We thank the reviewers for their careful review and helpful comments. Detailed responses to each comment raised by the reviewers are given in blue as follows.

On behalf of all co-authors,

Shushi PENG

Responses to Anonymous Referee #1

General comments: Peng et al. compares simulated change in soil temperature from 1960 to 2000 with nine process-based land surface models which are forced with historical data. Through a set of additional factorial simulations the effect of different forcing variables on soil temperature is explored as well as the uncertainty contribution from forcing versus model structure (and parameters). Finally, they study present estimates of how much the near surface permafrost area has decreased in this period and its sensitivity to soil temperature. Altogether it is my understanding that this study presents new results which fits with the scope of The Cryosphere. The aim is clearly stated and the manuscript is with few exceptions well written and logically structured. However, the discussion should in my opinion be deeper on some points before the conclusions can be drawn. Also, it seems that the spread in model results is not always reflected in the text and not shown on certain figures. I therefore recommend the paper to be published in The Cryosphere after major revisions.

[Response] Overall, we revised the manuscript following your suggestions. We expanded the discussion section, especially on the uncertainty of Ts trend between models. The details are provided below.

Specific comments:

I would have liked to see a deeper discussion about the differences between the models and to what extent this can be assumed to represent the uncertainties in our understanding of the actual system. This issue is briefly touched upon in the last part of the conclusion, but should in my opinion be expanded, and be discussed before conclusions are drawn.

[Response] We expanded the discussion about uncertainty of modeling Ts including climate forcing uncertainty and 'cold processes' in models as below.

"Meteorological stations are sparse in the cold permafrost regions. For example, there are only 8.8 stations per million km^2 north of 60°N in the CRU TS3.22 gridded air temperature product compared to 41.1 stations per million km^2 between 25°N and 60°N. This results in uncertainty in gridded climate products over Arctic regions, especially for trends of Arctic climate variables (Mitchell and Jones, 2005; Troy and Wood, 2009; Rawlins et al., 2010; Weedon et al., 2011). Troy and Wood (2009) reported 15-20 W m^{-2} of differences in radiative fluxes on seasonal timescales over

northern Eurasia, between six gridded products. Between different gridded observations and reanalysis precipitation products, the magnitude of Arctic precipitation ranges from 410 mm yr^{-1} to 520 mm yr^{-1}, and the trend of Arctic precipitation also has a large spread (Rawlins et al., 2010). These large uncertainties in climate forcing in Arctic undoubtedly can cause large spread of modeled T_s. We found that the FU dominates the total uncertainty of $\dot{T}_s$. This suggests that modelers not only need to improve their models, but also need better climate forcing data (or need to test the effects of different climate input data) when modeling long term changes of T_s in permafrost regions. However, to quantify the SU, simulations using the same agreed upon climate forcing data are highly recommended to further attribute the contribution of each process in the soil thermal dynamics of models such as organic carbon insulation effects, snow insulation effects, latent heat formation and emission, soil conductivity and surface properties (see Lawrence and Slater, 2008; Koven et al., 2009; Bonfils et al., 2012; Gouttevin et al., 2012). In addition, important processes in permafrost regions such as ice content (e.g. ice wedge) in permafrost and thermokarst lakes etc. should be developed in land surface models to improve the prediction of future permafrost feedbacks (e.g. van Huissteden et al., 2013; Lee et al., 2014).

The discussion of the different drivers of trend in soil temperature (section 3.3) does not reflect the importance of the different variables. In particular the importance of Ta and the spread in sensitivity to this variable should be discussed more. For LWDR it should be more clearly stated that this is based on only two models, which does not have a representative sensitivity to for instance Ta. This should also be reflected in the conclusions about which variables are most important in driving Ts.

[Response] We clarified that the contribution of LWDR on the trend of Ts is based on only two models (JULES and ORCHIDEE) in the Conclusions section. “Note that the relative contribution of LWDR is based on only two models in this study, and this needs further investigation.” was added into the revised version.

We also added the discussion about Ta and Ts as below.

“This indicates the importance of increasing Ta on the trend of Ts, and is consistent with observations. Based on 30 climate stations observations in Canada during the period 1958-2008, Ts at 10 cm significantly and positively correlates with Ta at most sites (>90%) in spring, but at fewer sites (<30%) in winter (Qian et al., 2011). For winter Ts, the winter snow depth was found to have significant and positive correlation with Ts in shallow soil layers (e.g. Zhang et al., 2001; Qian et al., 2011). Recent increases in Ta also explain the trend of Ts at 1.6m measured at Churapcha metrological station (N62.02, E132.36), and at 5 m measured in a borehole at Iqaluit (N63.47, W68.48) in Canada (Smith et al., 2005; Romanovsky et al., 2007). To some extent, the trend of Ta is a good indicator for the trend of deep permafrost ground temperature with some time lag (Romanovsky et al., 2007). For the modeled Ts in land surface models, the effects of Ta on Ts depend on surface energy balance and ground heat flux into soil; i.e. the extent of coupled Ta on Ts relates to the surface properties such as snow, organic soil horizons and rougness etc. in the models. The

different relative contributions of the trend of Ta to the trend of Ts in these models maybe mainly result from the different model parameterization and structures, as the trends of Ta (~0.03 °C yr-1) in the climate forcing do not have a large spread (Figure 7)."

The statement that in UW-VIC "nearly 100 %" of the trend in soil temperature can be explained with Ta (P2310 line 17-19) does not reflect the results presented in table 3.

[Response] We corrected it.

Figure 4 and 6 seems to show a too narrow range of values, excluding the extremes. This is clearly also the case in figure 5, but here I think it is sufficient to include a note about this in the figure caption.

[Response] We clarified this in the figure captions for these two figures.

P2306 line 25: The term "some of their non-linear interactions" should be explained better.

[Response] We used the difference between two simulations with and without the trend of one driver to identify the effects of such driver on Ts. For example, R01 includes the effects of all drivers, and R02 includes the effects of all drivers except Ta. In R02 with detrended Ta, the interactions between increasing Ta and CO_2 or other drivers were also excluded. Thus, the difference between R01 and R02 shows the effects of Ta and some interactions between Ta and other drivers. We removed "non-linear" here.

Technical comments:

P2306 line 6: Either "hereafter" or "in the following" should be sufficient.

[Response] We corrected it.

P2306: The sentence starting with "To separate . . ." is not very fluent. Dropping the ":" and writing the variables in a parenthesis would make it clearer.

[Response] We corrected it.

P2307: I found the sentence staring with "if the maximum soil depth. . ." hard to understand. Please clarify.

[Response] We revised this sentence as "Modeled Ts at depths deeper than 300 cm (six models modeled Ts deeper than 300 cm, except CoLM, JULES and LPJ-GUESS) was not extrapolated (the maximum soil depth of each model is shown in Table 1).".

P2307 line 7: differenced => differences

[Response] We corrected it.

P2307 line 27: Wei et al. 2013 not found in reference list. Is it Wei et al. 2014?

[Response] We corrected it.

P2308 line 3: which "seven models"?

[Response] We clarified here "seven models" out of the nine models except LPJ-GUESS and UVic because LWDR was not used by these two models.

P2308 line 22: Unclear sentence.

[Response] We revised this sentence as "The trend of Ts at different soil depths is shown in Figure 5 for each model.".

P2309 line 8-9/Figure 5: why is CLM only showed down to 35 m?

[Response] In the previous version, we only plotted the depth of soil node for each soil layer. In CLM, the node of deepest soil layer in CLM is 38.2 m, and the bottom of deepest soil layer is 45.1 m. We re-plotted Figure 5, and extended the soil depth to the bottom for each soil layer.

P2309 line 12-13: Please elaborate on the negative trend seen in UW-VIC below 2.5m.

[Response] The negative trend in UW-VIC below 2.5 m is about -0.035 °C yr^{-1}. We clarified this in the revised version.

P2311 line 12: Please rephrase. Does not all show trend since 1960?

[Response] We corrected it.

P2311 line 14: is => are

[Response] We corrected it.

P2311: Please rephrase last sentence.

[Response] We revised the last sentence as "The effects of snowfall trends and growing season precipitation trends may oppose each other as mentioned above. These two contrasting effects cannot be separated in this analysis, because models did not run simulations with seasonally detrended precipitation. But the different effects of seasonal precipitation on Ts should be studied in the future.".

P2315: Please rephrase first sentence.

[Response] We revised this sentence as "The total boreal NSPA during 1960-2000 estimated by the nine models ranges from 6.8 million km2 (CoLM) to 19.7 million km2 (ORCHIDEE).".

P2315 line 7: Lawrence and Slanter et al. 2013 => Slanter and Lawrence 2013 ?

[Response] We corrected it.

References

Bonfils, C. J. W., Phillips, T. J., Lawrence, D. M., Cameron-Smith, P., Riley, W. J., and Subin, Z. M.: On the influence of shrub height and expansion on northern high latitude climate, Environmental Research Letters, 7, 015503, doi:10.1088/1748-9326/7/1/015503, 2012.

Gouttevin, I., Krinner, G., Ciais, P., Polcher, J., and Legout, C.: Multi-scale validation of a new soil freezing scheme for a land-surface model with physically-based hydrology, The Cryosphere, 6, 407-430, doi:10.5194/tc-6-407-2012, 2012.

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, Geophysical Research Letters, 36, L21501, doi:10.1029/2009gl040150, 2009.

Lawrence, D., and Slater, A.: Incorporating organic soil into a global climate model, Clim Dyn, 30, 145-160, doi:10.1007/s00382-007-0278-1, 2008.

Lee, H., Swenson, S. C., Slater, A. G., and Lawrence, D. M.: Effects of excess ground ice on projections of permafrost in a warming climate, Environmental Research Letters, 9, 124006, doi:10.1088/1748-9326/9/12/124006, 2014.

Mitchell, T. D., and Jones, P. D.: An improved method of constructing a database of monthly climate observations and associated high-resolution grids, International Journal of Climatology, 25, 693-712, doi:10.1002/joc.1181, 2005.

Qian, B., Gregorich, E. G., Gameda, S., Hopkins, D. W., and Wang, X. L.: Observed soil temperature trends associated with climate change in Canada, Journal of Geophysical Research: Atmospheres, 116, D02106, doi:10.1029/2010jd015012, 2011.

Rawlins, M. A., Steele, M., Holland, M. M., Adam, J. C., Cherry, J. E., Francis, J. A., Groisman, P. Y., Hinzman, L. D., Huntington, T. G., Kane, D. L., Kimball, J. S., Kwok, R., Lammers, R. B., Lee, C. M., Lettenmaier, D. P., McDonald, K. C., Podest, E., Pundsack, J. W., Rudels, B., Serreze, M. C., Shiklomanov, A., Skagseth, O., Troy, T. J., Vorosmarty, C. J., Wensnahan, M., Wood, E. F., Woodgate, R., Yang, D. Q.,

Zhang, K., and Zhang, T. J.: Analysis of the Arctic System for Freshwater Cycle Intensification: Observations and Expectations, Journal of Climate, 23, 5715-5737, doi:10.1175/2010jcli3421.1, 2010.

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, Global and Planetary Change, 56, 399-413, doi:10.1016/j.gloplacha.2006.07.022, 2007.

Smith, S. L., Burgess, M. M., Riseborough, D., and Mark Nixon, F.: Recent trends from Canadian permafrost thermal monitoring network sites, Permafrost and Periglacial Processes, 16, 19-30, doi:10.1002/ppp.511, 2005.

Stevens, M. B., Smerdon, J. E., González-Rouco, J. F., Stieglitz, M., and Beltrami, H.: Effects of bottom boundary placement on subsurface heat storage: Implications for climate model simulations, Geophysical Research Letters, 34, L02702, doi:10.1029/2006gl028546, 2007.

Troy, T. J., and Wood, E. F.: Comparison and evaluation of gridded radiation products across northern Eurasia, Environmental Research Letters, 4, 045008, doi:10.1088/1748-9326/4/4/045008, 2009.

van Huissteden, J., Berrittella, C., Parmentier, F. J. W., Mi, Y., Maximov, T. C., and Dolman, A. J.: Methane emissions from permafrost thaw lakes limited by lake drainage, Nature Clim. Change, 1, 119-123, doi:dx.doi.org/10.1038/nclimate1101, 2011.

Weedon, G. P., Balsamo, G., Bellouin, N., Gomes, S., Best, M. J., and Viterbo, P.: The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data, Water Resources Research, 50, 7505-7514, doi:10.1002/2014wr015638, 2014.

Responses to Anonymous Referee #2

The manuscript presents a comparison of soil temperature trends simulated by nine numerical models. Texts are well organized and clearly written, and tables and diagrams are effective in showing the variability among simulated soil temperatures. It is an interesting exercise to demonstrate the large uncertainty in soil temperature simulation in the permafrost region. However, I am not sure what scientific advances we gain from this exercise. If I am not mistaken, the nine models are driven by different climatic forcings, and have vastly different structures and algorithms (Table 1). Therefore, the variability in model results is due to the variability in both forcing and model algorithms. As a result, the reader is left wondering what the results of this exercise really mean. For example, ColM and ISBA show completely different patterns of soil temperature trends (Figure 4). Is this caused by model algorithms or climate forcing?

As someone who has done much work on the comparison between permafrost model simulations with field observations, I am keenly aware of model sensitivity to subtle changes in surface variables (e.g. vegetation parameterization), subsurface variables (e.g. soil moisture), and boundary conditions (e.g. magnitude of geothermal flux). Permafrost models are also sensitive to the initial condition, as well as the thickness of the model domain. In order of the reader to understand the meaning and implication of model results, it would have been much more meaningful to conduct the model comparison exercises using a common set of forcing variables. For these reason, I cannot recommend publication of this manuscript in its present form. I suggest that the authors re-design and conduct new model comparison exercises, or present more meaning explanation for differences among the present simulation results and discuss how what causes these differences.

[Response] We agree that it is easier to understand the comparison exercise using the same climate forcing. However, as we shown in this manuscript, large uncertainty of climate forcing in high latitude regions also can cause the spread of modeled Ts, besides the spread of modeled Ts caused by model parameters and structure. Previous studies (e.g. Troy and Wood, 2009; Rawlins et al., 2010) shown that large spread of climate variables over Arctic regions. Troy and Wood (2009) reported 15-20 W m^{-2} of differences in radiative fluxes on seasonal timescales over northern Eurasia, between six gridded products. Between different gridded observations and reanalysis precipitation products, the magnitude of Arctic precipitation ranges from 410 mm yr^{-1} to 520 mm yr^{-1}, and the trend of Arctic precipitation also has a large spread (Rawlins et al., 2010). We do not know whether the spread of modeled trend of Ts among applications of land surface models applied to the permafrost region associated with uncertainty of climate forcing is larger than that caused by uncertainty associated different parameterization and processes in models. One protocol model inter-comparison exercise to quantify the relative uncertainty between these two sources would be to apply several climate forcing data sets with each model and then partition the variability among the climate drivers and the

models. Alternatively, another approach is allow the models to conduct simulations with their preferred historical climate data, and then use the methodology in this study to partition results key aspects of variability among climate drivers (as inferred from ancillary simulations with detrended climatic components) and models. For practical reasons, the simulation protocol in the model-intercomparison project of Permafrost Carbon Network (PCN, www.permafrostcarbon.org) allow each model group to choose climate forcing and other boundary files (such as vegetation map, soil map etc.) to fully quantify the uncertainty of model results among uncertainties from climate forcing (and other forcing files) and model parameterization and structure. We indicate in the revised manuscript that future model inter-comparisons on permafrost dynamics should investigate the full uncertainty by conducting simulations for multiple climate forcing data sets.

We clarified that the main objective of this study is to distinguish the uncertainty caused by assigned parameter values and model structure from the uncertainty attributable to uncertain climate forcing data. We accomplished this by an approach of separating functional uncertainty from structural uncertainty. For the explanation of difference between models, we discussed the possible causes. In Figure 7-9, we show that the climate forcing could explain most of the differences among models. We also added further discussion about the differences in the revised version.

In addition to the fundamental comments above, the following is specific comments.

Page 2305, Line 11-16. What are the scientific objectives of this work? The objectives (1)-(3) cannot be meaningfully achieved, if model simulation results are strongly dependent on model algorithms and structure.

[Response] As mentioned in the response to fundamental comments above, we try to use the information from this model inter-comparison as much as possible to show the range of modeled Ts trend, their possible drivers (Table 3) and observation-constraint Ts trend and permafrost retreat rate (Figure 8 and 9).

Page 2305, Line 19-23. In addition to surface forcing, the forcing from the bottom boundary of models needs to be explained clearly. Energy input in the form of geothermal flux has strong effects on soil temperature.

[Response] We agree that bottom geothermal flux has effects on simulated soil temperature. We added the number of geothermal flux for each model in Table 1.

Page 2305, Line 23-24. If I have understood correctly, three out of nine models do not consider the "effects of water in soil on phase change". Does that mean these three models do not simulate the freezing and thawing of soil water? Since permafrost is the phenomenon of pore water freezing and thawing, I am not sure if these three models are even suitable for the purpose of this exercise. Clear justification is needed for the inclusion of these models.

[Response] Most models used in this study are the land surface component of Earth System Models participating CMIP5 project. The original idea was to compare these models and to show possible guidelines for further permafrost simulations in CMIP6 or other future projects. Actually, two of the nine models used in this study (CoLM and MIROC-ESM) do not represent the effects of water phase change processes (freezing and thawing) in frozen soil. But these two models represent other important processes/issues such as maximum 3 and 5 snow layers in MIROC-ESM and CoLM, respectively. We think that it is important to include all the nine models to show their results for comparison.

Page 2306, Line 8-10. Why were different forcing data sets used for different models? Clear justification is needed in relation to the scientific objectives of the study.

[Response] Please see the response to general comments. We added two sentences in Introduction and Methods section as below.

"both uncertainties from climate forcing (and other forcing files) and from model parameterization and structure"

"The different modeling groups in this study used different forcing datasets for climate and other model boundary conditions (Table 1), which collectively represent both uncertainty from climate forcing (and other forcing files) and from model parameterization and structure in simulating soil thermal dynamics across the permafrost region."

Page 2307, Line 18-22. The thermal condition of top 3m is strongly dependent on the presence of absence of permafrost in the underlying zone. In some regions permafrost is more than 10-20m thick. Table 1 indicates that some of the models are not sufficiently deep to represent the effects of underlying permafrost.

[Response] Yes, we pointed this out in the text and suggested that models should have deeper soil depth to simulate permafrost as below.

"…This suggests that much deeper maximum soil depth than the currently prescribed maximum soil depths (Table 1) are needed for some models to calculate the heat flux into the entire soil profile (Stevens et al., 2007). CoLM, JULES and LPJ-GUESS have too shallow maximum soil depth for the calculation of permafrost soil temperature trends over the last four decades, which makes these models even less realistic for deeper Ts projections over the next century (e.g. Alexeev et al., 2007)."

References

Alexeev, V. A., Nicolsky, D. J., Romanovsky, V. E., and Lawrence, D. M.: An evaluation of deep soil configurations in the CLM3 for improved representation of permafrost, Geophysical Research Letters, 34, L09502, doi:10.1029/2007gl029536, 2007.

Rawlins, M. A., Steele, M., Holland, M. M., Adam, J. C., Cherry, J. E., Francis, J. A.,

Groisman, P. Y., Hinzman, L. D., Huntington, T. G., Kane, D. L., Kimball, J. S., Kwok, R., Lammers, R. B., Lee, C. M., Lettenmaier, D. P., McDonald, K. C., Podest, E., Pundsack, J. W., Rudels, B., Serreze, M. C., Shiklomanov, A., Skagseth, O., Troy, T. J., Vorosmarty, C. J., Wensnahan, M., Wood, E. F., Woodgate, R., Yang, D. Q., Zhang, K., and Zhang, T. J.: Analysis of the Arctic System for Freshwater Cycle Intensification: Observations and Expectations, Journal of Climate, 23, 5715-5737, doi:10.1175/2010jcli3421.1, 2010.

Stevens, M. B., Smerdon, J. E., González-Rouco, J. F., Stieglitz, M., and Beltrami, H.: Effects of bottom boundary placement on subsurface heat storage: Implications for climate model simulations, Geophysical Research Letters, 34, L02702, doi:10.1029/2006gl028546, 2007.

Troy, T. J., and Wood, E. F.: Comparison and evaluation of gridded radiation products across northern Eurasia, Environmental Research Letters, 4, 045008, doi:10.1088/1748-9326/4/4/045008, 2009.

Simulated high-latitude soil thermal dynamics during the past four decades

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

[1]UJF – Grenoble 1/CNRS, Laboratoire de Glaciologie et Géophysique de l'Environnement (LGGE), 38041 Grenoble, France

[2]Laboratoire des Sciences du Climat et de l'Environnement (LSCE), CEA-CNRS-UVSQ, 91191 Gif-sur-Yvette, France

[3]Irstea, UR HHLY, 5 rue de la Doua, CS 70077, 69626 Villeurbanne Cedex, France

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

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

[6]Met Office Hadley Centre, FitzRoy Road, Exeter, EX1 3PB, UK

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

[8]CNRM-GAME, Unitémixte de recherche CNRS/Meteo-France (UMR 3589), 42 avCoriolis, 31057 Toulouse cedex, France

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

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

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

[12]Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Potsdam, Germany

[13]Research Institute for Global Change, Japan Agency for Marine-Earth Science and Technology, Yokohama, Kanagawa, Japan

[14]Department of Physical Geography and Ecosystem Science, Lund University, Sölvegatan 12, SE-223 62 Lund, Sweden

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

[16]National Institute of Polar Research, Tachikawa, Tokyo, Japan

Revised Manuscript for *The Cryosphere*

Jul 31, 2015

392 words (abstract) + 5360 words (text) + 3 table + 9 figures + 60 references

[*]**Corresponding author:**
Shushi Peng
LGGE/LSCE, France
Tel: +0033-1 69 08 53 03
Email: Shushi.Peng@lsce.ipsl.fr

Abstract

Soil temperature (T_s) change is a key indicator of the dynamics of permafrost. On seasonal and inter-annual time scales, the variability of T_s determines the active layer depth, which regulates hydrological soil properties and biogeochemical processes. On the multi-decadal scale, increasing T_s not only drives permafrost thaw/retreat, but can also trigger and accelerate the decomposition of soil organic carbon. The magnitude of permafrost carbon feedbacks is thus closely linked to the rate of change of soil thermal regimes. In this study, we used nine process-based ecosystem models with permafrost processes, all forced by different observation-based climate forcing during the period 1960-2000, to characterize the warming rate of T_s in permafrost regions. There is a large spread of T_s trends at 20 cm depth across the models, with trend values ranging from 0.010 ± 0.003 °C yr^{-1} to 0.031 ± 0.005 °C yr^{-1}. Most models show smaller increase in T_s with increasing depth. Air temperature (T_a) and longwave downward radiation (LWDR) are the main drivers of T_s trends, but their relative contributions differ amongst the models. Different trends of LWDR used in the forcing of models can explain 61% of their differences in T_s trends, while trends of T_a only explain 5% of the differences in T_s trends. Uncertain climate forcing contributes a larger uncertainty in T_s trends (0.021 ± 0.008 °C yr^{-1}, mean ± standard deviation) than the uncertainty of model structure (0.012 ± 0.001 °C yr^{-1}), diagnosed from the range of response between different models, normalized to the same forcing. In addition, the loss rate of near-surface permafrost area, defined as total area where the maximum seasonal active layer thickness (ALT) is less than 3 m loss rate is found to be significantly correlated with the magnitude of the trends of T_s at 1 m depth across the models (R=-0.85, P=0.003), but not with the initial total near-surface permafrost area (R=-0.30, P=0.438). The sensitivity of the total boreal near-surface permafrost area to T_s at 1 m, is estimated to be of -2.80±0.67 million km^2 $°C^{-1}$. Finally, by using two long-term LWDR datasets and relationships between trends of LWDR and T_s across

models, we infer an observation-constrained total boreal near-surface permafrost area decrease comprised between 39 ± 14 $\times 10^3$ and 75 ± 14 $\times 10^3$ km^2 yr^{-1} from 1960 to 2000. This corresponds to 9% - 18% degradation of the current permafrost area.

Key words: soil temperature, permafrost, downward longwave radiation, climate warming, land surface model

1 Introduction

Arctic permafrost regions store ~1300 Pg carbon (C) in the soil, including ~1100 Pg C in frozen soil and deposits (Hugelius et al., 2014). Decomposition of these large carbon pools in response to permafrost thawing from projected future warming is expected to be a positive feedback on climate warming through increased emissions of CO_2 and CH_4 (Khvorostyanov et al., 2008; Schuur et al., 2008; McGuire et al., 2009; Koven et al., 2011; Schaefer et al., 2011). The magnitude of permafrost soil carbon feedbacks on climate depends on the rate of soil carbon decomposition, which is related to permafrost thaw, soil water and temperature changes, the quantity and quality of soil carbon available as a substrate for decomposition, and the concentration of oxygen in the soil, which determines CH_4 vs. CO_2 production ratio (Schuur et al., 2008; Schädel et al., 2014; Elberling et al., 2013). Both the rate of permafrost thaw and the rate of soil carbon decomposition are closely related to soil thermal dynamics (Koven et al., 2011; Schädel et al., 2014; Elberling et al., 2013).

Measurements of active layer depth across circumpolar regions and borehole temperature profiles indicate that active layer thickness on top of boreal permafrost has been increasing in response to the warming that occurred during recent decades in North America, Northern Europe and Russia (e.g. Zhang et al., 2001; Qian et al., 2011; Smith et al., 2005, 2010; Romanovsky et al., 2007, 2010). For example, the borehole record of Alert in Canada (82°30′N, 62°25′W) shows that soil temperature at 9 m, 15 m and 24 m increased at rates of 0.6 °C decade^{-1}, 0.4 °C decade^{-1} and 0.2 °C decade^{-1} from 1978 to 2007, respectively (Smith et al., 2012). These observations provide long-term local monitoring of changes in active layer thickness and soil temperature, but the measurement sites are sparse, and their temporal sampling frequency is often low (Romanovsky et al., 2010). Because site measurements cannot document permafrost area loss on a large scale, land surface models including ‘cold

processes', such as soil freeze-thaw and the thermal and radiative properties of snow, are important tools for quantifying the rate of permafrost degradation on a large scale, and its evolution in response to climate change scenarios.

However, there are large uncertainties in soil thermal dynamics in land surface models (e.g. Koven et al., 2013), and these uncertainties also impact predictions of carbon-cycle feedbacks on climate. To quantify and reduce the uncertainty of modeled soil temperature (T_s), the driving factors of T_s trends need to be investigated. Besides the uncertainty in model parameterization and structure, the gridded climate forcing for offline land surface models over high latitude regions have large uncertainty (e.g. Troy and Wood, 2009; Rawlins et al., 2010). It is also important to distinguish the uncertainty caused by assigned parameter values and model structure from the uncertainty attributable to uncertain climate forcing data.

In this study, nine process-based models that participated in the Permafrost Carbon Network (PCN, www.permafrostcarbon.org) were used (1) to compare trends of simulated T_s at different depths over the boreal permafrost regions during the past four decades and to assess the uncertainty of modeled T_s trends ~~across models~~; (2) to identify which factors drive trends of permafrost T_s; and (3) to quantify the sensitivity of changes in near-surface permafrost area to warming.

2 Methods

2.1 Models and simulations

The nine land surface models that were used for simulating T_s in permafrost regions organized by Permafrost Carbon Network (PCN, www.permafrostcarbon.org) are listed in Table 1. All the models used finite difference solution of heat equation with phase change to

simulate T_s, but models have different soil depths, snow parameterizations, and soil thermal conductivities~~y~~ (Table 1). Three models (CLM, ISBA, UW-VIC) explicitly considered organic soil insulation and ~~six~~ seven models explicitly considered the effect of water in soil on phase change. All models explicitly considered snow insulation but with different snow layers. The soil thermal conductivity depends on soil moisture in all models. More details can be found in Rawlins et al. (2015) and McGuire et al. (in preparation). We defined the Northern Hemisphere permafrost spatial domain as the definition in Figure 1, and the analysis considers three permafrost regions, Boreal North America (BONA), Boreal Europe (BOEU), Boreal Asia (BOAS) (Figure 1; Brown et al., 1998). We did not include the Tibetan plateau because not all the models covered this region. Hereafter, the term "boreal regions" is used ~~in the following~~ for the sum of the three sub-regions BONA, BOEU and BOAS in Figure 1.

Following the simulation protocol of the PCN project, nine land surface models performed historical simulations from 1960 to 2000, using different forcing data sets (Table 1). The different modeling groups in this study used different forcing datasets for climate and other model boundary conditions (Table 1), which collectively represent both uncertainty from climate forcing (and other forcing files) and from model parameterization and structure in simulating soil thermal dynamics across the permafrost region.~~Modelers used different forcing datasets for climate and other model boundary conditions (Table 1)~~. Climate forcing data chosen by each group are presented in Table 1, and the differences in the trend of T_a, precipitation, and radiative forcing are summarized in Figure S1 and S2. How differences between these drivers are related to differences of the modeled T_s is discussed in the Results and Discussion section.

To separate the contributions of the trends of four forcing variables~~:~~ (T_a, atmospheric

CO_2, precipitation, and LWDR) on permafrost thermal dynamics and carbon stocks, six out of the nine models conducted factorial simulations (R01-R04). The ORCHIDEE and JULES performed two additional simulations (R05-R06) to isolate the contribution of LWDR on T_s trends (Table 2 and 3). In the reference simulation R01, all drivers varied ~~vary~~ at the same time. In R02 T_a ~~is~~ was detrended; in R03 atmospheric CO_2 was ~~is~~ set constant to the observed 1960 level of 316 ppmv; In R04 both T_a and precipitation ~~are~~ were detrended; in R05 T_a and LWDR ~~are~~ were detrended; in R06 T_a, precipitation and LWDR were~~are~~ detrended. Differences between two simulations were~~are~~ used to separate the controlling effect of each driver on T_s, and some of their ~~non-linear~~ interactions.

2.2 Analysis

Modeled monthly T_s at 5, 20, 50, 100, 200 and 300 cm depths in every grid cell of each model were calculated by linear interpolation of T_s between the central depths of two adjacent layers. ~~If the maximum soil depth of a model is larger than 300 cm (which is the case for seven models, except CoLM and JULES), then~~ Modeled T_s at depths deeper than 300 cm (six models modeled Ts deeper than 300 cm, except CoLM, JULES and LPJ-GUESS) was not extrapolated ~~to layers deeper than 300 cm~~ (the maximum soil depth of each model is shown in Table 1). For each boreal sub-regions BONA, BOEU, BOAS (Figure 1) T_s was first averaged over all grid cells and the trend of regional mean T_s (denoted $\dot{T}_s$) was calculated from a linear regression. The statistical significance of $\dot{T}_s$ is evaluated by a *t-test*.

To estimate the uncertainty of $\dot{T}_s$ caused by differences~~d~~ in the trend of each climate input variable, we regressed $\dot{T}_s$ against the trends of T_a, precipitation and short-wave downward radiation (SWDR) and LWDR, respectively, using the output of R01. The

uncertainty of $\dot{T}_s$ attributed to each forcing variable was~~is~~ defined as the resulting range of $\dot{T}_s$ associated to different trends in each forcing variable in the models. To achieve this aim, we regressed $\dot{T}_s$ against forcing variable across the models, and the uncertainty of $\dot{T}_s$ resulting from uncertain forcing data was calculated as the range of $\dot{T}_s$ from the maximum and minimum values of forcing data in the regression equation. Then we define the $\dot{T}_s$ uncertainty attributed to model structure, which reflects the differences in model parameterizations and parameter values, as the uncertainty of $\dot{T}_s$ assuming all models were using the same climate forcing data.

Here, we defined near-surface permafrost as in previous studies (e.g. Schneider von Deimling et al., 2012): near-surface permafrost is defined as where the maximum seasonal thaw depth (i.e., the active layer thickness, ALT) is less than 3 m. The total near-surface permafrost area (NSPA) is the sum of the areas of grid cells that fulfill this condition.

We used monthly LWDR data from CRUNCEP v5.2 (http://dods.extra.cea.fr/data/p529viov/cruncep) and WATCH (Weedon et al., 2011) with a spatial resolution of 0.5º by 0.5º during the period 1960-2000 to derive the trend of LWDR. The CRUNCEP LWDR dataset was~~is~~ derived from CRU TS3.21 and NCEP reanalysis meteorology, and ancillary data sets (e.g. Wei et al., ~~2013~~2014). The WATCH LWDR dataset was~~is~~ derived from ERA-40 reanalysis (Weedon et al., 2011). Because there is no long-term large scale LWDR observation product available, we did an experiment using LWDR from CRUNCEP and WATCH data to estimate the loss of permafrost area during the period 1960-2000 by an empirical relationship ~~found~~ between the loss of permafrost area and LWDR

trends across the seven models out of the nine models (except LPJ-GUESS and UVic because LWDR was not used by these two models) (see section 3.4 below).

3 Results and Discussion

3.1 Trend in upper-layer soil temperature over boreal regions

The simulated values of $\dot{T}_s$ at 20 cm depth averaged over boreal regions range from 0.010 ± 0.003 °C yr^{-1} (CoLM) to 0.031 ± 0.005 °C yr^{-1} (UVic) during the period 1960-2000 (Figure 2). Figure 3 shows $\dot{T}_s$ at 20 cm for BONA, BOEU and BOAS regions. Six out of the nine models show the largest $\dot{T}_s$ at 20 cm in BOAS, followed by BONA and ~~the smallest $\dot{T}_s$ at 20 cm in~~ BOEU. The other three models (CoLM, JULES and UW-VIC) show the smallest $\dot{T}_s$ at 20 cm in BOAS. Among the six models with smaller $\dot{T}_s$ at 20 cm in BOEU, we found that $\dot{T}_s$ at 20 cm in BOEU is significantly lower than in BOAS and in BONA ($P<0.001$, two sample *t-test*). This is also shown in the spatial distribution of $\dot{T}_s$ at 20 cm (Figure 4). For example, in northern Siberia, T_s at 20 cm increased by more than 0.02 °C yr^{-1} in five out of the nine models (ISBA, LPJ-GUESS, MICRO-ESM, ORCHIDEE and UVic) but decreased in two models (CoLM and JULES). All models show an increase of T_s at 20 cm in northern BONA, but this increase is of different magnitude between models (Figure 4). Six models show significant $\dot{T}_s$ at 20 cm over northern and western Siberia, but all models show non-significant $\dot{T}_s$ at 20 cm over northern BOEU (Figure 4).

3.2 Attenuation of the trend in soil temperature with soil depth

The trend of T_s at different soil depths ~~in soil layers over 0 to 3 m and soil depth~~ is shown in Figure 5 for each model. Based on ground soil temperature observation, annual T_s at

1.6 m increased by 0.02-0.03 °C yr^{-1} from 1960s to 2000s in Russia (Park et al., 2014). The simulated trends of T_s at 1.6 m over BOAS in most models are within this range (Figure S3). Two models (CoLM and JULES) show vertically quasi-uniform $\dot{T}_s$ over the upper 3 m of soil, probably because of too quick soil thermal equilibrium in these two models. The seven other models show decreasing values of $\dot{T}_s$ with increasing soil depth, but the vertical gradient of $\dot{T}_s$ varies among them (Figure 5a). UW-VIC has the largest negative vertical gradient of $\dot{T}_s$ (-0.0052 ± 0.0001 °C yr^{-1} m^{-1}), followed by ISBA, MICRO-ESM, ORCHIDEE and UVic (~-0.0030 ± 0.0003 °C yr^{-1} m^{-1}) and by near-zero vertical gradient of $\dot{T}_s$ in CLM (-0.0009 ± 0.0003 °C yr^{-1} m^{-1}) and in LPJ-GUESS (-0.0014 ± 0.0000 °C yr^{-1} m^{-1}).

Figure 5b shows the trend of T_s in all soil layers over boreal regions. CLM and UVic show an increase of T_s even at depths deeper than 40m, but T_s exhibited no changes deeper than 22m in ORCHIDEE (Figure 5b). T_s increased in the deepest layer of ISBA (12m) and MIROC-ESM (14m), and the depth at which T_s exhibited no changes could not be deduced from these two models. UW-VIC shows a negative trend of T_s (i.e. cooling) at depths deeper than 2.5m. The trends of T_s ~~in all soil layers~~ over BONA, BOEU and BOAS regions decrease in magnitude with increasing soil depth, but show different vertical gradients. In Figure S3, vertical gradient of $\dot{T}_s$ ~~are~~ is shown to be larger in BONA and BOAS than that in BOEU for most models. Figure 6 shows the spatial distribution of the difference in $\dot{T}_s$ at depths between ~~20 c~~0.2 m and 3 m. $\dot{T}_s$ at 0.2m is larger than that at 3m over most regions in BONA, BOEU and BOAS in seven out of the nine models, except~~ions being~~ JULES and CoLM. Generally, borehole~~s~~ records show that mean annual soil temperature at depths between 10 m and 30 m

have increased during the last three decades over the circumpolar northern permafrost regions (Osterkamp, 2003; Romanovsky et al., 2010; Smith et al., 2005, 2012; Vaughan et al., 2013). In Alaska, T_s at 20 m from boreholes ~~even~~ increased by ~1 °C between the early 1980s and 2001 (Osterkamp, 2003). The observed value of $\dot{T}_s$ at one of Alert (BH3) boreholes is of ~0.04 °C yr^{-1} at ~2.5 m depth and nearly zero at ~27 m depth during the period 1979-2004 (see Figure 9 in Smith et al., 2012). Some boreholes (BH1 and BH2) at Alert however still indicated a small warming during the period 1979-2008 (Smith et al., 2012) at 37m. This suggests that much deeper maximum soil depth than the currently prescribed maximum soil depths (Table 1) are needed for some models to calculate the heat flux into the entire soil profile (Stevens et al., 2007). CoLM, JULES and LPJ-GUESS have too shallow maximum soil depth for the calculation of permafrost soil temperature trends over the last four decades, which makes these models even less realistic for deeper T_s projections over the next century (e.g. Alexeev et al., 2007). Compared to the increased ground temperature ~~observed to increase~~ at depths deeper than 20 m in boreholes during the past three decades (Vaughan et al., 2013), most models that do not have deeper soil depth seem to underestimate the penetration of heat into deep soil layers (Figure 5b). Note that this comparison may be biased because of different periods and climate records between sites and model grid cells. It is also recommended that simulations at site level using in-situ local climate forcing can be compared with temperature profiles of boreholes (Smith et al., 2012) to evaluate why models underestimate the warming of T_s at deeper depths.

3.3 Drivers of trend in soil temperature

We used the sensitivity runs (R02-R06) compared with the reference simulation with all drivers varying together (R01) to separate the effects of T_a, CO_2, precipitation, and LWDR on $\dot{T}_s$ during 1960-2000 (Table 3). Seven of the nine models only provided results from R02,

R03 and R04. Except for JULES, all the models show a positive response of T_s to increasing T_a, but with different sensitivities (Table 3). The fraction of the trend of T_s explained by air temperature increase alone (R01 – R02) is nearly 100% in CLM, ISBA and more than 100% in UW-VIC, against only 34%, 56% and 67% in ORCHIDEE, UVic and LPJ-GUESS. This indicates the importance of increasing T_a on the trend of T_s, and is consistent with observations. Based on 30 climate stations observations in Canada during the period 1958-2008, T_s at 10 cm significantly and positively correlates with T_a at most sites (>90%) in spring, but at fewer sites (<30%) in winter (Qian et al., 2011). For winter T_s, the winter snow depth was found to have significant and positive correlation with T_s in shallow soil layers (e.g. Zhang et al., 2001; Qian et al., 2011). Recent increases in T_a also explain the trend of T_s at 1.6m measured at Churapcha metrological station (N62.02, E132.36), and at 5 m measured in a borehole at Iqaluit (N63.47, W68.48) in Canada (Smith et al., 2005; Romanovsky et al., 2007). To some extent, the trend of T_a is a good indicator for the trend of deep permafrost ground temperature with some time lag (Romanovsky et al., 2007). For the modeled T_s in land surface models, the effects of T_a on T_s depend on surface energy balance and ground heat flux into soil; i.e. the extent of coupled T_a on T_s relates to the surface properties such as snow, organic soil horizons and rougness etc. in the models. The different relative contributions of the trend of T_a to the trend of T_s in these models maybe mainly result from the different model parameterization and structures, as the trends of T_a (~0.03 °C yr-1) in the climate forcing do not have a large spread (Figure 7).

The increase of atmospheric CO_2 concentration has almost no effect on the increase of T_s in most models (-5% to +4% of increase of T_s, Table 3). This is expected since CO_2 has no direct effect on T_s apart from its impact on climate. The only indirect effect of rising CO_2 on T_s trends could result from feedbacks between plant productivity driven by rising CO_2, soil

carbon changes and soil thermal properties. For instance, if models include heat production from microbial decomposition of soil organic carbon (Khvorostyanov et al., 2008) or if changes in soil organic carbon from the balance of NPP input and decomposition, these could impact the soil temperature directly or the profile of soil heat conductivity and capacity. In that case, the expected response is that a CO_2 driven increase of productivity will increase soil organic carbon, which will enhance the insulation effect of soil organic carbon in the soil and ~~act~~ lower the trend of T_s (Lawrence et al., 2008; Lawrence and Slater, 2008; Koven et al., 2009). Further, complex changes in the surface energy balance from changes in evapotranspiration under higher CO_2 concentrations can influence soil moisture content and affect T_s trends (e.g. Field et al., 1995). Most models do not have a feedback between soil organic carbon dynamics and soil thermal properties, and the increase in soil organic carbon due to rising CO_2 is relatively small in the models compared to the initial soil organic carbon storage (< 0.1%). The changes in evapotranspiration because of increasing CO_2 are also relatively small (-3% to +1%). Therefore, the increased CO_2 concentration has a very small effect on $\dot{T}_s$ from 1960 to 2000.

Precipitation shows an increase in BONA and BOEU and a decrease ~~since 1960~~ in BOAS in the climate forcing used by most models (Figure S1b). None of the ~~regional~~ trends of boreal precipitation are~~is~~ significant (P>0.05; except for the UW-VIC and JULES drivers). Changes in precipitation alone (R02 – R04) are found to cause a negative trend of T_s in CLM, JULES and UW-VIC, no effects in LPJ-GUESS and UVic, and a positive trend in ISBA and ORCHIDEE (Table 3). Increasing winter snow~~-~~fall can enhance T_s in winter through snow insulation effect (e.g. Smith et al., 2010; Koven et al., 2013). All models in this study indeed show higher winter T_s where winter snow depth became ~~thicker~~deeper, but with different magnitudes of snow insulation effects across the models. The snow insulation effects are

smaller in ISBA, LPJ-GUESS and UVic than that in the other models. A decrease in ~~snow fall~~snowfall could contribute a negative trend of T_s in CLM~~ and JULES~~, and ~~conversely,~~ an increase in ~~snow fall~~snowfall could enhance T_s in ORCHIDEE (Figure S4; Table 3). In addition, increased rainfall in summer can cause an increase in evapotranspiration during the growing seasons, which could depress the increase of T_s. The effects of snowfall trends and growing season precipitation trends may oppose each other as mentioned above. These two contrasting effects cannot be separated in this analysis, because models did not run simulations with seasonally detrended precipitation. But the different effects of seasonal precipitation on T_s should be studied in the future.~~Because models did not run simulations where precipitation is detrended only during a specific season, the effects of snowfall trends and growing season precipitation trends that may oppose each other, and cannot be separated in this analysis.~~

LWDR significantly increased since 1960 in all models yet with different trends in the forcing data used by each modeling group (0.058 ~ 0.200 W m^{-2} yr^{-1}) (Figure S2a)~~ reflecting the different climate forcing data input~~. LWDR forcing is mainly from two reanalysis datasets (ERA and NCEP) with corrections (e.g. Weedon et al., 2011; http://dods.extra.cea.fr/data/p529viov/cruncep). ORCHIDEE and JULES performed the simulation R05 with detrended LWDR. The results of ~~R01~~ R02 – R05 allowing to attribute $\dot{T}_s$ to trends of LWDR, indicate that the increase of LWDR explains 56% and 31% of the trend of T_s since 1960 in ORCHIDEE and JULES, respectively. Increased LWDR provides ~~addtional~~additional energy to the surface, and dominates the atmosphere-to-soil energy flux in winter over boreal regions when shortwave radiation is small. Even in summer, LWDR contributes ~60% of total downward radiation (SWDR+LWDR) over boreal regions in CRUNCEP. An increase of LWDR with time thus increases the surface energy input, which

accelerates the warming of T_s in case the extra energy is not dissipated by an increase of sensible and latent heat flux. The contribution of changes in LWDR, T_a and other factors on all components of the surface energy budget and on T_s could be further studied by testing models against observations from eddy-flux towers located in permafrost soils.

3.4 Uncertainty of modeled soil temperature trends

The uncertainty of modeled $\dot{T}_s$ at 20 cm is large, as given by the spread of model results (0.010 °C yr^{-1} - 0.031 °C yr^{-1}). The uncertainty of $\dot{T}_s$ across the models can be conceptually decomposed into two components, a forcing uncertainty (FU) reflecting how different climate input data used by each modeling group contribute to the spread of $\dot{T}_s$ (Table 1), and a structural uncertainty (SU) related to uncertain parameter values and different equations and parameterizations of processes in models. Since T_a and LWDR are the two main drivers of the increase of T_s in most of the models (Section 3.3), we regressed $\dot{T}_s$ during 1960-2000 against the trends of T_a and LWDR, in order to estimate the FU. We then estimated SU from the uncertainty of parameters in the regression equation for a normalized same climate forcing across all models.

We found no significant correlation between $\dot{T}_a$ and $\dot{T}_s$ over boreal regions or sub-regions across the nine models (Figure 7 and Figure S5), indicating that a bias of $\dot{T}_a$ forcing is not simply associated with the bias of $\dot{T}_s$ in a particular model compared to the others. We also found that trends of SWDR and precipitation do not significantly explain differences in $\dot{T}_s$ at 20 cm across the models ($P>0.05$; 21% and 19% explanation of

differences in $\dot{T}_s$ at 20 cm for trends of SWDR and precipitation respectively; Figure S6). The correlations between trends in winter ~~snow fall~~snowfall and trends of annual or winter T_s at 20 cm are not significant (P>0.05) across the models for boreal regions or sub-regions. However, the trend of LWDR ($\dot{LWDR}$) can explain 61% of the differences in $\dot{T}_s$ at 20 cm across the models (Figure 8). This result indicates that, across the model ensemble, differences of $\dot{T}_s$ at 20 cm between models are positively correlated (R=0.78, P=0.037) with differences of $\dot{LWDR}$ used by the different modeling groups. $\dot{T}_s$ at 1 m also significantly correlated with $\dot{LWDR}$ (R=0.79, P=0.034) across the models. The values of $\dot{LWDR}$ used by different models averaged over permafrost regions, range from 0.058 W m^{-2} yr^{-1} to 0.200 W m^{-2} yr^{-1}, statistically explaining a range of simulated $\dot{T}_s$ at 20 cm of 0.021 ± 0.005 °C yr^{-1} (solid blue arrow in Figure 8). This $\dot{T}_s$ range defines the FU (the range of $\dot{T}_s$ to $\dot{LWDR}$ from 0.058 W m^{-2} yr^{-1} to 0.200 W m^{-2} yr^{-1} based on the linear regression of Figure 8). We also used multiple linear regression between $\dot{T}_s$ at 20 cm depth and $\dot{T}_a$, $\dot{LWDR}$ as independent variables across the models, to derive an estimation of the FU on $\dot{T}_s$ of 0.021 ± 0.008 °C yr^{-1} (the deviation was derived from the uncertainty of regression coefficients in the multiple linear regression). However, the uncertainty of the linear regression of $\dot{T}_s$ at 20 cm by $\dot{LWDR}$ or $\dot{T}_a$ and $\dot{LWDR}$ shows that if all the models used the same climate forcing data, the SU would be of 0.012 ± 0.001 °C yr^{-1} (solid orange arrow in Figure 8). If all models used LWDR from CRUNCEP or WATCH, then applying the trend of annual LWDR (0.087 ± 0.023 W m^{-2} yr^{-1} from CRUNCEP and 0.187 ± 0.028 W m^{-2} yr^{-1} from WATCH) during the period 1960-2000 as an emerging observation constraint

393 empirical relationship in Figure 8, the posterior range is reduced compared with the prior $\dot{T}_s$
394 range (black curve in right panel of Figure 8). Overall, the total uncertainty range of $\dot{T}_s$ at 20
395 cm (~0.02 °C yr^{-1}, defined as the spread of $\dot{T}_s$ at 20 cm across the models) can be broken
396 down into FU (0.021 ± 0.008 °C yr^{-1}) and SU (0.012 ± 0.001 °C yr^{-1}). Since FU and SU are
397 not independent, the total uncertainty of $\dot{T}_s$ at 20 cm is not the sum of FU and SU.
398
399 Further, we found that correlation coefficients between trends of summer T_s at 20 cm and
400 at 1 m and summer LWDR over boreal regions are statistically significant ($P<0.05$) (Figure
401 S7). This is also found for winter (November to March) T_s at 20 cm and 1 m (Figure S8).
402 Trends of summer and winter T_s at 20 cm or 1 m are not significantly correlated with other
403 climate drivers than LWDR (snowfall, rainfall, T_a and SWDR) across the models ($P>0.05$).
404
405 Meteorological stations are sparse in the cold permafrost regions. For example, there are
406 only 8.8 stations per million km^2 north of 60°N in the CRU TS3.22 gridded air temperature
407 product compared to 41.1 stations per million km^2 between 25°N and 60°N. This results in
408 uncertainty in gridded climate products over Arctic regions, especially for trends of Arctic
409 climate variables (Mitchell and Jones, 2005; Troy and Wood, 2009; Rawlins et al., 2010;
410 Weedon et al., 2011). Troy and Wood (2009) reported 15-20 W m^{-2} of differences in radiative
411 fluxes on seasonal timescales over northern Eurasia, between six gridded products. Between
412 different gridded observations and reanalysis precipitation products, the magnitude of Arctic
413 precipitation ranges from 410 mm yr^{-1} to 520 mm yr^{-1}, and the trend of Arctic precipitation
414 also has a large spread (Rawlins et al., 2010). These large uncertainties in climate forcing in
415 Arctic undoubtedly can cause large spread of modeled T_s. We found that the FU dominates
416 the total uncertainty of $\dot{T}_s$. This suggests that modelers not only need to improve their

models, but also need better climate forcing data (or need to test the effects of different climate input data) when modeling long term changes of T_s in permafrost regions. However, to quantify the SU, simulations using the same agreed upon climate forcing data are highly recommended to further attribute the contribution of each process in the soil thermal dynamics of models such as organic carbon insulation effects, snow insulation effects, latent heat formation and emission, soil conductivity and surface properties (see Lawrence and Slater, 2008; Koven et al., 2009; Bonfils et al., 2012; Gouttevin et al., 2012). In addition, important processes in permafrost regions such as ice content (e.g. ice wedge) in permafrost and thermokarst lakes etc. should be developed and evaluated in land surface models to improve the prediction of future permafrost feedbacks (e.g. van Huissteden et al., 2013; Lee et al., 2014).

3.5 Emerging constraint on how much near-surface permafrost has disappeared.

~~All the nine models provide estimates of~~The total boreal NSPA during 1960-2000 estimated by the nine models ~~that~~ ranges from 6.8 million km^2 ~~in~~ (CoLM) to 19.7 million km^2 ~~in~~ (ORCHIDEE)~~, which is three times as large as the former~~. The average of total NSPA in the nine models ensemble (12.5 million km^2) is smaller than the estimate from the International Permafrost Association (IPA) map (16.2 million km^2; Brown et al, 1998; Slater and Lawrence ~~and Slater~~ et al., 2013). A statistic model based on relationships between air temperature and permafrost shows that permafrost extent over Northern Hemisphere was also estimated in the range 12.9 - 17.8 million km^2 (Gruber, 2012), and six out of the nine models are within this range. Eight out of the nine models show a significant decrease in NSPA with climate warming during 1960-2000 (except UW-VIC). The loss rate of NSPA is found to vary by a factor of 13 across the nine models, varying from -4×10^3 km^2 yr^{-1} in MIROC-ESM to -50×10^3 km^2 yr^{-1} in JULES (Figure 9a). The average of loss rate of NSPA across the models (-23

± 23 $\times 10^3$ km^2 yr^{-1}) is smaller than in the previous estimations of Burke et al. (2013) and Slater and Lawrence (2013). For example, the loss rate of NSPA was estimated at -81×10^3 – -55×10^3 km^2 yr^{-1} during the period 1967-2000 by JULES offline simulations with different climate forcing datasets (Burke et al., 2013). The ranges of loss rate of NSPA in BONA, BOEU and BOAS across the models are -16.6×10^3 – 2.2×10^3 km^2 yr^{-1}, -4.0×10^3 – 0.0×10^3 km^2 yr^{-1} and -34.2×10^3 – -1.1×10^3 km^2 yr^{-1}, respectively (Figure 9). This is consistent with the observed permafrost degradation (decrease in thickness) in these regions (Vaughan et al., 2013).

The retreat rate of NSPA is not correlated significantly with the initial NSPA of each model (R=-0.30, P=0.438), implying that the initial state of models is less important than their response to climate change in determining NSPA loss rates. On the contrary to the small effect of initial NSPA, the trend of summer T_s at 1 m is found to be strongly correlated with NSPA loss rates across the models of the ensemble. Figure 9 shows that the trend of summer T_s at 1 m can explain 73% of the differences in NSPA loss rates between models. The sensitivity of NSPA loss rate to summer $\dot{T}_s$ at 1 m is estimated to be -2.80 ± 0.67 million km^2 $°C^{-1}$, based on the linear regression between the loss rate of NSPA and the trend of summer T_s at 1 m across the nine models (Figure 9). For the BONA, BOEU and BOAS sub-regions, the sensitivities of NSPA loss rate to summer $\dot{T}_s$ at 1 m are -0.74 ± 0.10 million km^2 $°C^{-1}$, -0.09 ± 0.03 million km^2 $°C^{-1}$ and -1.74 ± 0.59 million km^2 $°C^{-1}$, respectively (Figure 9). The sensitivity of future total NSPA changes to T_a over Pan-Arctic regions was estimated to be -1.67 ± 0.7 million km^2 $°C^{-1}$, ranging from 0.2 million km^2 $°C^{-1}$ to 3.5 million km^2 $°C^{-1}$ in CMIP5 model ensembles (Slater and Lawrence, 2013; Koven et al., 2013). The average of trends in summer T_s ~~trends~~ at 1 m is only 70% (43%-100%) of $\dot{T}_a$ in the nine models, so that

the sensitivity of total NSPA to T_a over boreal regions in the nine models is about -2.00 ± 0.47 million km^2 °C^{-1}, which is larger than that from CMIP5 model ensemble, but comparable within the uncertainties of each estimate (Slater and Lawrence, 2013). Six out of the nine models of this study were also used as land surface schemes of the coupled CMIP5 models, but possibly for different versions.

A mean positive trend of summer LWDR of 0.073 ± 0.030 W m^2 yr^{-1} and 0.210 ± 0.027 W m^2 yr^{-1}over boreal regions from 1960 to 2000 are derived from the CRUNCEP and WATCH datasets respectively ~~described in Section 2~~. We applied this trend of LWDR to an emerging constraint on summer T_s trends from the relationship between the trend of summer LWDR and the trend of summer T_s at 1 m (Figure S7). This approach constrains the trend of summer T_s to 0.014 ± 0.004 °C yr^{-1} with CRUNCEP and to 0.027 ± 0.004 °C yr^{-1} with WATCH. The uncertainty is reduced by 50% from the prior range including different models and different forcings. A total NSPA loss rate of 39 ± 14 ×10^3 km^2 yr^{-1} can be constrained by multiplying the sensitivity of total NSPA loss rate to summer $\dot{T}_s$ at 1 m (-2.80 ± 0.67 million km^2 °C^{-1}) by the trend of T_s at 1 m itself empirically estimated by $\dot{LWDR}$ during 1960-2000 from CRUNCEP (0.014 ± 0.004 °C yr^{-1}). The constrained loss rate of NSPA over BONA, BOEU and BOAS based upon the CRUNCEP $\dot{LWDR}$ from 1960 to 2000 are 11 ± 5 ×10^3 km^2 yr^{-1}, 1 ± 1 ×10^3 km^2 yr^{-1} and 25 ± 11 ×10^3 km^2 yr^{-1}, respectively. Similarly, if WATCH $\dot{LWDR}$ is used to constrain NSPA loss rate, the total NSPA loss rate is 75 ± 14 ×10^3 km^2 yr^{-1}, and loss rate of NSPA over BONA, BOEU and BOAS are estimated to be 28 ± 10 ×10^3 km^2 yr^{-1}, 2 ± 1 ×10^3 km^2 yr^{-1} and 39 ± 19 ×10^3 km^2 yr^{-1}, respectively. The southern boundary of the discontinuous permafrost zone has been observed to shift northward during the recent~~past~~ decades ~~is also observed~~ (Vaughan et al., 2013), which is generally consistent

with the simulations reported in this study. The larger warming rate and higher sensitivity of NSPA loss to T_s over BOAS could explain the reason for significant degradation of permafrost over BOAS than the other boreal regions (Vaughan et al., 2013). The larger permafrost degradation rate in BOAS than that in BONA may have larger effects on changes in vegetation distribution and growth, and permafrost carbon in these two regions, and can be quantified in future studies. Obviously, there is a large difference in constrained NSPA between CRUNCEP and WATCH. In the future, long-term climate reanalysis including radiation evaluated against sites with long-term radiation measurements (http://www.geba.ethz.ch) would be extremely useful for land surface models to provide improved estimate of NSPA.

4 Conclusions

In this study, trends of T_s over boreal regions from nine process-based models were analyzed for the past 40 years. All models produce a warming of T_s, but the trends of T_s at 20 cm depth range from 0.010 ± 0.003 °C yr^{-1} (CoLM) to 0.031 ± 0.005 °C yr^{-1} (UVic) during 1960-2000. Most models show a smaller increase of T_s with deeper depth. T_a and LWDR are found to be the predominant drivers of the increase in T_s averaged across large spatial scales. The relative contribution of T_a and LWDR trends to the increase of T_s is however different across the models. Note that the relative contribution of LWDR is based on only two models in this study, and this needs further investigation. The total uncertainty of the trend of T_s at 20 cm is decomposed into the uncertainty contributed by uncertain climate forcing datasets (0.021 ± 0.008 °C yr^{-1}) and the uncertainty reflecting model structure (0.012 ± 0.001 °C yr^{-1}). The NSPA loss rate is significantly correlated among the model results with the simulated trend of T_s at 1 m, with a linear sensitivity of total NSPA loss rate to summer $\dot{T}_s$ at 1 m of -2.80 ± 0.67 million km^2 $°C^{-1}$. Based on LWDR from CRUNCEP and WATCH data, the total

NSPA decrease is estimated to be 39 ± 14 $\times 10^3$ km^2 yr^{-1} - 75 ± 14 $\times 10^3$ km^2 yr^{-1} from 1960 to 2000. The constraint method used in this study could be applied to estimate historical and future permafrost degradation rate, and further to quantify the permafrost carbon loss by permafrost carbon distribution map.

Given that meteorological stations are sparse in the cold permafrost regions, especially in Siberia and other unpopulated land in the north, the gridded climate products over high-latitude regions have a large uncertainty as well (Mitchell and Jones, 2005; Rawlins et al., 2010; Weedon et al., 2011). This large uncertainty could propagate into simulated permafrost dynamics and feedbacks. More sites are needed in high-latitude regions for reducing the climate uncertainty. Future model inter-comparisons on permafrost dynamics should investigate the full uncertainty by conducting simulations for multiple climate forcing data sets. Since the beginning of the satellite era, microwave emissivity data related to land surface temperature has become increasingly available (e.g. Smith et al., 2004). These images could be used to independently evaluate soil surface temperature in models on a large scale, although they have their own uncertainties. In addition, many complex processes affect permafrost thermal dynamics in the models, such as soil organic insulation effects, snow insulation effects, soil freeze-thaw etc., it is valuable to evaluate the uncertainty of each process effects on soil thermal dynamics simulations based on site measurements. This could be helpful for reducing permafrost simulation uncertainty.

Acknowledgments

This study has been supported by the PAGE21 project, funded by the European Commission FP7-ENV-2011 (grant agreement NO. 282700) and has been developed as part

of the modeling integration team of the Permafrost Carbon Network (PCN, www.permafrostcarbon.org) funded by the National Science Foundation. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government. Bertrand Decharme and Christine Delire were supported by the French Agence Nationale de la Recherche under agreement ANR-10-CEPL-012-03.

References

Adam, J. C., Clark, E. A., Lettenmaier, D. P., and Wood, E. F.: Correction of Global Precipitation Products for Orographic Effects, Journal of Climate, 19, 15-38, doi:10.1175/jcli3604.1, 2006.

Alexeev, V. A., Nicolsky, D. J., Romanovsky, V. E., and Lawrence, D. M.: An evaluation of deep soil configurations in the CLM3 for improved representation of permafrost, Geophysical Research Letters, 34, L09502, doi:10.1029/2007gl029536, 2007.

Avis, C. A., Weaver, A. J., and Meissner, K. J.: Reduction in areal extent of high-latitude wetlands in response to permafrost thaw, Nature Geosci, 4, 444-448, 2011.

Best, M. J., Pryor, M., Clark, D. B., Rooney, G. G., Essery, R. L. H., Ménard, C. B., Edwards, J. M., Hendry, M. A., Porson, A., Gedney, N., Mercado, L. M., Sitch, S., Blyth, E., Boucher, O., Cox, P. M., Grimmond, C. S. B., and Harding, R. J.: 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.

Bohn, T. J., Podest, E., Schroeder, R., Pinto, N., McDonald, K. C., Glagolev, M., Filippov, I., Maksyutov, S., Heimann, M., Chen, X., and Lettenmaier, D. P.: Modeling the large-scale effects of surface moisture heterogeneity on wetland carbon fluxes in the West Siberian Lowland, Biogeosciences, 10, 6559-6576, doi:10.5194/bg-10-6559-2013, 2013.

Bonfils, C. J. W., Phillips, T. J., Lawrence, D. M., Cameron-Smith, P., Riley, W. J., and Subin,

565 Z. M.: On the influence of shrub height and expansion on northern high latitude climate,
566 Environmental Research Letters, 7, 015503, doi:10.1088/1748-9326/7/1/015503, 2012.
567 Brown, J., Ferrians, O.J.Jr., Heginbottom, J.A.. and Melnikov, E.S. (1998, revised February
568 2001). Circum-Arctic map of permafrost and ground-ice conditions. Boulder, CO:
569 National Snow and Ice Data Center/World Data Center for Glaciology. Digital Media.
570 Burke, E., Dankers, R., Jones, C., and Wiltshire, A.: A retrospective analysis of pan Arctic
571 permafrost using the JULES land surface model, Clim Dyn, 41, 1025-1038,
572 doi:10.1007/s00382-012-1648-x, 2013.
573 Clark, D. B., Mercado, L. M., Sitch, S., Jones, C. D., Gedney, N., Best, M. J., Pryor, M.,
574 Rooney, G. G., Essery, R. L. H., Blyth, E., Boucher, O., Harding, R. J., Huntingford, C.,
575 and Cox, P. M.: The Joint UK Land Environment Simulator (JULES), model description
576 – Part 2: Carbon fluxes and vegetation dynamics, Geosci. Model Dev., 4, 701-722,
577 doi:10.5194/gmd-4-701-2011, 2011.
578 Dai, Y., Zeng, X., Dickinson, R. E., Baker, I., Bonan, G. B., Bosilovich, M. G., Denning, A. S.,
579 Dirmeyer, P. A., Houser, P. R., Niu, G., Oleson, K. W., Schlosser, C. A., and Yang, Z.-L.:
580 The Common Land Model, Bulletin of the American Meteorological Society, 84,
581 1013-1023, doi:10.1175/bams-84-8-1013, 2003.
582 Dai, Y., Dickinson, R. E., and Wang, Y.-P.: A Two-Big-Leaf Model for Canopy Temperature,
583 Photosynthesis, and Stomatal Conductance, Journal of Climate, 17, 2281-2299,
584 doi:10.1175/1520-0442(2004)017<2281:atmfct>2.0.co;2, 2004.
585 Decharme, B., Boone, A., Delire, C., and Noilhan, J.: Local evaluation of the Interaction
586 between Soil Biosphere Atmosphere soil multilayer diffusion scheme using four
587 pedotransfer functions, Journal of Geophysical Research: Atmospheres, 116, D20126,
588 doi:10.1029/2011jd016002, 2011.
589 Decharme, B., Martin, E., and Faroux, S.: Reconciling soil thermal and hydrological lower

boundary conditions in land surface models, Journal of Geophysical Research: Atmospheres, 118, 7819-7834, doi:10.1002/jgrd.50631, 2013.

Elberling, B., Michelsen, A., Schadel, C., Schuur, E. A. G., Christiansen, H. H., Berg, L., Tamstorf, M. P., and Sigsgaard, C.: Long-term CO2 production following permafrost thaw, Nature Clim. Change, 3, 890-894, doi:10.1038/nclimate1955, 2013.

Field, C. B., Jackson, R. B., and Mooney, H. A.: Stomatal responses to increased CO2: implications from the plant to the global scale, Plant, Cell & Environment, 18, 1214-1225, doi:10.1111/j.1365-3040.1995.tb00630.x, 1995.

Gouttevin, I., Krinner, G., Ciais, P., Polcher, J., and Legout, C.: Multi-scale validation of a new soil freezing scheme for a land-surface model with physically-based hydrology, The Cryosphere, 6, 407-430, doi:10.5194/tc-6-407-2012, 2012.

Gruber, S.: Derivation and analysis of a high-resolution estimate of global permafrost zonation, The Cryosphere, 6, 221-233, doi:10.5194/tc-6-221-2012, 2012.

Harris, I., Jones, P. D., Osborn, T. J., and Lister, D. H.: Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset, International Journal of Climatology, 34, 623-642, doi:10.1002/joc.3711, 2014.

Hugelius, G., Strauss, J., Zubrzycki, S., Harden, J. W., Schuur, E. A. G., Ping, C. L., Schirrmeister, L., Grosse, G., Michaelson, G. J., Koven, C. D., O'Donnell, J. A., Elberling, B., Mishra, U., Camill, P., Yu, Z., Palmtag, J., and Kuhry, P.: Estimated stocks of circumpolar permafrost carbon with quantified uncertainty ranges and identified data gaps, Biogeosciences, 11, 6573-6593, doi:10.5194/bg-11-6573-2014, 2014.

Ji, D., Wang, L., Feng, J., Wu, Q., Cheng, H., Zhang, Q., Yang, J., Dong, W., Dai, Y., Gong, D., Zhang, R.-H., Wang, X., Liu, J., Moore, J. C., Chen, D., and Zhou, M.: Description and basic evaluation of Beijing Normal University Earth System Model (BNU-ESM)

615 version 1, Geosci. Model Dev., 7, 2039-2064, doi:10.5194/gmd-7-2039-2014, 2014.
616 Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha,
617 S., White, G., Woollen, J., Zhu, Y., Leetmaa, A., Reynolds, R., Chelliah, M., Ebisuzaki,
618 W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Jenne, R., and
619 Joseph, D.: The NCEP/NCAR 40-Year Reanalysis Project, Bulletin of the American
620 Meteorological Society, 77, 437-471,
621 doi:10.1175/1520-0477(1996)077<0437:tnyrp>2.0.co;2, 1996.
622 Khvorostyanov, D. V., Ciais, P., Krinner, G., Zimov, S. A., Corradi, C., and Guggenberger,
623 G.: Vulnerability of permafrost carbon to global warming. Part II: sensitivity of
624 permafrost carbon stock to global warming, Tellus B, 60, 265-275,
625 doi:10.1111/j.1600-0889.2007.00336.x, 2008.
626 Koven, C., Friedlingstein, P., Ciais, P., Khvorostyanov, D., Krinner, G., and Tarnocai, C.: On
627 the formation of high - latitude soil carbon stocks: Effe
628 insulation by organic matter in a land surface model, Geophysical Research Letters, 36,
629 L21501, doi:10.1029/2009gl040150, 2009.
630 Koven, C. D., Ringeval, B., Friedlingstein, P., Ciais, P., Cadule, P., Khvorostyanov, D.,
631 Krinner, G., and Tarnocai, C.: Permafrost carbon-climate feedbacks accelerate global
632 warming, P. Natl. Acad. Sci. USA, 14769-14774, doi:10.1073/pnas.1103910108, 2011.
633 Koven, C. D., Riley, W. J., and Stern, A.: Analysis of Permafrost Thermal Dynamics and
634 Response to Climate Change in the CMIP5 Earth System Models, Journal of Climate, 26,
635 1877-1900, doi:10.1175/jcli-d-12-00228.1, 2013.
636 Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais,
637 P., Sitch, S., and Prentice, I. C.: A dynamic global vegetation model for studies of the
638 coupled atmosphere-biosphere system, Global Biogeochemical Cycles, 19, GB1015,
639 doi:10.1029/2003gb002199, 2005

640 Lawrence, D. M., Slater, A. G., Romanovsky, V. E., and Nicolsky, D. J.: Sensitivity of a model
641 projection of near-surface permafrost degradation to soil column depth and
642 representation of soil organic matter, Journal of Geophysical Research: Earth Surface,
643 113, F02011, doi:10.1029/2007jf000883, 2008.
644 Lawrence, D., and Slater, A.: Incorporating organic soil into a global climate model, Clim
645 Dyn, 30, 145-160, doi:10.1007/s00382-007-0278-1, 2008.
646 Lee, H., Swenson, S. C., Slater, A. G., and Lawrence, D. M.: Effects of excess ground ice on
647 projections of permafrost in a warming climate, Environmental Research Letters, 9,
648 124006, doi:10.1088/1748-9326/9/12/124006, 2014.
649 MacDougall, A. H., Avis, C. A., and Weaver, A. J.: Significant contribution to climate
650 warming from the permafrost carbon feedback, Nature Geosci., 5, 719-721, doi:
651 10.1038/ngeo1573, 2012.
652 McGuire, A. D., Anderson, L. G., Christensen, T. R., Dallimore, S., Guo, L., Hayes, D. J.,
653 Heimann, M., Lorenson, T. D., Macdonald, R. W., and Roulet, N.: Sensitivity of the
654 carbon cycle in the Arctic to climate change, Ecological Monographs, 79, 523-555,
655 doi:10.1890/08-2025.1, 2009.
656 McGuire, A. D., Christensen, T. R., Hayes, D., Heroult, A., Euskirchen, E., Kimball, J. S.,
657 Koven, C., Lafleur, P., Miller, P. A., Oechel, W., Peylin, P., Williams, M., and Yi, Y.: An
658 assessment of the carbon balance of Arctic tundra: comparisons among observations,
659 process models, and atmospheric inversions, Biogeosciences, 9, 3185–3204,
660 doi:10.5194/bg-9-3185-2012, 2012.
661 McGuire, D. Lawrence, D., Burke, E., Chen, X., Delire, C., Koven, C., MacDougall, A., Peng,
662 S., Rinke, A., Saito, K., Zhang, W., Alkama, R., J. Bohn, T., Ciais, P., Decharme, B.,
663 Gouttevin, I., Hajima, T., Ji, D., Krinner, G., Lettenmaier, D. P., Miller, P., Moore, J. C.,
664 Smith, B., and Sueyoshiet, T.: An retrospective assessment of the vulnerability of

665 permafrost carbon in the earth system: comparison of dynamics among process-based
666 models, in preparation, 2015.
667 Mitchell, T. D., and Jones, P. D.: An improved method of constructing a database of monthly
668 climate observations and associated high-resolution grids, International Journal of
669 Climatology, 25, 693-712, doi:10.1002/joc.1181, 2005.
670 Oleson, K.W., Lawrence, D. M., Bonan, G. B., Drewniak, B., Huang, M., Koven, C. D., Levis,
671 S., Li, F., Riley, W. J., Subin, Z. M., Swenson, S. C., Thornton, P. E., Bozbiyik, A., Fisher,
672 R., Heald, C. L., Kluzek, E., Lamarque, J., Lawrence, P. J., Leung, L. R., Lipscomb, W.,
673 Muszala, S., Ricciuto, D. M., Sacks, W., Tang, J., Yang, Z. Technical Description of
674 version 4.5 of the Community Land Model (CLM). P. O. Box 3000 BOULDER,
675 COLORADO 80307-3000: National Center for Atmospheric Research. 2013 June.
676 NCAR Technical Note #NCAR/TN-503+STR.
677 Osterkamp, T. E. A thermal history of permafrost in Alaska. Proceedings of Eighth
678 International Conference on Permafrost, Zurich, pp. 863-868, July 21-25, 2003.
679 Osterkamp, T. E.: Characteristics of the recent warming of permafrost in Alaska, Journal of
680 Geophysical Research: Earth Surface, 112, F02S02, doi:10.1029/2006jf000578, 2007.
681 Park, H., Sherstiukov. A. B., Fedorov, A. N., Polyakov, I. V., and Walsh, J. E.: An
682 observation-based assessment of the influences of air temperature and snow depth on soil
683 temperature in Russia, Environmental Research Letters, 9, 064026,
684 doi:10.1088/1748-9326/9/6/064026, 2014.
685 Qian, B., Gregorich, E. G., Gameda, S., Hopkins, D. W., and Wang, X. L.: Observed soil
686 temperature trends associated with climate change in Canada, Journal of Geophysical
687 Research: Atmospheres, 116, D02106, doi:10.1029/2010jd015012, 2011.
688 Rawlins, M. A., Steele, M., Holland, M. M., Adam, J. C., Cherry, J. E., Francis, J. A.,
689 Groisman, P. Y., Hinzman, L. D., Huntington, T. G., Kane, D. L., Kimball, J. S., Kwok,

690 R., Lammers, R. B., Lee, C. M., Lettenmaier, D. P., McDonald, K. C., Podest, E.,
691 Pundsack, J. W., Rudels, B., Serreze, M. C., Shiklomanov, A., Skagseth, O., Troy, T. J.,
692 Vorosmarty, C. J., Wensnahan, M., Wood, E. F., Woodgate, R., Yang, D. Q., Zhang, K.,
693 and Zhang, T. J.: Analysis of the Arctic System for Freshwater Cycle Intensification:
694 Observations and Expectations, Journal of Climate, 23, 5715-5737,
695 doi:10.1175/2010jcli3421.1, 2010.
696 Rawlins, M. A., McGuire, A. D., Kimball, J. K., Dass, P., Lawrence, D., Burke, E., Chen, X.,
697 Delire, C., Koven, C., MacDougall, A., Peng, S., Rinke, A., Saito, K., Zhang, W.,
698 Alkama, R., J. Bohn, T., Ciais, P., Decharme, B., Gouttevin, I., Hajima, T., Ji, D., Krinner,
699 G., Lettenmaier, D. P., Miller, P., Moore, J. C., Smith, B., and Sueyoshi, T.: Assessment
700 of model estimates of land–atmosphere CO2 exchange across Northern Eurasia,
701 Biogeosciences Discuss., 12, 2257-2305, doi:10.5194/bgd-12-2257-2015, 2015.
702 Romanovsky, V. E., Sazonova, T. S., Balobaev, V. T., Shender, N. I., and Sergueev, D. O.: Past
703 and recent changes in air and permafrost temperatures in eastern Siberia, Global and
704 Planetary Change, 56, 399-413, doi:10.1016/j.gloplacha.2006.07.022, 2007.
705 Romanovsky, V. E., Smith, S. L., and Christiansen, H. H.: Permafrost thermal state in the
706 polar Northern Hemisphere during the international polar year 2007–2009: a synthesis,
707 Permafrost and Periglacial Processes, 21, 106-116, doi:10.1002/ppp.689, 2010.
708 Schädel, C., Schuur, E. A. G., Bracho, R., Elberling, B., Knoblauch, C., Lee, H., Luo, Y.,
709 Shaver, G. R., and Turetsky, M. R.: Circumpolar assessment of permafrost C quality and
710 its vulnerability over time using long-term incubation data, Global Change Biology, 20,
711 641-652, doi:10.1111/gcb.12417, 2014.
712 Schaefer, K., Zhang, T., Bruhwiler, L., and Barrett, A. P.: Amount and timing of permafrost
713 carbon release in response to climate warming, Tellus B, 63, 165-180,
714 doi:10.1111/j.1600-0889.2011.00527.x, 2011.

715 Schneider von Deimling, T., Meinshausen, M., Levermann, A., Huber, V., Frieler, K.,
716 Lawrence, D. M., and Brovkin, V.: Estimating the near-surface permafrost-carbon
717 feedback on global warming, Biogeosciences, 9, 649-665, doi:10.5194/bg-9-649-2012,
718 2012.
719 Schuur, E. A. G., Bockheim, J., Canadell, J. G., Euskirchen, E., Field, C. B., Goryachkin, S.
720 V., Hagemann, S., Kuhry, P., Lafleur, P. M., Lee, H., Mazhitova, G., Nelson, F. E.,
721 Rinke, A., Romanovsky, V. E., Shiklomanov, N., Tarnocai, C., Venevsky, S., Vogel, J.
722 G., and Zimov, S. A.: Vulnerability of Permafrost Carbon to Climate Change:
723 Implications for the Global Carbon Cycle, BioScience, 58, 701-714,
724 doi:10.1641/b580807, 2008.
725 Sheffield, J., Goteti, G., and Wood, E. F.: Development of a 50-Year High-Resolution Global
726 Dataset of Meteorological Forcings for Land Surface Modeling, Journal of Climate, 19,
727 3088-3111, 10.1175/jcli3790.1, 2006.
728 Slater, A. G., and Lawrence, D. M.: Diagnosing Present and Future Permafrost from Climate
729 Models, Journal of Climate, 26, 5608-5623, doi:10.1175/jcli-d-12-00341.1, 2013.
730 Smith, B., Prentice, I. C., and Sykes, M. T.: Representation of vegetation dynamics in the
731 modelling of terrestrial ecosystems: comparing two contrasting approaches within
732 European climate space, Global Ecology and Biogeography, 10, 621-637,
733 doi:10.1046/j.1466-822X.2001.t01-1-00256.x, 2001.
734 Smith, N. V., Saatchi, S. S. and Randerson, J. T.: Trends in high northern latitude soil freeze
735 and thaw cycles from 1988 to 2002, Journal of Geophysical Research-Atmospheres, 109,
736 D12101, doi:10.1029/2003JD004472, 2004.
737 Smith, S. L., Burgess, M. M., Riseborough, D., and Mark Nixon, F.: Recent trends from
738 Canadian permafrost thermal monitoring network sites, Permafrost and Periglacial
739 Processes, 16, 19-30, doi:10.1002/ppp.511, 2005.

740 Smith, S. L., Romanovsky, V. E., Lewkowicz, A. G., Burn, C. R., Allard, M., Clow, G. D.,
741 Yoshikawa, K., and Throop, J.: Thermal state of permafrost in North America: a
742 contribution to the international polar year, Permafrost and Periglacial Processes, 21,
743 117-135, doi:10.1002/ppp.690, 2010.
744 Smith, S. L., Throop, J., and Lewkowicz, A. G.: Recent changes in climate and permafrost
745 temperatures at forested and polar desert sites in northern Canada, Canadian Journal of
746 Earth Sciences, 49, 914-924, doi:10.1139/e2012-019, 2012.
747 Stevens, M. B., Smerdon, J. E., González-Rouco, J. F., Stieglitz, M., and Beltrami, H.: Effects
748 of bottom boundary placement on subsurface heat storage: Implications for climate
749 model simulations, Geophysical Research Letters, 34, L02702,
750 doi:10.1029/2006gl028546, 2007.
751 Troy, T. J., and Wood, E. F.: Comparison and evaluation of gridded radiation products across
752 northern Eurasia, Environmental Research Letters, 4, 045008,
753 doi:10.1088/1748-9326/4/4/045008, 2009.
754 van Huissteden, J., Berrittella, C., Parmentier, F. J. W., Mi, Y., Maximov, T. C., and Dolman,
755 A. J.: Methane emissions from permafrost thaw lakes limited by lake drainage, Nature
756 Clim. Change, 1, 119-123, doi:dx.doi.org/10.1038/nclimate1101, 2011.
757 Vaughan, D.G., Comiso, J.C., Allison, I., Carrasco, J., Kaser, G., Kwok, R., Mote, P., Murray,
758 T., Paul, F., Ren, J., Rignot, E., Solomina, O., Steffen, K. and Zhang, T. (2013)
759 *Observations: Cryosphere*. In: *Climate Change 2013: The Physical Science Basis*.
760 Contribution of Working Group I to the Fifth Assessment Report of the
761 Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M.
762 Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)].
763 Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp.
764 317–382, doi:10.1017/CBO9781107415324.012.

765 Watanabe, S., Hajima, T., Sudo, K., Nagashima, T., Takemura, T., Okajima, H., Nozawa, T.,
766 Kawase, H., Abe, M., Yokohata, T., Ise, T., Sato, H., Kato, E., Takata, K., Emori, S., and
767 Kawamiya, M.: MIROC-ESM 2010: model description and basic results of
768 CMIP5-20c3m experiments, Geosci. Model Dev., 4, 845-872,
769 doi:10.5194/gmd-4-845-2011, 2011.
770 Weedon, G. P., Balsamo, G., Bellouin, N., Gomes, S., Best, M. J., and Viterbo, P.: The
771 WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to
772 ERA-Interim reanalysis data, Water Resources Research, 50, 7505-7514,
773 doi:10.1002/2014wr015638, 2014.
774 Weedon, G. P., Gomes, S., Viterbo, P., Shuttleworth, W. J., Blyth, E., Österle, H., Adam, J. C.,
775 Bellouin, N., Boucher, O., and Best, M.: Creation of the WATCH Forcing Data and Its
776 Use to Assess Global and Regional Reference Crop Evaporation over Land during the
777 Twentieth Century, Journal of Hydrometeorology, 12, 823-848,
778 doi:10.1175/2011jhm1369.1, 2011.
779 Wei, Y., Liu, S., Huntzinger, D. N., Michalak, A. M., Viovy, N., Post, W. M., Schwalm, C. R.,
780 Schaefer, K., Jacobson, A. R., Lu, C., Tian, H., Ricciuto, D. M., Cook, R. B., Mao, J.,
781 and Shi, X.: The North American Carbon Program Multi-scale Synthesis and Terrestrial
782 Model Intercomparison Project – Part 2: Environmental driver data, Geosci. Model Dev.
783 Discuss., 6, 5375-5422, doi:10.5194/gmdd-6-5375-2013, 2013.
784 Willmott, C. J. and K. Matsuura (2001) Terrestrial Air Temperature and Precipitation:
785 Monthly and Annual Time Series (1950 - 1999),
786 http://climate.geog.udel.edu/~climate/html_pages/README.ghcn_ts2.html.
787 Zhang, T., Barry, R., Gilichinsky, D., Bykhovets, S. S., Sorokovikov, V. A., and Ye, J.: An
788 Amplified Signal of Climatic Change in Soil Temperatures during the Last Century at
789 Irkutsk, Russia, Climatic Change, 49, 41-76, doi:10.1023/a:1010790203146, 2001.

790

Table 1. Soil depth for soil thermal dynamics and climate forcing used in each model.

Model	Soil depth (m)	Soil discretization layers	Bottom boundary geothermal heat flux (mW m^{-2})	Climate forcing (Reference)	Model reference	Note
CLM	45.1	30	0	CRUNCEP v4 (http://dods.extra.cea.fr/)	Oleson et al., 2013	带格式表格
CoLM	3.4	10	0	Princeton (Sheffield et al., 2006)	Dai et al., 2003, 2004	
ISBA	12.0	14	0	WATCH (1901-1978) WFDEI (1978-2009) (Weedon et al., 2011; 2014)	Decharme et al., 2011, 2013	-
JULES	20.8	16	0	WATCH (1901-2001) (Weedon et al., 2011)	Best et al., 2011; Clark et al., 2011	
LPJ-GUESS	3.0	8	0	CRU TS 3.1 (Harris et al., 2014)	Smith et al., 2001; McGuire et al., 2012	Soil temperature in the top 3 meter is based on another 6 pa the bottom layer condition. Surface shortwave downward cloudiness data set; No longwave downward radiation a
MIROC-ESM	14.0	6	0	CMIP5 Drivers (Watanabe et al., 2011)	Watanabe et al., 2011	
ORCHIDEE	47.4	32	58	WATCH (1901-1978) WFDEI (1978-2009) (Weedon et al., 2011; 2014)	Krinner et al., 2005; Koven et al., 2011; Gouttevin et al., 2012	
UVic	250.3	14	0	CRUNCEP v4 (http://dods.extra.cea.fr/)	Avis et al., 2011, MacDougall et al., 2012	Surface shortwave and longwave downward radiation
UW-VIC	25.0	25	0	temperature from CRU TS3.1, precipitation from UDel, wind speed from NCEP-NCAR (Mitchell and Jones, 2005; Willmott and Matsura, 2001; Adam et al., 2006; Kalnay et	Bohn et al., 2013	Surface shortwave and longwave downward radiation

al., 1996)

Table 2. Description of simulations used in this study.

Simulation ID	Climate	CO_2
R01	variable	variable
R02	variable, but with detrended T_a	variable
R03	variable	constant in the year of 1960
R04	variable, but with detrended T_a and precipitation	variable
R05	variable, but with detrended T_a and LWDR	variable
R06	variable, but with detrended T_a, precipitation and LWDR	variable

Table 3. The trends of annual air temperature (T_a), precipitation and longwave downward radiation (LWDR) in the second to fourth columns. The fifth column shows the trends of annual T_s at 20 cm in the reference simulation (R01). The last four columns show the contributions of drivers (T_a, precipitation, CO_2 and LWDR) on the trend of T_s as mentioned in Methods section. The relative contributions (divided by the trend of T_s in Ref) are shown in the parentheses. The bold font indicates statistically significant ($P<0.05$).

Model	Trend of T_a ($^oC\ yr^{-1}$)	Trend of precipitation ($mm\ yr^{-2}$)	Trend of LWDR ($W\ m^{-2}\ yr^{-1}$)	Simulated Trend of T_s (R01) ($^oC\ yr^{-1}$)	Contribution from T_a (R01-R02) ($^oC\ yr^{-1}$)	Contribution from precipitation (R02-R04) ($^oC\ yr^{-1}$)	Contribution from CO_2 (R01-R03) ($^oC\ yr^{-1}$)	Contribution from LWDR (R02-R05) ($^oC\ yr^{-1}$)
CLM	**0.031**	0.13	**0.114**	**0.016(100%)**	**0.015(92%)**	**-0.002(-12%)**	**0.001(4%)**	-
CoLM	**0.031**	-0.05	**0.058**	**0.010(100%)**	-	-	-	-
ISBA	**0.033**	-0.17	**0.183**	**0.030(100%)**	**0.030(99%)**	**0.001(2%)**	**0.000(-1%)**	-
JULES	**0.034**	**0.31**	**0.189**	**0.017(100%)**	**-0.001(-6%)**	-0.005(-28%)	**0.000(0%)**	**0.005(31%)**
LPJ-GUESS	**0.033**	0.11		**0.026(100%)**	**0.018(67%)**	0.000(-1%)	**-0.001(-5%)**	-
MIROC-ESM	**0.025**	0.44	**0.140**	**0.024(100%)**	-	-	-	-
ORCHIDEE	**0.045**	0.00	**0.201**	**0.030(100%)**	**0.010(34%)**	**0.002(7%)**	**0.001(2%)**	**0.017(56%)**
UVic	**0.031**	0.11		**0.031(100%)**	**0.017(56%)**	**0.000(0%)**	0.000(-1%)	-
UW-VIC	**0.031**	**2.01**	**0.125**	**0.011(100%)**	**0.029(266%)**	-0.005(-47%)	**0.000(0%)**	-

808

Figure legends

Figure 1. The spatial extent of regions defined in this study. Red, green, blue and magenta indicate the regions of boreal North America (BONA), boreal Europe (BOEU) and boreal Asia (BOAS), other permafrost areas (Other), respectively. We only selected BONA, BOEU and BOAS sub-regions for analysis in this study.

Figure 2. Simulated anomaly of annual T_s at 20 cm averaged over boreal regions of each model, during the period of 1960-2000.

Figure 3. Simulated trends of annual T_s at 20 cm averaged over boreal regions and sub-regions of each model, from 1960 to 2000. * indicates significant trend of T_s ($P<0.05$).

Figure 4. Spatial distributions of trends of annual T_s at 20 cm over boreal regions from 1960 to 2000 in (a) CLM, (b) CoLM, (c) ISBA, (d) JULES, (e) LPJ-GUESS, (f) MICRO-ESM, (g) ORCHIDEE, (h) UVic and (i) UW-VIC models. The black dots indicate regions with significant trends of T_s ($P<0.05$). Note that extreme values outside of the range of -0.05 °C yr^{-1} - 0.05 °C yr^{-1} are shown in deepest blue and red in the color bar.

Figure 5. Simulated trends of annual T_s over boreal regions as a function of soil depths (a) 0 - 3 m and (b) 0 - 40 m for the nine models. Note the different total soil depths of the models and negative trends for UW-VIC (~ -0.01 - -0.03 °C yr^{-1}) below 2.3 m are not shown in the plots.

Figure 6. Spatial distributions of difference in trends of annual T_s at 0.2 m and 3 m over boreal regions from 1960 to 2000 in (a) CLM, (b) CoLM, (c) ISBA, (d) JULES, (e) LPJ-GUESS, (f) MICRO-ESM, (g) ORCHIDEE, (h) UVic and (i) UW-VIC models. The black dots indicate statistically significant difference by t-test ($P<0.05$). Note that extreme values outside of the range of -0.005 °C yr^{-1} - 0.005 °C yr^{-1} are shown in deepest

blue and red in the color bar.

Figure 7. Simulated trends of annual T_s at 20 cm and T_a in the climate forcing data across the nine models.

Figure 8. Simulated trends of annual T_s at 20 cm and annual LWDR in the climate forcing data over boreal regions across the seven models which used and provided LWDR in their climate forcing. The black dotted lines indicate the linear regression and 95% confidence interval. The gray dashed line indicates the uncertainty of trend of LWDR in the climate forcing data. The solid blue and orange lines with double arrows indicate FU and SU, respectively. The red solid line with shade area shows the trend of LWDR (0.087 ± 0.023 W m^{-2} yr^{-1}) during the period 1960-2000 from CRUNCEP v5.2 dataset. The purple solid line with shade area shows the trend of LWDR (0.187 ± 0.028 W m^{-2} yr^{-1}) during the period 1960-2000 from WATCH dataset. The right panel shows the prior normal probability density function (PDF) with modeled mean and standard deviation (black solid line) and posteiror normal PDF (red and purple dotted line) with given trend of LWDR from CRUNCEP and WATCH respectively.

Figure 9. Simulated trends of summer T_s at 1 m and loss rate of NSPA over (a) boreal regions, (b) BONA, (c) BOEU and (d) BOAS across the nine models.

Figure 1. The spatial extent of regions defined in this study. Red, green, blue and magenta indicate the regions of boreal North America (BONA), boreal Europe (BOEU) and boreal Asia (BOAS), other permafrost areas (Other), respectively. We only selected BONA, BOEU and BOAS sub-regions for analysis in this study.

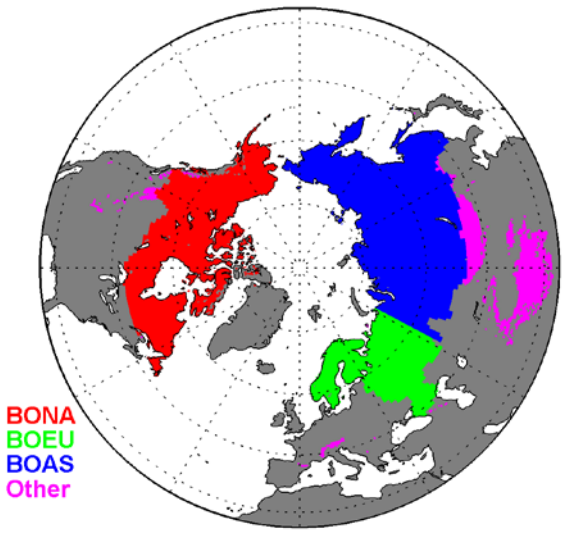

Figure 2. Simulated anomaly of annual T_s at 20 cm averaged over boreal regions of each model, during the period of 1960-2000.

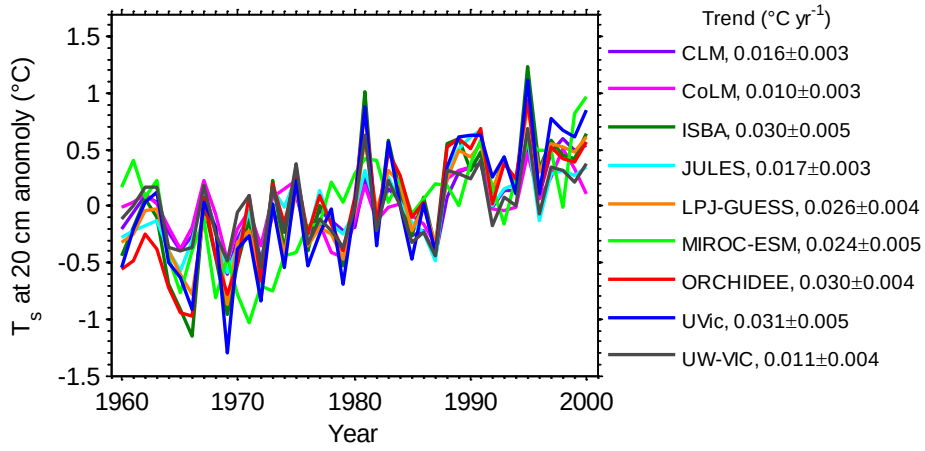

Figure 3. Simulated trends of annual T_s at 20 cm averaged over boreal regions and sub-regions of each model, from 1960 to 2000. * indicates significant trend of T_s ($P<0.05$).

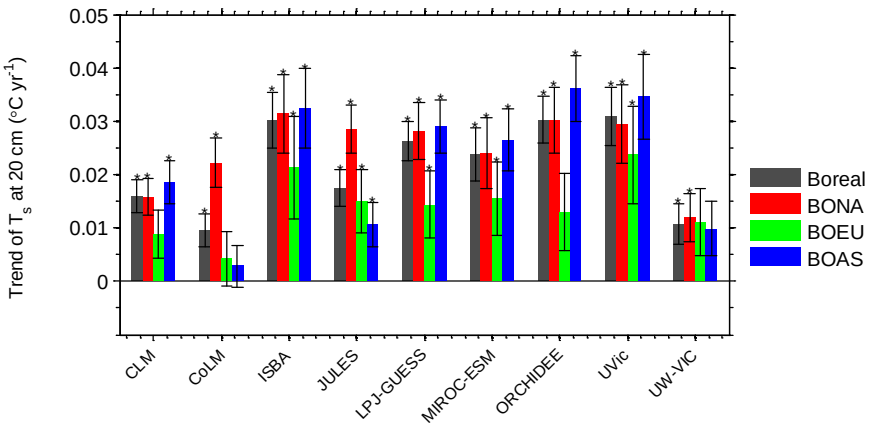

Figure 4. Spatial distributions of trends of annual T_s at 20 cm over boreal regions from 1960 to 2000 in (a) CLM, (b) CoLM, (c) ISBA, (d) JULES, (e) LPJ-GUESS, (f) MICRO-ESM, (g) ORCHIDEE, (h) UVic and (i) UW-VIC models. The black dots indicate regions with significant trends of T_s ($P<0.05$). Note that extreme values outside of the range of -0.05 °C yr^{-1} - 0.05 °C yr^{-1} are shown in deepest blue and red in the color bar.

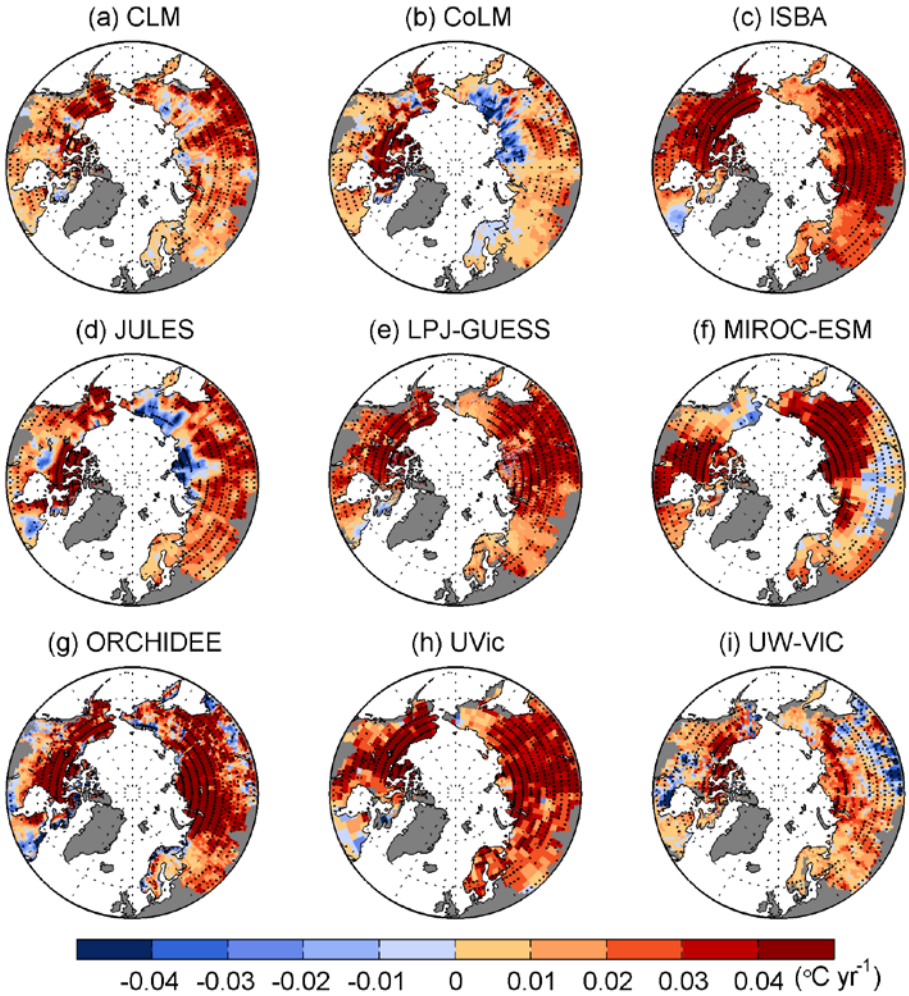

Figure 5. Simulated trends of annual T_s over boreal regions as a function of soil depths (a) 0 - 3 m and (b) 0 - 40 m for the nine models. Note the different total soil depths of the models and negative trends for UW-VIC (~ -0.01 - -0.03 °C yr^{-1}) below 2.3 m are not shown in the plots.

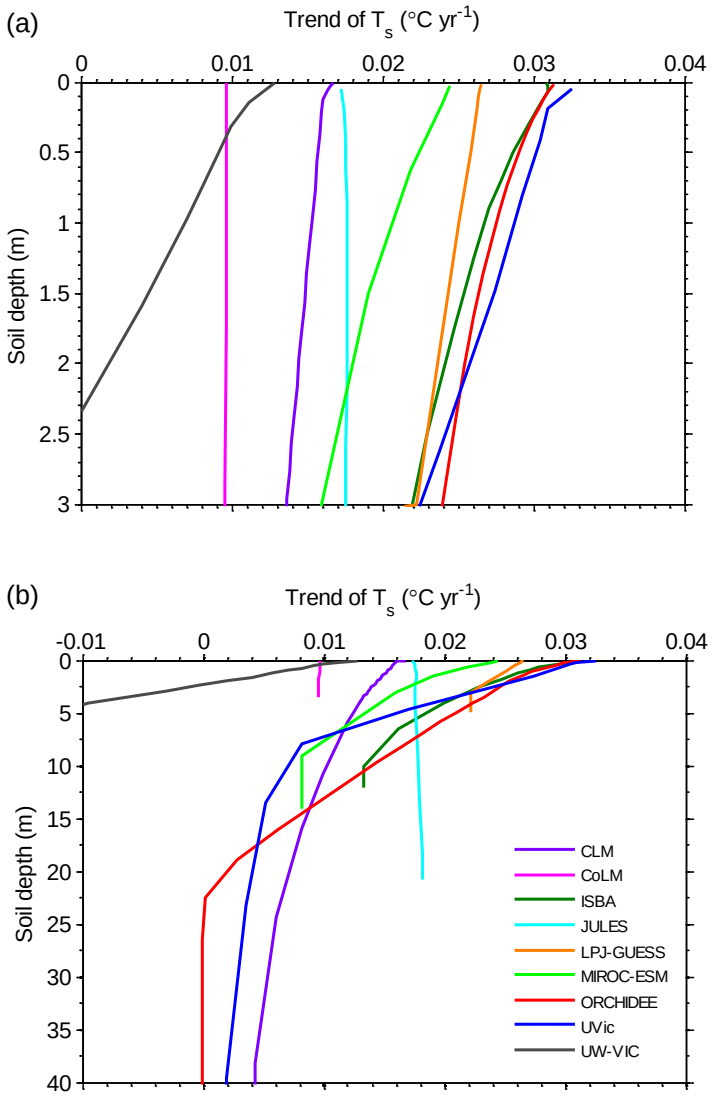

Figure 6. Spatial distributions of difference in trends of annual T_s at 0.2 m and 3 m over boreal regions from 1960 to 2000 in (a) CLM, (b) CoLM, (c) ISBA, (d) JULES, (e) LPJ-GUESS, (f) MICRO-ESM, (g) ORCHIDEE, (h) UVic and (i) UW-VIC models. The black dots indicate statistically significant difference by t-test (P<0.05). Note that extreme values outside of the range of -0.005 °C yr^{-1} - 0.005 °C yr^{-1} are shown in deepest blue and red in the color bar.

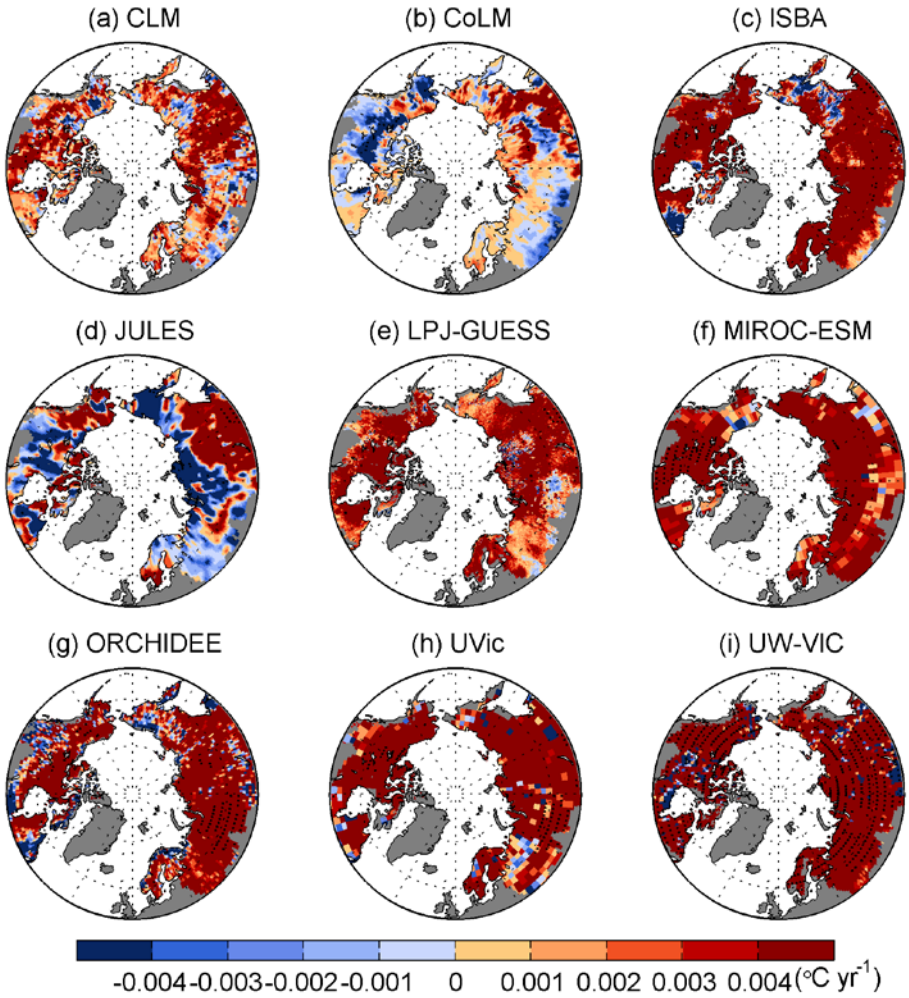

Figure 7. Simulated trends of annual T_s at 20 cm and T_a in the climate forcing data across the nine models.

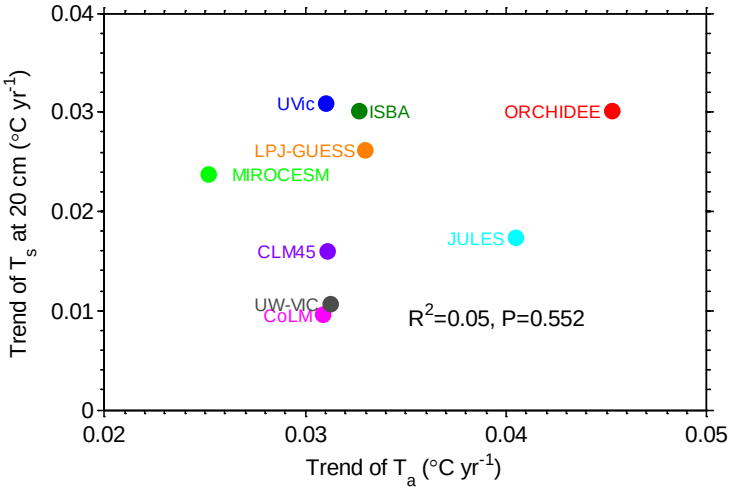

Figure 8. Simulated trends of annual T_s at 20 cm and annual LWDR in the climate forcing data over boreal regions across the seven models which used and provided LWDR in their climate forcing. The black dotted lines indicate the linear regression and 95% confidence interval. The gray dashed line indicates the uncertainty of trend of LWDR in the climate forcing data. The solid blue and orange lines with double arrows indicate FU and SU, respectively. The red solid line with shade area shows the trend of LWDR (0.087 ± 0.023 W m^{-2} yr^{-1}) during the period 1960-2000 from CRUNCEP v5.2 dataset. The purple solid line with shade area shows the trend of LWDR (0.187 ± 0.028 W m^{-2} yr^{-1}) during the period 1960-2000 from WATCH dataset. The right panel shows the prior normal probability density function (PDF) with modeled mean and standard deviation (black solid line) and posteiror normal PDF (red and purple dotted line) with given trend of LWDR from CRUNCEP and WATCH respectively.

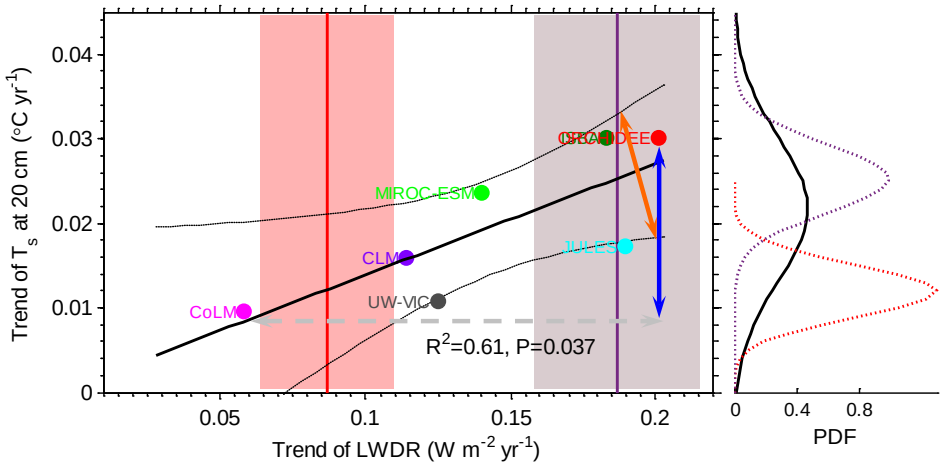

Figure 9. Simulated trends of summer T_s at 1 m and loss rate of NSPA over (a) boreal regions, (b) BONA, (c) BOEU and (d) BOAS across the nine models.