.

Best, M.J., and 16 co-authors: The Joint UK Land Environment Simulator (JULES), model description—Part 1: energy and water fluxes, Geosci. Model. Dev., 4, 677–699, doi:10.5194/gmd-4-677-2011, 2011.

Boone, A., and Etchevers, P.: An intercomparison of three snow schemes of varying complexity coupled to the same land-surface model: Local scale evaluation at an Alpine site, J. Hydrometeor., 2, 374-394, 2001.

Brown, J., Ferrians, O.J., Heginbottom, J.A. and Melnikov, S.E.: International Permafrost Association Circum-Arctic Map of Permafrost and Ground Ice Conditions, scale 1:10,000,000, Circum-Pacific Map Series, USGS Circum-Pacific Map Series, Map CP-45, 1997.

Brun, E., David, P., Sudul, M. and Brunot, G.: A numerical model to simulate snow cover stratigraphy for operational avalanche forecasting, J. Glaciol., 38, 13–22, 1992.

Brun, E., Vionnet, V., Boone, A., Decharme, B., Peings, Y., Valette, R., Karbou, F. and Morin, S.: Simulation of northern Eurasian local snow depth, mass and density using a detailed snowpack model and meteorological reanalysis, J. Hydrometeorol., 14, 203–214, doi:10.1175/jhm-d-12-012.1, 2013.

Cook, B.I., Bonan, G.B., Levis, S. and Epstein, H.E.: The thermoinsulation effect of snow cover within a climate model, Clim. Dyn., 31, 107-124, doi:10.1007/s00382-007-0341-y, 2008.

Dai, Y., Zeng, X., Dickinson, R.E., Baker, I., Bonan, G.B., Bosilovich, M.G., Denning, A.S., Dirmeyer, P.A., Houser, P.R., Niu, G., Oleson, K.W., Schlosser, C.A. and Yang, Z.: The Common Land Model (CLM), Bull. Am. Meteorol. Soc., 84, 1013–1023, doi:10.1175/BAMS-84-8-1013, 2003.

Dankers, R., Burke, E.J., and Price, J.: Simulation of permafrost and seasonal thaw depth in the JULES land surface scheme, The Cryosphere, 5, 773-790, doi:10.5194/tc-5-773-2011, 2011.

Decharme, B., Brun, E., Boone, A., Delire, C., Le Moigne, P. and Morin, S.: Impacts of snow and organic soils parameterization on North-Eurasian soil temperature profiles simulated by the ISBA land surface model, The Cryosphere, 10, 853–877, doi: 10.5194/tc-10-853-2016, 2016.

Douville, H., Royer, J.-F. and Mahfouf, J.-F.: A new snow parameterization for the Meteo-France climate model. Part 1: Validation in stand-alone experiments, Clim. Dyn., 12, 21–35, 1995.

Dutra, E., Viterbo, P., Miranda, P.M.A. and Balsamo, G.: Complexity of snow schemes in a climate model and its impact on surface energy and hydrology, J. Hydrometeorol., 13, 521–538, doi:10.1175/jhm-d-11-072.1, 2012.

Dutra, E., Balsamo, G., Viterbo, P., Miranda, P.M.A., Beljaars, A., Schär, C. and Elder, K.: An improved snow scheme for the ECMWF land surface model: description and offline validation, J. Hydrometeorol., 11, 899-916, 2010.

Essery, R., Morin, S., Lejeune, Y. and Ménard, C.B.: A comparison of 1701 snow models using observations from an alpine site, Adv. Water Resour., 55, 131–148, doi:10.1016/j.advwatres.2012.07.013, 2013.

Essery, R.L.H, Rutter, N., Pomeroy, J., Baxter, R., Staehli, M., Gustafsson, D., Barr, A., Bartlett, P. and Elder, K.: SnowMIP2: An evaluation of forest snow process simulations, Bull. Am. Meteorol. Soc., 90, 1120-1135, doi:10.1175/2009BAMS2629.1, 2009.

Ge, Y. and Gong, G.: Land surface insulation response to snow depth variability, J. Geophys. Res., 115, D08107, doi:10.1029/2009JD012798, 2010.

Gerten, D., Schaphoff, S., Haberlandt, U., Lucht, W. and Sitch, S.: Terrestrial vegetation and water balance: Hydrological evaluation of a dynamic global vegetation model, J. Hydrol., 286, 249–270, 2004.

Gisnas, K., Westermann, S., Schuler, T., Litherland, T., Isaksen, K., Boike, J. and Etzelmuller, B.: A statistical approach to represent small-scale variability of permafrost temperatures due to snow cover, The Cryosphere, 8, 2063-2074, doi: 10.5194/tc-8-2063-2014, 2014.

Gouttevin, I., Menegoz, M., Domine, F., Krinner, G., Koven, C.D., Ciais, P., Tarnocai, C. and Boike, J.: How the insulating properties of snow affect soil carbon distribution in the continental pan-Arctic area, J. Geophys. Res., 117, G02020, doi:10.1029/2011JG001916, 2012.

Gubler, S., Endrizzi, S., Gruber, S. and Purves, R.S.: Sensitivities and uncertainties of modeled ground temperatures in mountain environments, Geosci. Model Dev., 6, 1319-1336, doi:10.5194/gmd-6-1319-2013, 2013.

Jafarov, E.E., Nicolsky, D.J., Romanovsky, V.E., Walsh, J.E., Panda, S.K. and Serreze, M.C.: The effect of snow: How to better model ground surface temperatures, Cold Regions Sci. Technol., 102, 63-77, doi:10.1016/j.coldregions.2014.02.007, 2014.

Ji, D., Wang, L., Feng, J., Wu, Q., Cheng, H., et al.: Description and basic evaluation of Beijing Normal University Earth System Model (BNU-ESM) version 1, Geosci. Model Dev., 7, 2039-2064, 2014.

Jordan, R.: A one-dimensional temperature model for a snow cover, technical documentation for SNTHERM.89, Special Report 91-16, U.S. Army Cold Regions Research and Engineering Laboratory, Hanover, N.H, 1991.

Karunaratne, K.C., and Burn, C.R.: Freezing n-factors in discontinuous permafrost terrain, Takhini River, Yukon Territory, Canada Proc. 8th Int. Conf. on Permafrost, Zurich, Eds. M. Phillips, S.M. Springman and L.U.Arenson, pp 519–24, 2003.

Klehmet, K., Geyer, B., and Rockel, B.: A regional climate model hindcast for Siberia: analysis of snow water equivalent, The Cryosphere, 7, 1017-1034, doi:10.5194/tc-7-1017-2013, 2013.

Koven, C.D., Riley, W.J. and Stern, A.: Analysis of Permafrost Thermal Dynamics and Response to Climate Change in the CMIP5 Earth System Models, J. Clim., 26, 1877–1900. doi:10.1175/JCLI-D-12-00228.1, 2013.

Koven, C., Friedlingstein, P., Ciais, P., Khvorostyanov, D., Krinner, G. and Tarnocai, C.: On the formation of high-latitude soil carbon stocks: Effects of cryoturbation and insulation by organic matter in a land surface model, Geophys. Res. Lett., 36, L21501, doi:10.1029/2009GL040150, 2009.

Langer, M., Westermann, S., Heikenfeld, M., Dorn, W., Boike, J.: Satellite-based modeling of permafrost temperatures in a tundra lowland landscape, Remote Sensing of Environment, 135 , pp. 12-24, doi:10.1016/j.rse.2013.03.011, 2013.

Lawrence, D.M., and Slater, A.G.: The contribution of snow condition trends to future ground climate, Clim. Dyn., 34, 969-981, doi:10.1007/s00382-009-0537-4, 2010.

Ling, F., and Zhang, T.: Sensitivity study of tundra snow density on surface energy fluxes and ground thermal regime in Northernmost Alaska, Cold Regions Sci. Technol., 44, 121–130, 2006.

Marchenko, S. and Etzelmüller, B.: Permafrost: Formation and Distribution, Thermal and Mechanical Properties. In: John F. Shroder (ed.) Treatise on Geomorphology, Volume 8, pp. 202-222. San Diego: Academic Press, 2013.

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

Meissner, K.J., Weaver, A.J., Matthews, H.D. and Cox, P.M.: The role of land-surface dynamics in glacial inception: A study with the UVic earth system model, Clim. Dyn., 21, 515-537, 2003.

Morse, P.D., Burn, C.R., and Kokelj, S.V.: Influence of snow on near-surface ground temperatures in upland and alluvial environments of the outer Mackenzie Delta, Northwest Territories, Can. J. Earth Sci., 49, 895–913, doi:10.1139/E2012-012, 2012.

Muskett, R.: Remote sensing, model-derived and ground measurements of snow water equivalent and snow density in Alaska, Int. J. Geosci., 3, 1127-1136, doi:10.4236/ijg.2012.35114, 2012.

Nicolsky, D.J., Romanovsky, V.E., Alexeev, V.A. and Lawrence, D.M.: Improved modelling of permafrost dynamics in a GCM land-surface scheme, Geophys. Res. Lett., 34, L08501, doi:10.1029/2007GL029525, 2007.

Oleson, K.W., Lawrence, D.M., Bonan, G.B., Drewniak, B., Huang, M., Koven, C.D., Levis, S., Li, F., Riley, W.J., Subin, Z.M., Swenson, S.C., Thornton, P.E., Bozbiyik, A., Fisher, R., Kluzek, E., Lamarque, ,J.-F., Lawrence, P.J., Leung, L.R., Lipscomb, W., Muszala, S., Ricciuto, D.M., Sacks, W., Sun, Y., Tang, J., Yang, Z.-L.: Technical description of version 4.5 of the Community Land Model (CLM). NCAR Technical Note NCAR/TN-503+STR, doi: 10.5065/D6RR1W7M, 2013.

PaiMazumder, D., Miller, J., Li, Z., Walsh, J. E., Etringer, A., McCreight, J., Zhang, T., Mölders, N. Evaluation of Community Climate System Model soil temperatures using observations from Russia, Theoretical and Applied Climatology , 94(3),187-213, 2008.

Paquin, J.-P., and Sushama, L.: On the Arctic near-surface permafrost and climate sensitivities to soil and snow model formulations in climate models, Clim. Dyn., 44, 203-228, doi:10.1007/s00382-014-2185-6, 2015.

Park, H., Fedorov, A.N., Zheleznyak, M.N., Konstantinov, P.Y. and Walsh, J.E.: Effect of snow cover on pan-Arctic permafrost thermal regimes, Clim. Dyn., 44, 2873-2895, doi:10.1007/s00382-014-2356-5, 2015.

Park, H., A.B. Sherstiukov, A.N. Fedorov, I.V. Polyakov, and J.E Walsh: An observation-based assessment of the influences of air temperature and snow depth on soil temperature in Russia, Environ. Res. Lett. 9, doi:10.1088/1748-9326/9/6/064026, 2014.

Pavlov, A.V., Malkova, G.V. Small-scale mapping of trends of the contemporary ground temperature changes in the Russian North, Earth's Cryosphere, XIII(4), 32-39, 2009.

Peng, S., Ciais, P., Krinner, G., Wang, T., Gouttevin, I., McGuire, A., Lawrence, D., Burke, E., Chen, X., Decharme, B., Koven, C., MacDougall, A., Rinke, A., Saito, K., Zhang, W., Alkama, R., Bohn, T.J., Delire, C., Hajima, T., Ji, D., Lettenmaier, D.P., Miller, P.A., Moore, J.C., Smith, B. and Sueyoshi, T.: Simulated high-latitude soil thermal dynamics during the past four decades, The Cryosphere, 10, 1–14, doi:10.5194/tc-10-1-2016, 2015.

Rawlins, M., McGuire, A., Kimball, J., Dass, P., Lawrence, D., Burke, E., Chen, X., Delire, C., Koven, C., MacDougall, A., Peng, S., Rinke, A., Saito, K., Zhang, W., Alkama, R., Bohn, T.J., Ciais, P., Decharme, B., Gouttevin, I., Hajima, T., Ji, D., Krinner, G., Lettenmaier, D.P., Miller, P.A., Moore, J.C., Smith, B. and Sueyoshi, T.: Assessment of model estimates of land–atmosphere CO 2 exchange across Northern Eurasia, Biogeosciences, 12, 4385-4405, doi:10.5194/bg-12-4385-2015, 2015.

Riseborough, D.W.: An analytical model of the ground surface temperature under snow cover with soil freezing, 58th Eastern Snow Conference, Ottawa, Ontario, Canada, 2001

Romanovsky, V.E., Sazonova, T.S., Balobaev, V.T., Shender, N.I. and Sergueev, D.O.: Past and recent changes in air and permafrost temperatures in eastern Siberia, Glob. Planet. Change, 56, 399-413, doi:10.1016/j.gloplacha.2006.07.022, 2007.

Saha, S., Rinke, A., Dethloff, K. and Kuhry, P.: Influence of complex land surface scheme on Arctic climate simulations, J. Geophys. Res., 111, D22104, doi:10.1029/2006JD007188, 2006.

Schaefer, K., Zhang, T., Bruhwiler, L. and Barrett, A.P.: Amount and timing of permafrost carbon release in response to climate warming, Tellus B, 63, 165–180, doi:10.1111/j.1600-0889.2011.00527.x, 2011.

Schuur, E.A.G., and 18 coauthors: Vulnerability of permafrost carbon to climate change: Implications for the global carbon cycle, Bioscience, 58, 701-714. doi:10.1641/b580807, 2008.

Sherstiukov, A. Dataset of daily soil temperature up to 320 cm depth based on meteorological stations of Russian Federation, RIHMI-WDC, 176, 224-232, 2012a.

Sherstiukov, A. Statistical quality control of soil temperature dataset, RIHMI-WDC, 176, 224-232, 2012b.

Singh, V.P., Singh, P. and Haritashya, U.: Encyclopedia of snow, ice and glaciers, Springer, 1240p, 2011.

Slater, A.G., Schlosser, C.A. and Desborough, C.E.: The representation of snow in land-surface schemes: Results from PILPS 2(d), J. Hydrometeorol., 2, 7-25, 2001.

Slater, A., and Lawrence, D.: Diagnosing present and future permafrost from climate models, J. Clim., doi:10.1175/JCLI-D-12-00341.1, 2013.

Smith, M.W., and Riseborough, D.W.: Climate and the limits of permafrost: a zonal analysis, Permafrost Periglac. Process., 13, 1-15, doi: 10.1002/ppp.410, 2002.

Sokratov, S.A., and Barry, R.G.: Intraseasonal variation in the thermoinsulation effect of snow cover on soil temperatures and energy balance, J. Geophys. Res., 107, doi:10.1029/2001JD000489, 2002.

Sturm, M., Holmgren, J., König, M. and Morrris, K.: The thermal conductivity of seasonal snow, J. Glaciol., 43 (143), 26–41, 1997.

Sturm, M., and Stuefer, S.: Wind-blown flux rates derived from drifts at arctic snow fences, J. Glaciol., 59 (213), 21-34, 2013.

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

Takala, M., Luojus, K., Pulliainen, J., Derksen, C., Lemmetyinen, J., Karna, J.P., Koskinen, J. and Bojkov, B.: Estimating Northern Hemisphere snow water equivalent for climate research through assimilation of space-borne radiometer data and ground-based measurements, Remote Sens. Environ., 115, 3517-3529, 2011.

Takata, K., Emori, S. and Watanabe, T.: Development of the minimal advanced treatments of surface interaction and runoff, Glob. Planet. Change,38, 209–222, 2003.

Vavrus, S.J.: The role of terrestrial snow cover in the climate system, Clim. Dyn., 20, 73-88, doi:10.1007/s00382-007-0226-0, 2007.

Vionnet, V., Brun, E., Morin, S., Boone, A., Faroux, S., Moigne, P.L., Martin, E. and Willemet, J.-M.: The detailed snowpack scheme Crocus and its implementation in SURFEX v7.2, Geosci. Model Dev., 5, 773–791, 2012.

Vionnet, V., Guyomarch, G., Martin, E., Durand, Y., Bellot, H., Bel, C. and Puglièse, P.: Occurrence of blowing snow events at an alpine site over a 10-year period: observations and modelling, Adv. Water Resour., 55, 53-63, 2013.

Von Storch, H. and Zwiers, F.W.: Statistical Analysis in Climate Research, Cambridge University Press, Cambridge, 484pp., 1999.

Wang, T., Ottle, C., Boone, A., Ciais, P., Brun, E., Morin, S., Krinner, G., Piao, S. and Peng, S.: Evaluation of an improved intermediate complexity snow scheme in the ORCHIDEE land surface model, J. Geophys. Res., 118, 6064-6079, doi:10.1002/jgrd.50395, 2013.

Wania, R., Ross, I. and Prentice, I.C.: Integrating peatlands and permafrost into a dynamic global vegetation model: 2. Evaluation and sensitivity of vegetation and carbon cycle processes, Global Biogeochem. Cycles, 23, GB3015, doi:10.1029/2008GB003413, 2009.

Woo, M., Heron, R., Marsh, P. and Steer, P.: Comparison of weather station snowfall with winter snow accumulation in high arctic basins, Atmosph-Ocean, 21 (3), 312-325, doi:10.1080/07055900.1983.9649171, 1983.

Zhang, T.: Influence of the seasonal snow cover on the ground thermal regime: An overview, Rev. Geophys., 43, RG4002, doi:10.1029/2004RG000157, 2005.

Zhang, T., Osterkamp, T.E. and Stamnes, K.: Influence of the depth hoar layer of the seasonal snow cover on the ground thermal regime, Water Resour. Res., 32, 2075–2086, doi:10.1029/96WR00996, 1996.

Zhong, X., Zhang, T. and Wang, K.: Snow density climatology across the former USSR, The Cryosphere, 8, 785-799, doi:10.5194/tc-8-785-2014, 2013.

Tables

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 ρ
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 ρ
ISBA Boone and Etchevers, 2001	ML	Static (3)	Liquid, Ice, Vapor	Diagnosed from snow temperature, mass, density	compaction [3) a,b]	quadratic equation on ρ, 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 ρ
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 $Wm^{-1}K^{-1}$

[1] ML: Multi-layer, BL: Bulk-layer, I: Implicit; according to Slater et al. (2001)

[2] Not represented means dry snow

[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 metamorphism (crystal breakdown due to wind, thermodynamic stress); Anderson (1976), Jordan (1991), Kojima (1967)

[4] quadratic equation on ρ according to Jordan (1991), Anderson (1976); contribution due to vapor transfer according to Sun et al.(1999)

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

	Snow depth regimes			
	Shallow $d_{snow} \leq 20$ cm		Thick $d_{snow} \geq 45$ cm	
	T_{soil} vs. T_{air} (°C/°C)	R^2	T_{soil} vs. T_{air} (°C/°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

Table 3. Russian-station-location averaged error statistics for snow depth (cm) and temperature difference between 20 cm soil and air temperature (ΔT; K) for winter 1980-2000. For each variable, the maximum available number of observations (n) is used. $Mean^{St,GS}$ and $stdev^{St,GS}$ are the observed mean and interannual variability (standard deviation), while stdev is the standard deviations of each model. Bias is the mean error 'simulation minus observation' and rmse is the root-mean-square error. The statistics for snow depth is given based on both station observation (St) and GlobSnow (GS) data.

	Snow depth (n=579) meanSt= 26.4 cm, mean^{GS}=23.4 cm stdevSt= 9.0 cm, stdev^{GS}= 6.5 cm					ΔT (n=268) meanSt= 11.9 K stdevSt= 2.3 K		
	biasSt	rmseSt	bias^{GS}	rmse^{GS}	stdev	biasSt	rmseSt	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

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

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

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

Figure 2. Conditional probability density functions (PDFs) of ΔT (K), the difference between soil temperature at 20 cm depth and air temperature for (a) different snow depth classes and (b) air temperature regimes, for winter 1980-2000.

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

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

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

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