https://doi.org/10.5194/egusphere-2025-389
Preprint. Discussion started: 10 February 2025

Assessing Climate Modeling Uncertainties in the Siberian Frozen Soil Regions by Contrasting CMIP6 and LS3MIP

Zhicheng Luo[1], Duoying Ji[2], and Bodo Ahrens[1]

[1]Institute for Atmospheric and Environmental Sciences, Goethe University Frankfurt, Frankfurt a. M., Germany
[2]College of Global Change and Earth System Science, Beijing Normal University, Beijing, China

Correspondence: Bodo Ahrens (Bodo.Ahrens@iau.uni-frankfurt.de)

Abstract. Climate models and their land components still show pervasive discrepancies in frozen soil simulations. Contrasting the historical runs of seven land-only models of the Land Surface, Snow, and Soil Moisture Model Intercomparison Project (LS3MIP) with their Coupled Model Intercomparison Project Phase 6 (CMIP6) counterparts allows quantifying the contributions of the land surface parameterization scheme and the atmospheric forcing to the discrepancies. The simulation capabilities are assessed using observational data from 152 sites in Siberia and reanalysis data. On average, the 0.2-m soil temperatures in the CMIP6 simulations are 5.4 °C colder than the observations if the simulated soil temperature drops below -5 °C. The LS3MIP simulations are even colder with a bias of -6.7 °C. In the winter months (December, January, and February), the LS3MIP ensemble diversity in 2-m temperature is less than the CMIP6 diversity (0.8°C vs 3.2°C). In contrast, the diversity of winter 0.2-m soil temperatures is larger in the LS3MIP ensemble (5.7°C) than in the CMIP6 ensemble (3.6°C). For permafrost sites, the spatial correlation of the simulations of winter soil temperature against observation is not better than 0.7, and spring/autumn spatial correlations of snow depth is less than 0.75 for the CMIP6 models. The biases of 2-m temperature have a different sign and are amplified in magnitude compared to the biases of the soil temperatures, especially below 0 °C. Four of the climate models and their land components underestimate the snow insulation effect. We conclude that land surface models struggle to well simulate soil temperatures and snow depth under low-temperature conditions. The CMIP6 models tend to compensate for errors in their land component by errors through errors in the atmospheric model component. In shallow snow depth (0 to 0.2 m) cases, all models show between 1 to 8 °C less air-soil temperature difference than in situ data. Therefore, a better representation of surface-soil insulation is essential for improvements in frozen soil land modeling.

1 Introduction

The recent rise in near-surface air temperature is almost twice the global average in the Arctic. Climate simulations show that the Arctic could warm by as much as 7.5°C in 1900-2100 (Lawrence and Slater, 2005; Lawrence et al., 2008). Higher temperatures drive the degradation of permafrost, especially in discontinuous and sporadic permafrost regions. The greatest changes

https://doi.org/10.5194/egusphere-2025-389
Preprint. Discussion started: 10 February 2025

occur in regions where the mean annual air temperatures are around 0 °C (Romanovsky et al., 2007; Åkerman and Johansson, 2008; You et al., 2021). The soil temperature at zero annual amplitude depth of continuous permafrost sites on a global scale warmed for 0.39 ± 0.15 °C during 2007-2016, which is two to three times the warming at discontinuous permafrost sites (Biskaborn et al., 2019; Smith et al., 2022), warning about the potential of more substantial permafrost degradation in the future. Large amounts of soil carbon are emitted into the atmosphere during the degradation of permafrost. Permafrost's abrupt thawing results in 1.5 times more carbon emissions than normal degradation (Turetsky et al., 2019; Miner et al., 2022). The low-oxygen environment further increases the share of methane in such carbon emissions (Koven et al., 2015). The winter contribution to total Arctic methane emissions is predicted to reach 39% (Rößger et al., 2022). Changes in hydrothermal conditions that occur during permafrost thawing can lead to changes in surface vegetation types based on different soil-frozen water contents (Heijmans et al., 2022).

The model parameters used to calculate hydrothermal transport are controlled by soil texture, surface organic matter, frequently varying soil water content, and freeze-thaw conditions, which determine the thermal offset or soil insulation effect. There are differences in the time scales of major physical processes between the soil and the atmosphere, and the soil surface is the window through which the atmosphere interacts with the soil (Beringer et al., 2001; Langer et al., 2011a, b). The presence of solid and liquid water in frozen soil greatly affects the hydrothermal properties of the soil, which is described in various manners by different models (Niu and Yang, 2006; Li et al., 2010).

Snow cover insulates the soil and affects the surface energy balance by changing local albedo and other conditions such as emissivity and roughness. Its effect on soil temperature is spatially heterogeneous, depending on its characteristics, including thickness, density, and duration (Zhang, 2005; Zhang et al., 2018). In northeastern Siberia, the change in snow conditions accounts for more than 50% of soil temperature changes (Park et al., 2014, 2015).

CMIP6, launched by the World Climate Research Program (WCRP), aims to explore various topics related to climate change (Eyring et al., 2016). It is currently the most suitable dataset for evaluating the ability of the latest generation of climate models to simulate frozen soil, as it provides an ensemble of climate models at resolutions fine enough for distinguishing different frozen soil regions. The models have been further developed relative to CMIP5 to varying degrees, including biological and physical processes in frozen soil areas. LS3MIP provides the opportunity to exclude the impact of distinct climate variability produced by different atmospheric models (Van Den Hurk et al., 2016)

In CMIP6 and LS3MIP, the set-up of the land cover/land use scenario and the radiative forcing conditions are based on the same protocol. However, the parameterization schemes of the climate models differ, and this is considered a main source of uncertainty in climate modeling (Deng et al., 2021; de Vrese et al., 2023; Kuma et al., 2023). In addition, the presence of internal climate variability can also lead to differences in uncertainty (Ye, 2021; Rashid, 2021; Schwarzwald and Lenssen, 2022; Jain et al., 2023).

Here, we focus on the Siberian region with frozen soils, which constitute a significant portion of the Eurasian continent's frozen terrain. The potential degradation of permafrost in Siberia could have far-reaching consequences for climate and ecosystems throughout Eurasia and globally. Within this region, the observation dataset provided by the All-Russian Scientific Re-

https://doi.org/10.5194/egusphere-2025-389
Preprint. Discussion started: 10 February 2025

search Institute of Hydrometeorological Information-World Data Center (RIHMI-WDC) (Frauenfeld and Zhang, 2011; Sherstiukov, 2012; Zhang et al., 2018) can be applied, which provides consistent soil temperature measurements at standardized depths and thus can be used as reference in climate model evaluation.

The characteristics of frozen soil surface dynamics are assessed by comparing model outputs with benchmarks, including reanalysis and observational data. We will analyze the discrepancies between the same model in CMIP6 and LS3MIP to quantify the bias and uncertainty present in frozen soil regions, attributing them to land surface models versus those resulting from atmospheric forcings. With identical and more realistic atmospheric conditions, we anticipate that the LS3MIP models will more accurately simulate soil conditions. If these models fail to produce soil variable outputs that align better with observed data than the CMIP6 simulations, it is regarded as an error in the land surface models. We will discuss the variations among different models of LS3MIP and try to establish a connection between model performance and their specific features.

2 Data and Methods

We used the data from climate models, reanalysis, observations, and processing methods of target variables for our analysis. We only included data from 1985 to 2014 in this research, as this period offers the best collection of observation records, and CMIP6 historical experiments are limited to 2014.

2.1 CMIP6 and LS3MIP Simulations

The CMIP6 multi-model ensemble provides historical climate simulations based on the same external forcing (solar radiation, greenhouse gases, aerosols, etc.) (Eyring et al., 2016). Our study uses the *historical* simulations with predefined CO_2 concentrations, which contain distinctive combinations of atmospheric and land models. The *Land-Hist* experiments from LS3MIP are offline land surface simulations with no feedback to the atmosphere and no dynamic forcing from atmospheric models (Van Den Hurk et al., 2016). All the LS3MIP simulations employed the same atmospheric forcing derived from the Global Soil Wetness Project Phase 3 (GSWP3) and the same land surface setup as in the CMIP6 experiments. This enables us to directly compare and assess the impact of errors due to coupling versus deficiencies in the land model on the simulation results in the frozen soil region.

We chose seven climate models/earth system models involved in both projects, incorporating six different land models. The selected models are listed in Table 1. Other climate models, which also participated in both projects, can not be considered as they turned off the freeze option in frozen soil in the CMIP6 version or did not provide data for all our target variables. Hereafter, we refer to the CMIP6 runs from CMIP6 historical simulations as Group C and LS3MIP runs from LS3MIP as Group L. Four variables are collected for this study: 2-meter air temperature (tas), soil temperature in 0.2-meter depth (tsl), snow depth (snd), and precipitation (pr).

Both CMIP6 and LS3MIP data can be accessed at https://aims2.llnl.gov/search/cmip6/.

https://doi.org/10.5194/egusphere-2025-389
Preprint. Discussion started: 10 February 2025

Table 1. Selected CMIP6/LS3MIP experiment pairs, the layering and resolution. For other features and references, see Table 2.

Model Name	**Land Surface Model**	**Total Soil Layers** (max. node depth (m))	**Soil Layers** in Top 3 m	**max. Snow Layers**	**Resolution** (lat×lon)
CESM2	CLM5	25 (42.0)	14	12	0.9°×1.25°
CNRM-CM6.1	Surfex 8.0c	14 (10.0)	11	10	1.4°×1.4°
CNRM-ESM2.1	Surfex 8.0c	14 (10.0)	11	10	1.4°×1.4°
HadGEM3-GC31-LL	JULES-HadGEM3-GL7.1	4 (2.0)	4	3	1.25°×1.875°
IPSL-CM6A-LR	ORCHIDEE v2.0	18 (65.56)	12	3	1.25°×1.875°
MIROC6	MATSIRO6.0	6 (9.0)	5	3	1.4°×1.4°
UKESM1.0-LL	JULES-ES-1.0	4 (2.0)	4	3	1.25°×1.875°

https://doi.org/10.5194/egusphere-2025-389
Preprint. Discussion started: 10 February 2025

Table 2 highlights several attributes of each land model, focusing on their key characteristics related to the surface energy balance and associated processes (Menard et al., 2021). These land models differ in their representation of critical processes related to surface energy balance, snow physics, and soil-vegetation-atmosphere interactions. Albedo is treated either spectrally, accounting for wavelength-dependent variations, or as a time-dependent parameter. Snow conductivity is modeled using density-dependent formulations, empirical power functions, or fixed values. Similarly, snow density is treated dynamically through mechanical compaction or as a constant. The soil bottom boundary is handled with zero-flux assumption, fixed values, or fixed temperature gradients. Models incorporate surface organic matter to varying degrees of complexity. Some focus on its hydrological and thermal impacts, while others emphasize its role in modifying surface albedo and roughness or as a component of the carbon cycle. For instance, models with a "Carbon Cycle Focus" simulate organic matter decomposition and carbon fluxes, whereas those with a "Hydrological Focus" prioritize water retention and flow.

https://doi.org/10.5194/egusphere-2025-389
Preprint. Discussion started: 10 February 2025

Table 2. Features of the land surface models. SA indicates spectrally averaged, and MC indicates mechanical compaction. Surface organic matter indicates how models consider the function of surface organic matter.

Land Model	Albedo	Snow Conductivity	Snow Density	Bottom Boundary Condition	Surface Organic Matter	References
CLM5	Spectral	Density-Dependent	MC	Zero Flux	Hydro-thermodynamic	Van Kampenhout et al. (2017) Lawrence et al. (2019)
Surfex 8.0c	Spectral	Power Function	MC	Fixed	Hydro-thermodynamic	Vionnet et al. (2012) Decharme et al. (2019)
JULES-HadGEM3-GL7.1	Spectral	Power Function	MC	Zero-flux	Carbon Cycle	Clark et al. (2011) Walters et al. (2019) Wiltshire et al. (2020c)
ORCHIDEE v2.0	SA	Quadratic Equation	MC	Fixed Flux	Water/Carbon Cycle	Wang et al. (2013) Bowring et al. (2019)
MATSIRO6.0	Spectral	Fixed	Fixed	Fixed	Hydrology	Takata et al. (2003)
JULES-ES-1.0	Spectral	Power Function	MC	Zero-flux	Carbon Cycle	Sellar et al. (2019)

https://doi.org/10.5194/egusphere-2025-389
Preprint. Discussion started: 10 February 2025

2.2 ERA5-Land Reanalysis

We utilized monthly averaged ERA5-Land reanalysis data from the European Centre for Medium-Range Weather Forecasts, as it can assist our assessment in determining the impact of comparing gridded data to site data. ERA5-Land is a numerical land surface model product forced by atmospheric variables of ERA5, featuring a high horizontal resolution of $0.1° \times 0.1°$ (Muñoz-Sabater et al., 2021). Monthly data provided by ERA5-Land include 2-meter temperature, snow depth, and soil temperature and moisture at four depths (0 to 0.07 m, 0.07 to 0.28 m, 0.28 to 1.0 m, 1.0 to 2.89 m) are provided by ERA5-Land. The dataset can be obtained from https://cds.climate.copernicus.eu/.

2.3 Observational Data

Observational daily data are gathered from 236 meteorology sites by RIHMI-WDC. We filtered the data from 1985 to 2014 based on the quality flag provided in the dataset and employed only sites with a minimum of 330 valid days per year for all four target variables and at least 15 years of valid data. Stations west of $60°\mathrm{E}$, east of $120°\mathrm{E}$, and south of $45°\mathrm{N}$ are excluded to eliminate stations with warmer climates. These criteria are put in place to ensure the accuracy and reliability of the data analyzed. A total of 152 stations have been selected. The data is collected from http://meteo.ru/.

2.4 Data Preprocessing

We interpolate tas, tsl, pr, and snd for all model simulations at the station sites' coordinates by choosing the nearest-neighbor grid cell. This introduces additional uncertainty when comparing modeled data with observed data. Notably, the CMIP6 historical runs can differ in phase from the actual climate phase due to internal climate variability. To minimize this uncertainty source, we consider only 30-year averaged data for evaluations.

2.5 Evaluation Metrics

To evaluate models, we quantify their ability to simulate a reasonable climate mean state and internal climate variability.

The variability of target variables is quantified using the Inter-Quartile Range (IQR), which measures the spread of the simulations. If the model over- or underestimates the observed IQR_o, the model over- or underestimates climate variability, respectively. Thus, we use the relative spread

$$RS_{m,i} = \frac{\mathrm{IQR}_{m,i}}{\mathrm{IQR}_{o,i}} \tag{1}$$

with m indicating the model, o the observation, and i the target variable.

The central climate state of the models and the observations is quantified by the median (med). We use the standardised model medians $\mathrm{med}_{m,i}$, naming it relative bias RB,

$$RB_{m,i} = \frac{\mathrm{med}_{m,i} - \mathrm{med}_{o,i}}{\mathrm{IQR}_{o,i}} \tag{2}$$

as a measure for systematic model errors.

https://doi.org/10.5194/egusphere-2025-389
Preprint. Discussion started: 10 February 2025

RS assesses a model's capability to reproduce the observed variability. This is essential for determining a model's reliability in simulating dynamic climate systems. RB, on the other hand, addresses systematic deviations in the central tendency of the data. It is standardized using observed variability (IQR_o), which facilitates comparisons of biases across different variables. RB emphasizes whether systematic errors are pronounced relative to natural variability (informs if the error is smaller ($RB_{m,i} < 1$) or larger (> 1) compared to the observed climate variability), helping prioritize improvements in model development.

For qualification of the error heterogeneity of the model ensembles at the sites' locations, we define the ensemble means of the models' 30-year median biases as follows

$$EB_{i,s} = \frac{\sum_{m=1}^{M}(\text{med}_{m,i,s} - \text{med}_{o,i,s})}{M} \tag{3}$$

and the ensemble-spreads of median biases $ES_{i,s}$, which are calculated as standard deviations.

The ensemble mean biases $EB_{i,s}$ and spreads are calculated for both the CMIP6 (Group C) and LS3MIP (Group L) model ensembles with $M = 7$ members each. The analyses are done in different seasons to distinguish the impact of different freeze/thaw periods on ensemble performance.

3 Results and Discussions

3.1 Winter 2-m Temperature in Target Area

Fig. 1 is based on average winter (DJF from 1985 to 2014) 2-meter air temperatures (tas) from ERA5-Land, and the symbols correspond to the matching observation data from RIHMI-WDC. As shown in the map, winter-time tas in the target area is colder in the northeast and warmer in the southwest. Within the area 50 °E to 185 °E, north of 45 °N, the average DJF tas is generally below 0 °C. The region with less than -25 °C has a large overlap with the continuous permafrost region (Brown et al., 1997; Obu et al., 2019).

We aim to assess the models' performance under varying climate conditions to determine whether simulation uncertainties increase at lower temperatures or remain similar. To distinguish different climate regimes, a practical approach to categorize the stations is by using their average DJF 2-m air temperature following Wang et al. (2016). By focusing on winter temperatures, we can further link the outcomes to the insulating effect of snow. The temperature categories are listed in the legend of Fig. 1. There are 3 sites with average tas warmer than -5 °C, 25 sites between -15 and -5 °C, 79 sites between -25 and -15 °C, and 40 stations with tas below -25 °C. Besides, five sites are identified as permafrost ('perma') sites using the method introduced by Lawrence and Slater (2005) (soil layer temperatures continuously below 0 °C for at least two years) and labeled by stars. All 'perma' sites have average winter tas values less than -30 °C.

The spatial distribution of average winter tas conditions in ERA5-Land exhibits high consistency with the observations. For example, the site observations colder than -25 °C are also colder than -25 °C in ERA5-Land, including observations in isolated locations south of 55 °N. This proves that ERA5-Land can be a solid benchmark that supports observation as gridded data.

https://doi.org/10.5194/egusphere-2025-389
Preprint. Discussion started: 10 February 2025

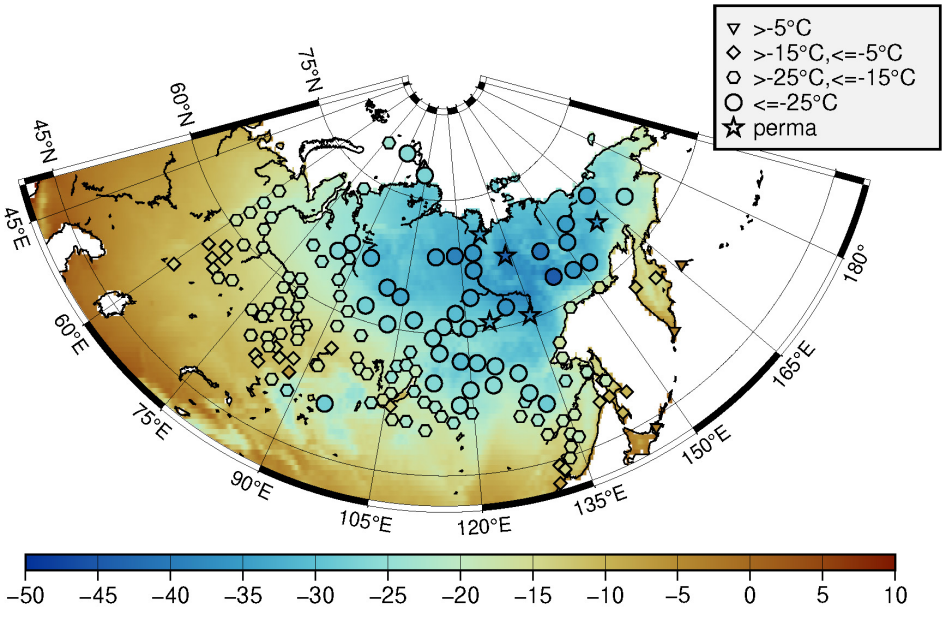

Figure 1. Mean near-surface air temperature (in °C) as given by ERA5-Land and the observational sites for winter (DJF) in 1985-2014. Symbols indicate the climate state at the observational sites (see legend).

https://doi.org/10.5194/egusphere-2025-389
Preprint. Discussion started: 10 February 2025

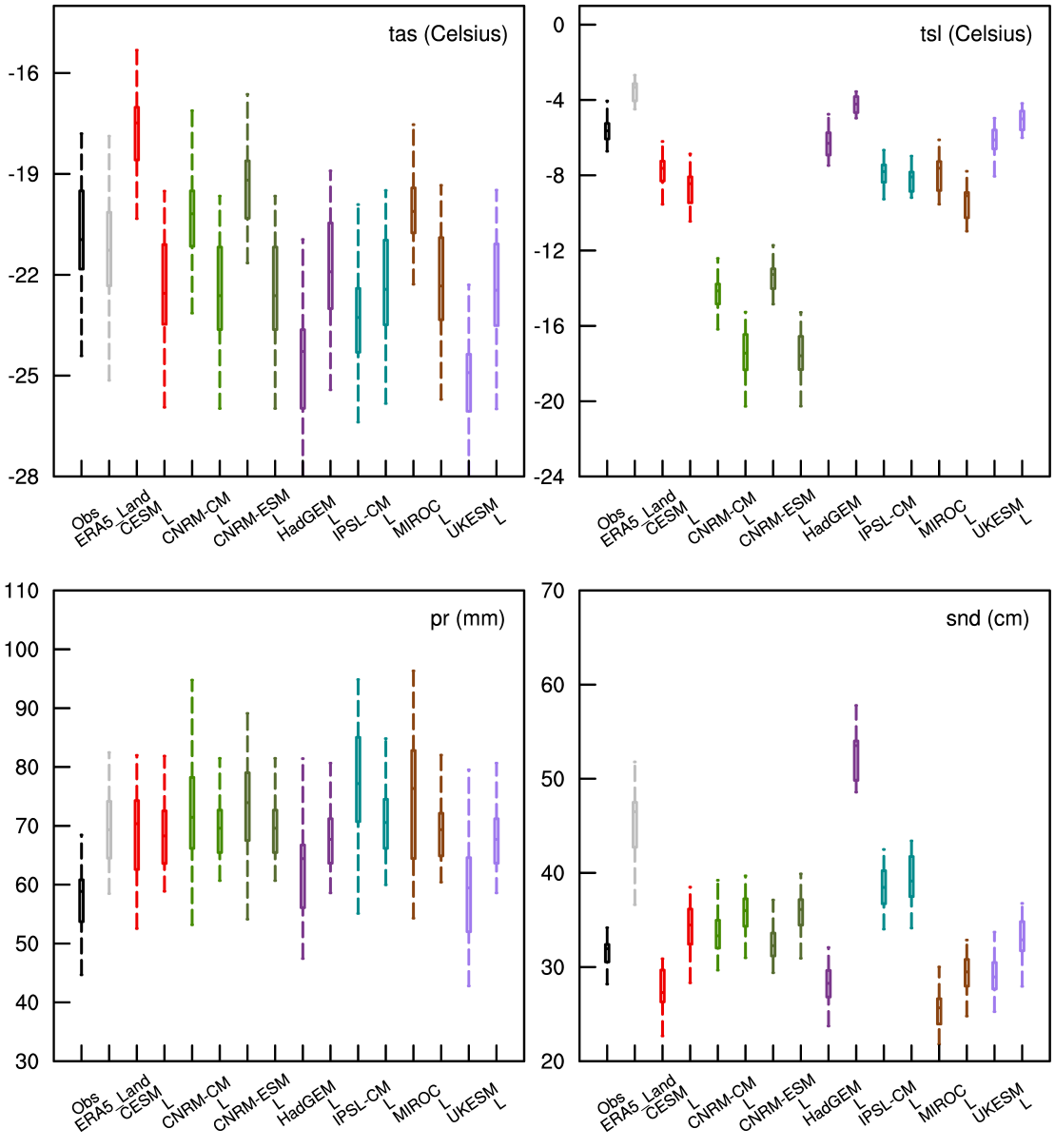

Figure 2. Sites-averaged DJF–climates (1985–2014) of hydrothermal variables as observed and simulated. Names at the abscissas and the colors indicate the different data sources. Observations (Obs) and ERA5-Land are in situ and reanalysis data, respectively; model names indicate CMIP6 model output, and Ls indicate the corresponding LS3MIP simulations (see Table 1). The boxes represent the medians, first and third quartiles; the $\pm$ 1.5$\times$IQR or the maximum and minimum values, if within the former range, are taken as the whiskers' length.

3.2 Model climatologies

https://doi.org/10.5194/egusphere-2025-389
Preprint. Discussion started: 10 February 2025

Figure 2 shows the site-averaged winter, DJF climatologies from 1985 to 2014 of the four target variables for the different datasets. The figure shows that ERA5-Land's tas and tsl climatologies are closer to the Obs climatologies than most of the models', where this is not the case for pr and snd.

The LS3MIP's tas and pr are interpolated from the same forcing dataset. Still, Fig. 2 shows slight differences between different land models because of interpolation uncertainties using different model grids with different setups. This illustrates how carefully a comparison of coarse-grid model output against point-like station data has to be interpreted. Additionally, the LS3MIPs' tas climatologies are systematically colder by more than 1 °C than Obs and ERA5, and their pr climatologies align better with ERA5-Land than with Obs (which has on average about 15% smaller values).

The CMIP6 models' tas climatologies scatter substantially. CESM2's tas median is shifted by more than $+3$ °C against Obs, while tas medians of HadGEM3-GC31-LL and UKESM1.0-LL are shifted by more than -3 °C.

Variations of tsl are controlled by changes in overlying tas and by geothermal heat from below (Smith and Riseborough, 2002). Most models' tsl climatologies are colder than observed (with the CMIP6 and LS3MIP climatologies of CNRM-CM6.1 and CNRM-ESM2.1 more than -8 °C). Models belonging to the same family (the two CNRM models and the models HadGEM and UKESM, see Tab. 1) exhibit similarities in their ability to simulate tsl. As to be expected, the temporal variability, quantified by the box plots' IQR, is smaller in tsl than in tas.

In the other seasons, the diversity in tas and tsl climatologies, respectively, is smaller than in DJF (not shown). In June, July, and August (JJA), the models' different tas variabilities are clearly reflected in tsl (with most sites lacking insulating snow).

In both CMIP6 and LS3MIP runs, DJF pr values are higher than in observations by typically 10 mm and thus closer to ERA5-Land. The diversity in the models' median values is within about ± 10 mm. Some CMIP6 models show up to 1.5 times greater interannual variability in DJF climatology compared to observation, suggesting an overestimation of the year-to-year fluctuations in winter climate. The observed snd of about 30 cm is overestimated by 70% in the HadGEM3-GC31-LL LS3MIP simulation and by 50% in ERA5-Land. All the other model medians are within ± 10 cm from Obs but with larger temporal variability.

3.2.1 Relative Spread and Relative Bias

This subsection compares the relative spread (RS) and relative biases (RB) of ERA5-Land and the different models compared to Obs for all four target variables and all seasons. The RS assesses the sites-averaged temporal climate variability, and RB is the shift of the climatologies.

https://doi.org/10.5194/egusphere-2025-389
Preprint. Discussion started: 10 February 2025

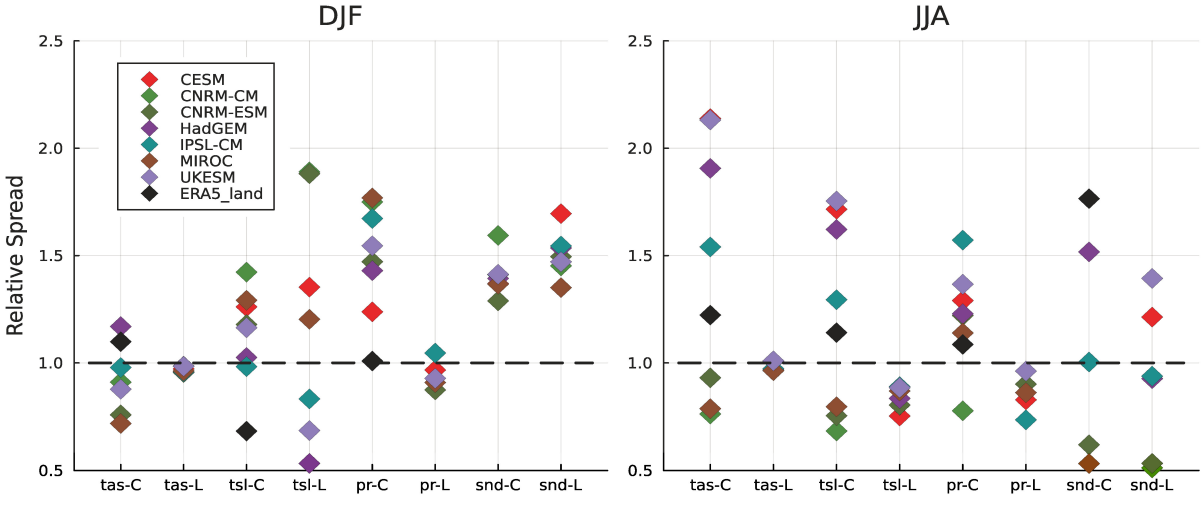

Figure 3. Relative spread (RS) of the sites-averaged climates (1985–2014) of the four variables in both ensembles and in ERA5-Land with reference Obs for all seasons. The colors indicate the models, and the abscissas show the variables and ensembles (C indicates CMIP6, and L indicates LS3MIP runs).

https://doi.org/10.5194/egusphere-2025-389
Preprint. Discussion started: 10 February 2025

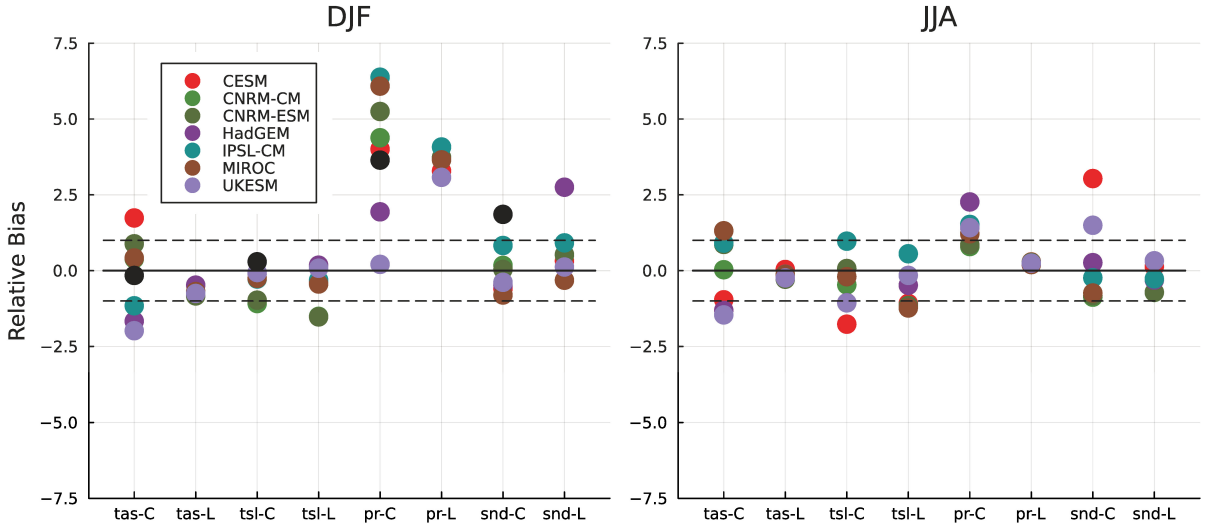

Figure 4. Same as in Fig. 3, but for relative biases (RB). The shaded areas (from -1 to 1) indicate the range of absolute median differences smaller than the observation's IQR.

Fig. 3 shows that most CMIP6 models underestimate tas DJF climate variability, besides in JJA with four models (CESM2, HadGEM3-GC31-LL, IPSL-CM6A-LR, UKESM1.0-LL) more than 1.5 times the variability of Obs. Most CMIP6 models overestimate tsl climate variability, but the LS3MIP models show a large diversity in winter (from 0.5 to almost 2, and it's a more extensive spread than the CMIP6 models despite their larger diversity in tas) and systematically underestimate variability in summer. For DJF, MAM, and SON, the tas relative spreads for all CMIP6 models are within the range of 0.5 to 1.2. HadGEM3-GC31-LL, IPSL-CM6A-LR, and UKESM1.0-LL exceed 1.5 in JJA. Larger variability differences in simulated tsl are observed for LS3MIP simulations even though their tas share an identical relative spread.

The winter pr relative spread of Group C is all higher than 1.2. The pr in Group C exhibits more extensive group diversity than in Group L. In spring and autumn, most climate models simulated good inter-annual variability of pr, with less than a 20% difference to observation (not shown). Both groups and ERA5-Land overestimate the spread in winter snd with RS-values larger than 1.2. The RS-values scatter between 0.5 and about 2 in the other seasons.

https://doi.org/10.5194/egusphere-2025-389
Preprint. Discussion started: 10 February 2025

Fig. 4 shows the relative bias RB of all variables calculated in all seasons. The shaded zone indicates a one-time IQR of the observed value. If a model's absolute RB is less than one, the model's bias is considered relatively small. For example, the winter RB of MIROC is within the shaded area for tas, tsl, and snd, but not for pr (note that Obs' pr is smaller than all models', including ERA5-Land's). Almost all CMIP6 and LS3MIP models have a positive pr-bias but a smaller relative and non-systematic snd-bias in winter. The CMIP6 runs of HadGEM3-GC31-LL and UKESM1.0-LL underestimate the values of tas in all seasons. For tsl, the relative bias in DJF for LS3MIP CNRM-CM6.1 and CNRM-ESM2.1 exceeds the observed IQR-range while other models are all within an acceptable range in both groups. Negative RB in CMIP6 and LS3MIP tsl is the largest in the transition seasons, which also shows a large spread in snd RB-values. With better RS and RB of tas and pr, models still have diverse abilities simulating tsl and snd.

3.2.2 Spatial Heterogeneity

https://doi.org/10.5194/egusphere-2025-389
Preprint. Discussion started: 10 February 2025

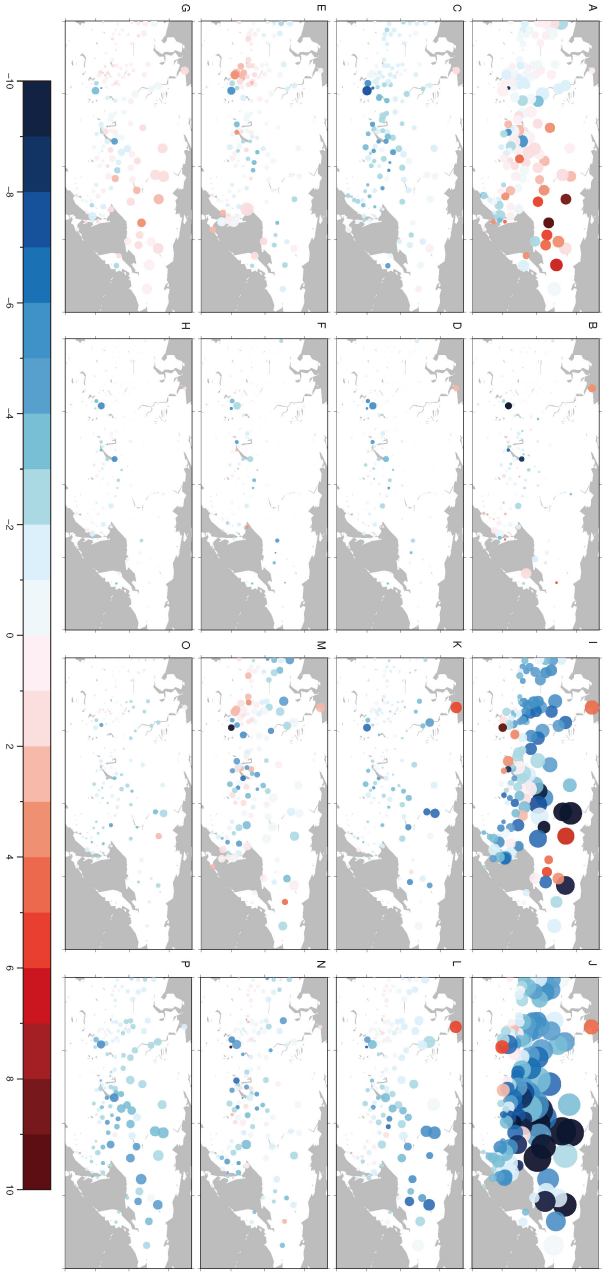

Figure 5. Mean biases and their standard deviations (°C) of the models' tas and tsl at Obs sites. Panels A–H and I–P are for tas and tsl, respectively. The first and third columns are the results of CMIP6 models, and the second and fourth columns are the results of LS3MIP models. The first to fourth rows show the DJF, MAM, JJA, and SON seasons, respectively. Colors indicate the bias value, while the circle radii represent biases' standard deviations, i.e., the ensemble spread.

https://doi.org/10.5194/egusphere-2025-389
Preprint. Discussion started: 10 February 2025

Figure 5 shows the model ensemble mean biases (EB) and the ensemble spread of biases (ES) for tas and tsl at the Obs sites. EB and bias diversity are largest in winter and larger in tsl than in tas. In northeastern Siberia's winter, the CMIP6 models overestimate tas on average by more than 4 °C but with a large model bias standard deviation of 3.2 °C.

In the other seasons, EB and ES are distinctly smaller, with most CMIP6 runs' tas slightly too cold in spring at most stations and slightly too warm in autumn at northeastern Siberia. Differences in grid cell scale among models can lead to biases in the tas state over the grid. The tas EB of the forced models are small at most sites (exceptions are near water bodies, e.g., Lake Baikal and sea coasts, which indicate interpolation artifacts because of selected grids). The magnitude of Group L's tas ES is almost unchanged among seasons, implying its bias mainly relates to the geographical characteristics of stations' sites.

The tsl EB and ES are larger in magnitude than for tas, especially in winter with spatial EB of -3.0 °C and -4.0 °C for CMIP6 and LS3MIP runs, respectively. Additionally, the LS3MIP runs are notably more diverse than the CMIP6 runs throughout all seasons but in summer. In winter, the ensemble bias spreads are 3.6 °C for the CMIP6 and as large as 5.7 °C for the LS3MIP runs. Given the smaller tas EB and the larger tsl EB in the LS3MIP than in the CMIP6 simulations, the modeling diversity introduced by the land surface models is large and partly compensated for by atmospheric diversity in the CMIP6 simulations.

3.3 Permafrost Region

https://doi.org/10.5194/egusphere-2025-389
Preprint. Discussion started: 10 February 2025

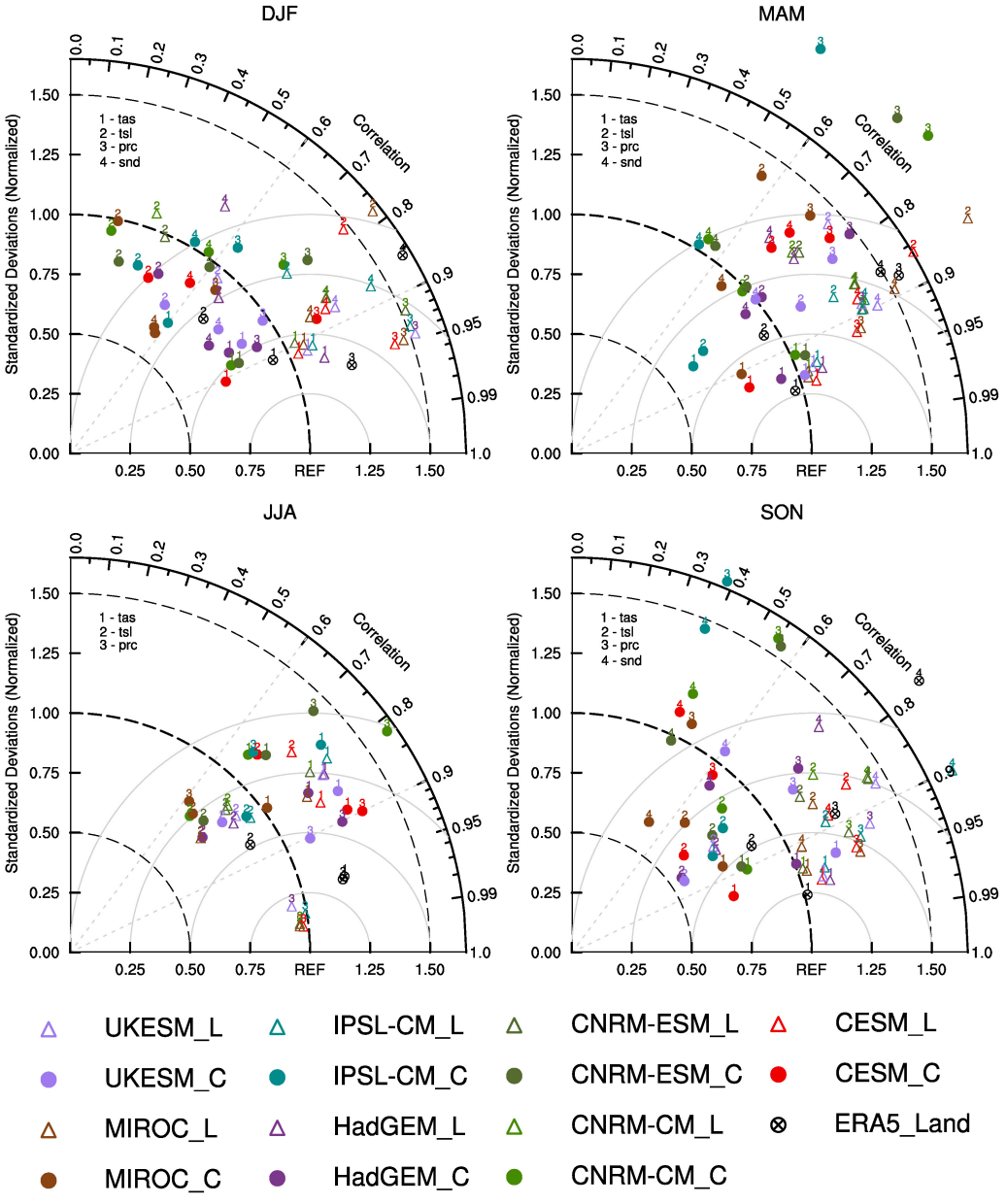

Figure 6. Taylor diagrams indicating the spatial correlation, standard deviation, and unbiased root-mean-square (RMS) difference (grey circles) for the four seasons of the simulated variables tas, tsl, pr, and snd (labeled by numbers) against observations at permafrost sites (here: sites with mean winter $tas < -25°C$). Normalization is applied using the standard deviations of the observations. Colors indicate different models and black crossed circles ERA5-Land. Triangles show the results of the LS3MIP runs and solid circles of the CMIP6 runs. Snow depth is not shown in JJA.

https://doi.org/10.5194/egusphere-2025-389
Preprint. Discussion started: 10 February 2025

Here, we show the models' results at the 45 observational sites with average winter tas below -25 °C, at which we find the largest ES (see Fig. 5). This research focuses on the shallow soil response to atmospheric forcing, explicitly targeting a depth of 20 cm. Using the authentic definition of permafrost, which requires more than two consecutive years of temperatures below 0 °C, only five of the sites qualify at this depth. However, under similar climatic conditions (where winter tas are below -25 °C), we observe 40 more sites distributed within permafrost areas as defined in Brown et al. (1997), indicating that permafrost may exist at deeper soil layers. Incorporating these sites for statistical analysis will enhance the robustness of the findings.

Figure 6 shows spatial Taylor diagrams (Taylor, 2001), which compare the 30-year-average of simulated data to the observations in all seasons at all sites in the permafrost region. Almost all CMIP6 and LS3MIP models struggled with simulating soil temperature, tsl, with large RMS differences and low correlations during DJF. Correlations are generally better than 0.5 for all variables in the other seasons. The tas correlations are better than ca. 0.8 for all models (except for MIROC and IPSL-CM of Group C in DJF), which indicates well-simulated spatial patterns by the CMIP6 models.

In the spring and autumn (MAM, SON), the CMIP6 models tend to overestimate precipitation variability (figure not shown). Overall, the Group L models and ERA5-Land tend to perform only slightly better than the Group C models in tsl and MAM snd. In SON, there is a more considerable difference in correlation performance of snd than in MAM, less than 0.6 for the Group C models and higher than 0.7 for the Group L models. Meanwhile, the correlations of pr are similarly low in both seasons. This indicates that the accumulation of snd in climate models is better related to pr condition than melting.

In JJA, tas are less correlated with observation in JJA than in MAM and SON, while tsl maintains the similar accuracy of tas simulations. The normalized RMS differences remain mainly within $\pm$ 0.25, while correlation coefficients range from 0.6 to 0.9.

3.4 Climate Dependency of Modeled Temperatures

https://doi.org/10.5194/egusphere-2025-389
Preprint. Discussion started: 10 February 2025

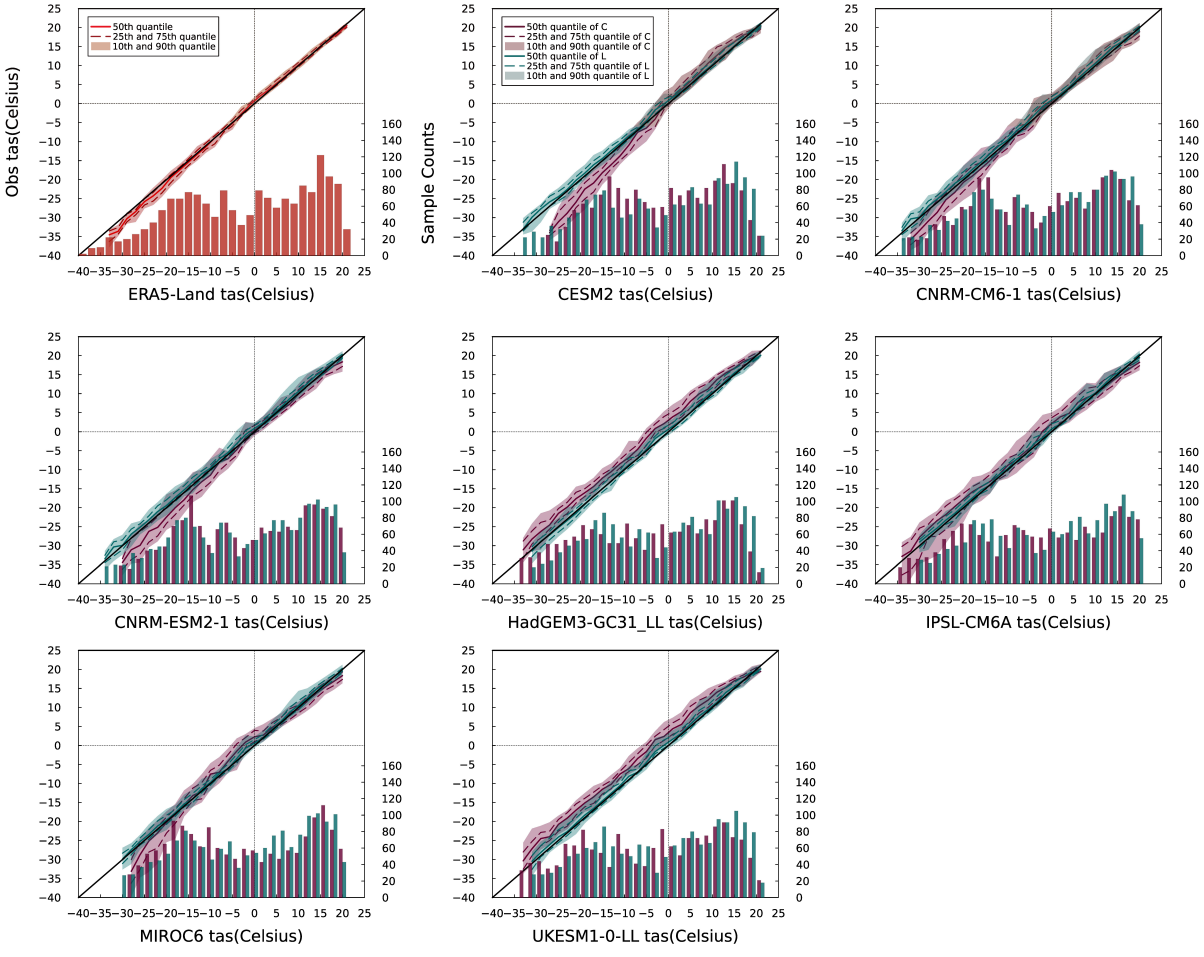

Figure 7. Quantile-Quantile (QQ) and histogram plots for tas (Wilks, 2019). The plots show 30-year mean monthly tas data of models at all sites and their observational values. The model simulation outputs are binned at 2 °C intervals. The colored solid and dash curves are the median and 1st/3rd quartiles of the corresponding observations for all data points in the temperature interval, and the shaded area is the inter-decide range. The histograms represent the sample size within each temperature interval, and temperature intervals with sample sizes smaller than 20 are excluded from the Q-Q plots. The further away the data is from the diagonal line, the larger the model's simulation bias is at certain temperature states. A higher vertical quantile range indicates more inconsistency with observation under identical temperature conditions.

https://doi.org/10.5194/egusphere-2025-389
Preprint. Discussion started: 10 February 2025

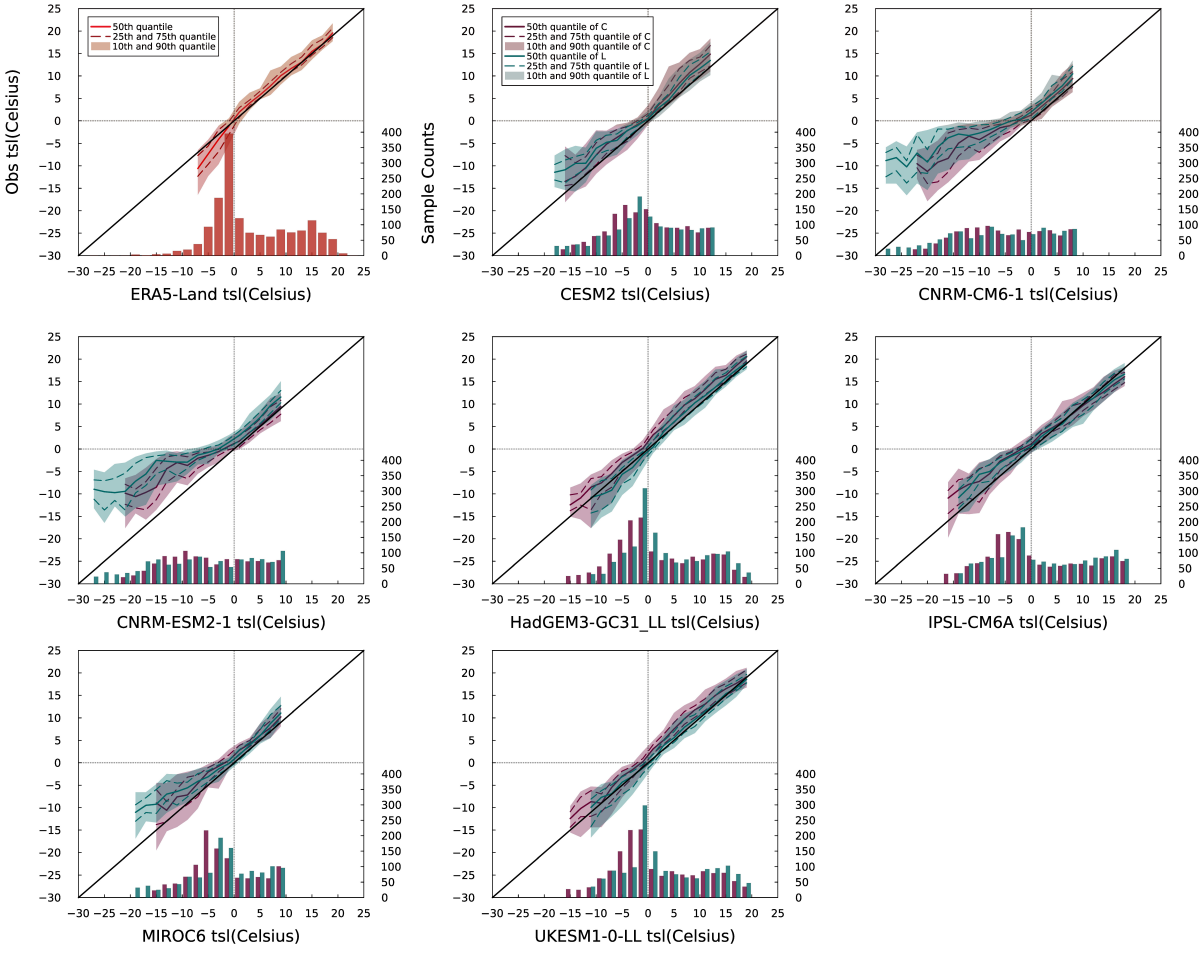

Figure 8. Same as Fig. 7, but for tsl.

https://doi.org/10.5194/egusphere-2025-389
Preprint. Discussion started: 10 February 2025

Since larger modeled biases and discrepancies are found in permafrost regions, we want to further evaluate how modeled errors depend on the temperature state. So, we categorize all the model outputs of tas and tsl by their values and then compare the relationship between model outputs and corresponding Obs. For each model, the monthly data of every site is an average of over 30 years. Then, we look for the corresponding value of the same site and month with valid observation. The mean value, quartiles, and deciles are calculated for each interval.

As shown in Fig. 7, Group L models simulate similar histograms as observation and small quantile spreads across all temperature ranges. However, a slight cold bias below -30 °C exists in Group L. Conversely, ERA5-Land has almost no bias when tas is above -20 °C, but has a warm bias below -20 °C. Group C's simulation results exhibit larger errors, with CESM2, CNRM-CM6.1, and CNRM-ESM2.1 showing warm bias when tas is below 0 °C, and HadGEM3-GC31-LL and UKESM1.0-LL displaying a comprehensive cold bias of 5 °C. In most temperature states, IPSL-CM6A-LR shows a slight cold bias of less than 5 °C. MIROC6 exhibits a cold tas bias from -10 to 2 °C and a warm tas bias that gradually increases below -20 °C. A common characteristic among the Group C models in tas simulation is that a larger vertical quantile range appears when tas is below -15 °C, emphasizing the requirement to improve the accuracy under such surface air temperature conditions.

Larger vertical quantile ranges appear in tsl results than in tas. The tas distribution of ERA5-Land shows two main peaks (Fig. 7), one around 15 °C and the other around -15 °C. However, the tsl of ERA5-Land has one major peak (Fig. 8), which is within -5 to 0 °C. The ERA5-Land vertical quantile range is smaller and more unbiased above 0 °C. Still, it shows a warming bias and increased vertical quantile range below 0 °C. Cao et al. (2022) discussed an unreasonable warm tsl bias that is possibly due to the overestimation of permafrost snow depth in ERA5-Land. Similar warming of tsl due to excessive snd occurs in the LS3MIP simulation of HadGEM3-GC31-LL. CNRM-CM6.1 and CNRM-ESM2.1 are the only ones that do not show a peak in soil temperature samples near 0 °C. The tsl in CNRM-CM6.1 and CNRM-ESM2.1 are more likely distributed and show a strong cold bias from 2 to 15 °C. CESM2 and MIROC6 have an overall cold bias, -5 °C and -6 °C at -10 °C of observed temperature, respectively. The Group L models have varying warm bias below 0 °C tsl, indicating that the tsl simulated by these models have a larger bias.

There are notable discrepancies between the two ensembles simulating cold temperature extremes. For example, the minimum tas in the CMIP6 run for CESM2 is 6 °C higher than that of the LS3MIP run, whereas the tas for CMIP6 IPSL-CM6A-LR is 4 °C lower than that of the land-only run. Such tas differences in the lower extreme directly impact that of tsl, creating a gap of 2 to 6 °C in the lowest value between two ensembles. For models (CESM2, CNRM-CM6.1, CNRM-ESM2.1, and MIROC6) underestimate tsl while exhibiting warm biases of tas, one conceivable explanation is that they underestimate the snow insulation effect(in line with results by (Dutch et al., 2022)) and surface insulation effect (the thermal offset, especially if tas is below -15 °C), allowing a too large energy loss from the soil to the atmosphere.

There is an excessively low tsl shown in Fig.8, possibly due to insufficient geothermal (functions as upward energy flux from the bottom of soil columns). As the decrease in tas has a limited influence on the tsl through high snd, the main source of error is likely from the other side of energy transportation (thermal conditions in the bottom of the soil column). The melting/accumulating snow in MAM and SON and the prevailing phase change processes in the soil are conceivably the major reason for higher tsl inter-annual variability in these seasons.

https://doi.org/10.5194/egusphere-2025-389
Preprint. Discussion started: 10 February 2025

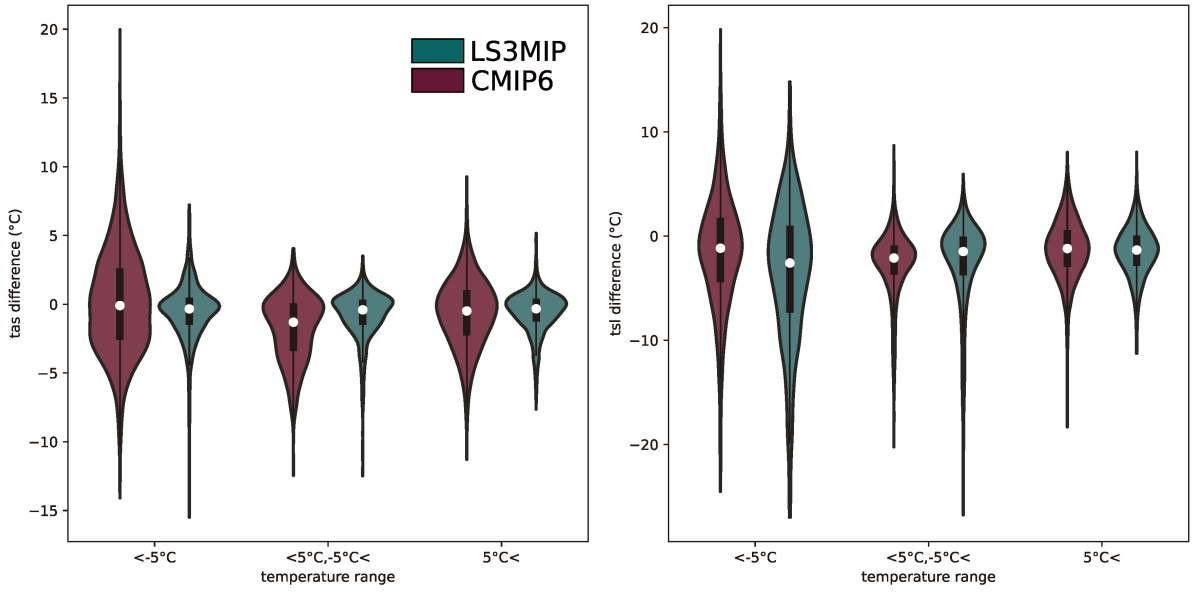

Figure 9. Difference between model tas (left), tsl (right) and corresponding observations. The x-axis represents different temperature intervals, and the y-axis is the 30-year average temperature difference (T_{model} - T_{obs}). Differences are categorized into three sets according to the 30-year average temperatures of every month at the observation sites: below -5 °C (Set Frozen), -5 °C to 5 °C (Set Intermediate), and above 5 °C (Set Warm).

https://doi.org/10.5194/egusphere-2025-389
Preprint. Discussion started: 10 February 2025

In addition, we categorized the model output by the freeze/thaw state of observation. This enables us to compare the performance of two groups in different temperature states. In order to avoid assessment errors caused by different sample sizes, we discuss the overall uncertainty exhibited by the model in the thawed state, the freeze-thaw transition state, and the frozen state, using the boundaries of -5 °C and 5 °C, as the phase change process occurs most frequently between the boundaries.

We subtract the 30-year average of monthly observational data from all stations with the corresponding simulated values and then sort them into observed temperature intervals. Results are shown in Fig. 9. Set Frozen and Set Warm tas data sample sizes are more than twice as large as in Set Intermediate. Set Frozen has the largest standard deviations, 3.84 °C in CMIP6 runs, and 2.14 °C in LS3MIP runs. The mean values of CMIP6 runs are rather divided, with -1.76 °C and 0.11 °C in Set Intermediate and Set Warm, respectively. In contrast, the difference in LS3MIP runs is negligible, suggesting that the climate model will likely have a cold and small deviation under these temperature states.

The tsl samples are mainly concentrated in Set Intermediate and Set Warm. In contrast, there are also higher standard deviations in Set Frozen, 5.39 °C and 6.73 °C for CMIP6 and LS3MIP runs, respectively. The standard deviations of Set Frozen and Set Intermediate in LS3MIP runs are higher than those in CMIP6 runs by 1.33 °C and 1.12 °C, respectively. This again indicates that the land model alone caused higher variability. The mean and minimum value of tsl bias is much lower than that of tas bias, and this negative bias is shown in all Sets of both groups, suggesting that the land models tend to simulate tsl colder than it should be. Better accuracy of tas of Group L in Set Frozen and Set Intermediate does not lead to better tsl results for the models. And tsl below -5 °C shows higher overall variability in Group L than in Group C.

3.5 Snow Insulation

https://doi.org/10.5194/egusphere-2025-389
Preprint. Discussion started: 10 February 2025

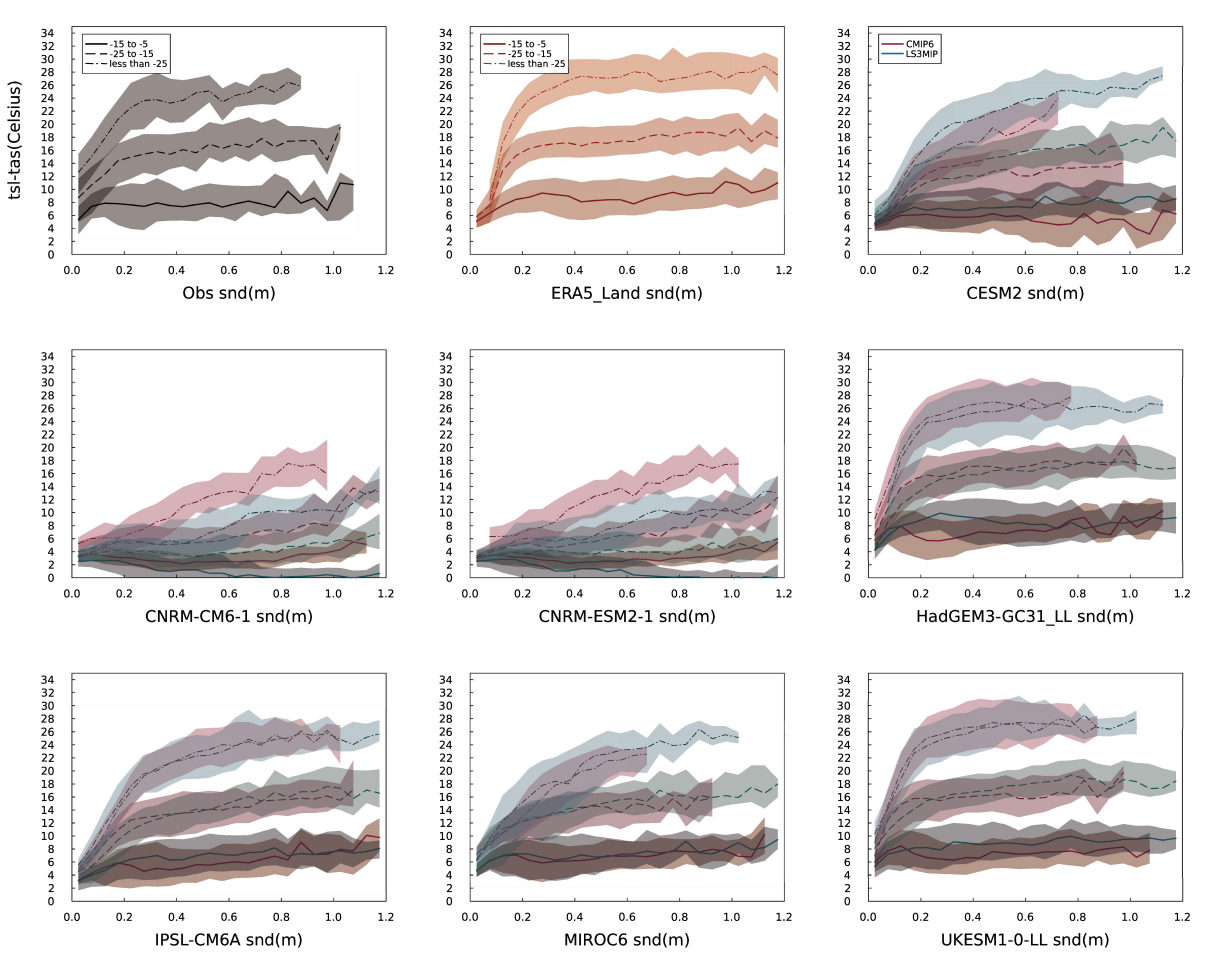

Figure 10. Snow depth and soil-surface temperature differences. The temperature categorizing method of Wang et al. (2016) and encompasses all seasons. The plots use different line styles and color schemes to differentiate between the various temperature categories and ensembles. Our sampling technique involves binning data at 0.05 m intervals, with a coverage of 0–1.2 m *snd*, and each interval contains a minimum of ten samples for accuracy. The plot curves represent the median, with the shaded areas indicating the 25th–75th percentile.

https://doi.org/10.5194/egusphere-2025-389
Preprint. Discussion started: 10 February 2025

To better understand the insulating effect of snow in the models, we refer to the method applied in previous model evaluations (Wang et al., 2016; Burke et al., 2020). We collected the monthly observational data samples which contain valid tas, tsl and snd, characterize them by three different tas intervals (-15 to -5 °C, -25 to -15 °C and less than -25 °C). Analyzing the relationship between temperature difference and snow depth allows us to understand each model's snow insulation effect under different tas and snd conditions.

According to the observation shown in Fig. 10, two key factors should be considered regarding the snow insulation effect. Firstly, even when snd is very thin, the insulating effect of the soil surface layer itself is still influential. Its intensity can vary from 3 °C to 15 °C according to tas. Secondly, the impact of snow on ΔT ($tsl - tas$) has the highest diversity when the snow depth is shallow; under sufficiently thick snow, the tsl gradually convergences near 0 °C and is primarily impacted by tas in a limited manner. For temperatures ranging from -15 to -5 °C, this inflection point is reached when snd is about 0.1 m high. For colder temperatures, it is reached at depths around 0.2 m to 0.3 m. However, snow shields most of the impacts from the overlying tas in thicker snow conditions, as shown by a strong relationship between ΔT and tas when snd is higher than 0.3 m. The insulation effect of snow becomes stronger at lower tas. The median value of maximum ΔT for the -5 °C to -15 °C category is approximately 7.5 °C, while for the -15 °C to -25 °C category, it is 15 °C, and it reaches 24 °C when tas is below -25 °C. When snd exceeds 0.3 m, tsl drops by 1 °C for every 4 °C to 5 °C decrease in tas, taking into account the range of interquartile variations.

Only HadGEM3-GC31-LL, IPSL-CM6A-LR, and UKESM1.0-LL exhibit ΔT curves which are similar to observation. Despite the generally good performance, HadGEM3-GC31-LL and UKESM1.0-LL have deficiencies in simulating soil insulation below -15 °C, with ΔT 1 to 2 °C higher than observation when snd is higher than 0.3 m. IPSL-CM6A-LR also underestimates the insulating effect when there is almost no snow. IPSL-CM6A-LR accurately reproduces the inflection point of the snd-ΔT curve. After reaching the inflection point, the insulation effect still increases in IPSL-CM6A-LR. While HadGEM3-GC31-LL and UKESM1.0-LL show a slower rate of increasing ΔT that is more comparable with the observations. All other models fail to reproduce the observation-like curve, underestimating the snow insulation effect under most conditions.

None of the models can accurately depict soil insulation. When snd is below 0.05 m, all models simulate ΔT that is lower than the observed value. This could be due to low resolutions of land surface models, which make it difficult to accurately determine the distribution of organic matter on the surface (de Vrese et al., 2023), resulting in a miscalculation of the surface insulation effect. This shortage poses a potential issue where increasing model soil surface insulation could lead to overestimating the total ΔT. It is especially true when some models adequately reproduce the insulation magnitude at high snd. UKESM1.0-LL simulates large insulation effects of 4 to 12 °C under low snd (lower than 0.05 m), which is closest to the benchmark and which might guide in further improvements.

Out of all the models tested, UKESM1.0-LL consistently demonstrated similar snow insulation effects in both ensembles. However, the results of the other models varied, suggesting that factors beyond tas and snd also impact the snow layer's ability to impede energy transfer. Evaluating the snow insulation effect based on LS3MIP or CMIP6 runs alone may lead to different conclusions. For example, CESM2 shows a snow insulation effect closer to observations in its LS3MIP run, but the CMIP6 run of CESM2 simulates a much lower ΔT curve. Despite cold conditions, an increase in snd still affects the snow insulation

https://doi.org/10.5194/egusphere-2025-389
Preprint. Discussion started: 10 February 2025

effect of LS3MIP CESM2. Similar results are observed in MIROC6 as well. Both models show realistic snow insulation effect values when snd is above 0.4 m. The issue with their snow insulation dynamics is the slow rise in ΔT when snd is below 0.2 m. On the other hand, CNRM-CM6.1 and CNRM-ESM2.1 have a low curve of ΔT-snd, even with a negative ΔT outcome for snd above 0.2 m under the -5 °C to -15 °C category of LS3MIP runs.

Four land models mentioned in this study are newer versions of the land models studied by Wang et al. (2016). We evaluate the different versions by comparing their snow insulation effect. The study reveals that the newer models have made noteworthy progress. CLM5 (here indicated by CESM2) generates ΔT profiles that are more similar to those observed in comparison to CLM4.5, particularly for temperatures ranging from -5 °C to -15 °C. The two newer versions of JULES (used in HadGEM3-GC31-LL and UKESM1.0-LL) outperform their predecessors when snd is less than 0.1 m. Compared to its older versions, the snow insulation effect for ORCHIDEE is also improved. The most obvious improvement is observed in the IPSL-CM6A-LR version of ORCHIDEE used in LS3MIP, where a substantial reduction in the lack of snow insulation is seen.

3.6 Impact of Land Model Features on Performance

All the selected models use multi-layer snow schemes to calculate the accumulation and melting of snowpack. The multi-layer schemes show good performance in reproducing accurate snd. And better spatial correlation with observation in LS3MIP simulations indicates the snow layering schemes adequately respond to forcing.

Most models use spectral-related methods to calculate albedo. Although IPSL-CM6A-LR employs a simpler spectral-averaged albedo scheme than other land surface models, it does not have an observable impact on its tsl simulation. Albedo has a large uncertainty beyond methods employed by land models since the albedo calculation strongly relies on the simulation of snow. For example, if the coverage and thickness of snow are not well simulated in regions of low snow depth or forest area, albedo cannot be precisely calculated (Krinner et al., 2018; Menard et al., 2021). Considering snow conductivity, the Power Function could be why CNRM-CM6.1 and CNRM-ESM2.1 have a negative bias of larger than -6 °C in the SON (figure not shown). The uncertainty of CNRM-CM6.1 and CNRM-ESM2.1 soil temperatures in spring and autumn is the largest. In autumn, all models underestimate soil temperature, which proves again models are poor in simulating snow thermal insulation effect when snow depths (also snow density) are low. IPSL-CM6A-LR simulates the most stable snd, but with the Quadratic Equation, its snow thermal conductivity is strong when snd is under 0.2 m, reflected by its lower snow insulation effect than Obs. Using similar schemes of snow does not guarantee similar performances. Better snow and soil temperature simulation are also related to other processes in land models, including soil bottom schemes. HadGEM3-GC31-LL and UKESM1.0-LL have good performance in simulating tsl with their shallow soil column. Moreover, the zero-flux assumption is possibly more influential on the soil surface when used in shallower soil column bottom; it constrains RS of soil temperature in winter. Generally speaking, considering the impact of bottom boundary conditions on soil temperature, defining a bottom flux shows better results than just giving fixed values.

Taking account of RS and RB of summer tsl, HadGEM3-GC31-LL, UKESM1.0-LL, and IPSL-CM6A-LR models exhibit the best organic layer thermal insulation. However, the significance of organic layers continues to be underestimated. Current considerations of the hydro-thermodynamic in CLM5 and Surfex 8.0c do not improve their insulation performances.

https://doi.org/10.5194/egusphere-2025-389
Preprint. Discussion started: 10 February 2025

It is challenging to identify how model features influence the vertical energy transportation process without conducting sensitivity experiments. Different physical processes represented in land surface models may interact in complex ways, either synergistically or oppositely, affecting the model's simulation capability. Without the ability to isolate the effects of these various processes, it becomes difficult to ascertain whether simulation errors are the result of one specific scheme or multiple overlapping processes.

4 Conclusions

This research investigated coupled CMIP6 and land-only LS3MIP historical climate simulations in frozen soil areas. Errors caused by the land surface models versus the errors caused by atmospheric forcings or coupled models were quantified and discussed.

Except in summer months, inaccurate inter-annual variability in the simulation of soil temperature by CMIP6 models is mainly caused by deficiencies in the land surface models and less inherited from atmospheric components. Biases in the land surface model even partially compensate for the influence of air temperature biases. Similarly, better precipitation simulation does not ensure snow depth results improve, especially in winter and spring. Soil temperature biases and their spread between models are much more evident in winter than in the other seasons. Spatially, the models exhibit larger disagreements in reproducing the soil temperature of sites in permafrost regions. The largest model biases of tas and tsl are witnessed under -5 $^{\circ}\mathrm{C}$. These indicate a weakness for models reproducing the tsl relationship with tas in freezing conditions. The deficiency of land surface models is reflected in the near-surface energy transport process, particularly during winter. The primary sources are the simulated snow amount and the snow insulation effect. Land models tend to simulate lower tsl when overlying snow exists. Further improvement of parameterization of the land model surface insulation process is necessary.

Note that the scope of this study is limited to soil depths down to 0.2 m and that the thermal state of frozen soils is not determined solely by temperature (Groenke et al., 2023). It is essential to consider various hydrothermal processes within the deeper soil, including thermal offset, permafrost active layers, seasonally frozen soil freezing depth, and heat and water transport. Additionally, the impact of the applied bottom boundary conditions is not investigated. These factors play a critical role in capturing frozen soil dynamics and have to be further investigated.

Data availability. CMIP6 and LS3MIP multi-model ensemble data (http://doi.org/10.22033/ESGF/CMIP6.4066, Voldoire (2018), http://doi.org/10.22033/ESGF/CMIP6.4095, Voldoire (2019b), http://doi.org/10.22033/ESGF/CMIP6.4095, Seferian (2018), http://doi.org/10.22033/ESGF/CMIP6.9599, Voldoire (2019a), http://doi.org/10.22033/ESGF/CMIP6.5195, Boucher et al. (2018), http://doi.org/10.22033/ESGF/CMIP6.5205, Boucher et al. (2019), http://doi.org/10.22033/ESGF/CMIP6.5603, Tatebe and Watanabe (2018), http://doi.org/10.22033/ESGF/CMIP6.5622, Onuma and Kim (2020), http://doi.org/10.22033/ESGF/CMIP6.6109, Ridley et al. (2019), http://doi.org/10.22033/ESGF/CMIP6.14460, Wiltshire et al. (2020b), http://doi.org/10.22033/ESGF/CMIP6.6113, Tang et al. (2019), http://doi.org/10.22033/ESGF/CMIP6.14462, Wiltshire et al. (2020a), http://doi.org/10.22033/ESGF/CMIP6.7627, Danabasoglu (2019b), http://doi.org/10.22033/ESGF/CMIP6.7650, Danabasoglu (2019a)) were downloaded from https://esgf-node.llnl.gov/projects/cmip6/ .

https://doi.org/10.5194/egusphere-2025-389
Preprint. Discussion started: 10 February 2025

Author contributions. ZL, BA, and DJ determined the research outline. ZL and BA wrote the initial manuscript. BA provided guidance on data analysis. ZL collected the data and was responsible for the code, calculations, and analysis of the results.

Competing interests. The contact author has declared that none of the authors has any competing interests

Acknowledgements. Zhicheng Luo gratefully acknowledges the China Scholarship Council (CSC) sponsorship for Z.L. (No.202006040064). BA acknowledges support by DWD IDEA S4S – project FS-SF (4823IDEAP2). This work used resources of the Deutsches Klimarechenzentrum (DKRZ) granted by its Scientific Steering Committee (WLA) under project ID bb1064 and of Goethe-HLR. Thanks to Danny Risto and Mittal Parmar for their contribution to this research. We express our gratitude to the World Climate Research Programme for its coordination and support of CMIP6 through the efforts of its Working Group on Coupled Modelling. We appreciate the climate modeling groups for producing and providing their model output, the Earth System Grid Federation (ESGF) for archiving the data and ensuring access, and the multiple funding agencies that support CMIP6 and ESGF.

https://doi.org/10.5194/egusphere-2025-389
Preprint. Discussion started: 10 February 2025

References

Beringer, J., Lynch, A. H., Chapin, F. S., Mack, M., and Bonan, G. B.: The Representation of Arctic Soils in the Land Surface Model: The Importance of Mosses, Journal of Climate, 14, 3324–3335, https://doi.org/10.1175/1520-0442(2001)014<3324:TROASI>2.0.CO;2, 2001.

Biskaborn, B. K., Smith, S. L., Noetzli, J., Matthes, H., Vieira, G., Streletskiy, D. A., Schoeneich, P., Romanovsky, V. E., Lewkowicz, A. G., Abramov, A., Allard, M., Boike, J., Cable, W. L., Christiansen, H. H., Delaloye, R., Diekmann, B., Drozdov, D., Etzelmüller, B., Grosse, G., Guglielmin, M., Ingeman-Nielsen, T., Isaksen, K., Ishikawa, M., Johansson, M., Johannsson, H., Joo, A., Kaverin, D., Kholodov, A., Konstantinov, P., Kröger, T., Lambiel, C., Lanckman, J.-P., Luo, D., Malkova, G., Meiklejohn, I., Moskalenko, N., Oliva, M., Phillips, M., Ramos, M., Sannel, A. B. K., Sergeev, D., Seybold, C., Skryabin, P., Vasiliev, A., Wu, Q., Yoshikawa, K., Zheleznyak, M., and Lantuit, H.: Permafrost is warming at a global scale, Nature Communications, 10, 264, https://doi.org/10.1038/s41467-018-08240-4, 2019.

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

Boucher, O., Denvil, S., Levavasseur, G., Cozic, A., Caubel, A., Foujols, M.-A., Meurdesoif, Y., Ghattas, J., Cadule, P., Ducharne, A., Vuichard, N., and Cheruy, F.: IPSL IPSL-CM6A-LR model output prepared for CMIP6 LS3MIP land-hist, https://doi.org/10.22033/ESGF/CMIP6.5205, 2019.

Bowring, S. P. K., Lauerwald, R., Guenet, B., Zhu, D., Guimberteau, M., Tootchi, A., Ducharne, A., and Ciais, P.: ORCHIDEE MICT-LEAK (r5459), a global model for the production, transport, and transformation of dissolved organic carbon from Arctic permafrost regions – Part 1: Rationale, model description, and simulation protocol, Geoscientific Model Development, 12, 3503–3521, https://doi.org/10.5194/gmd-12-3503-2019, 2019.

Brown, J., Sidlauskas, F. J., and Delinski, G.: International Permafrost Association circum-Arctic map of permafrost and ground ice conditions, The Survey ; For sale by Information Services, Reston, Va., Denver, Colo., ISBN 978-0-607-88745-7, oCLC: 38148545, 1997.

Burke, E. J., Zhang, Y., and Krinner, G.: Evaluating permafrost physics in the Coupled Model Intercomparison Project 6 (CMIP6) models and their sensitivity to climate change, The Cryosphere, 14, 3155–3174, https://doi.org/10.5194/tc-14-3155-2020, 2020.

Cao, B., Arduini, G., and Zsoter, E.: Brief communication: Improving ERA5-Land soil temperature in permafrost regions using an optimized multi-layer snow scheme, preprint, Frozen ground/Frozen Ground, https://doi.org/10.5194/tc-2022-71, 2022.

Clark, D. B., Mercado, L. M., Sitch, S., Jones, C. D., Gedney, N., Best, M. J., Pryor, M., Rooney, G. G., Essery, R. L. H., Blyth, E., Boucher, O., Harding, R. J., Huntingford, C., and Cox, P. M.: The Joint UK Land Environment Simulator (JULES), model description – Part 2: Carbon fluxes and vegetation dynamics, Geoscientific Model Development, 4, 701–722, https://doi.org/10.5194/gmd-4-701-2011, 2011.

Danabasoglu, G.: NCAR CESM2 model output prepared for CMIP6 LS3MIP land-hist, https://doi.org/10.22033/ESGF/CMIP6.7650, 2019a.

Danabasoglu, G.: NCAR CESM2 model output prepared for CMIP6 CMIP historical, https://doi.org/10.22033/ESGF/CMIP6.7627, 2019b.

de Vrese, P., Georgievski, G., Gonzalez Rouco, J. F., Notz, D., Stacke, T., Steinert, N. J., Wilkenskjeld, S., and Brovkin, V.: Representation of soil hydrology in permafrost regions may explain large part of inter-model spread in simulated Arctic and subarctic climate, The Cryosphere, 17, 2095–2118, https://doi.org/10.5194/tc-17-2095-2023, 2023.

Decharme, B., Delire, C., Minvielle, M., Colin, J., Vergnes, J., Alias, A., Saint-Martin, D., Séférian, R., Sénési, S., and Voldoire, A.: Recent Changes in the ISBA-CTRIP Land Surface System for Use in the CNRM-CM6 Climate Model and in Global Off-Line Hydrological Applications, Journal of Advances in Modeling Earth Systems, 11, 1207–1252, https://doi.org/10.1029/2018MS001545, 2019.

https://doi.org/10.5194/egusphere-2025-389
Preprint. Discussion started: 10 February 2025

Deng, M., Meng, X., Lu, Y., Li, Z., Zhao, L., Hu, Z., Chen, H., Shang, L., Wang, S., and Li, Q.: Impact and Sensitivity Analysis of Soil Water and Heat Transfer Parameterizations in Community Land Surface Model on the Tibetan Plateau, Journal of Advances in Modeling Earth Systems, 13, e2021MS002 670, https://doi.org/10.1029/2021MS002670, 2021.

Dutch, V. R., Rutter, N., Wake, L., Sandells, M., Derksen, C., Walker, B., Hould Gosselin, G., Sonnentag, O., Essery, R., Kelly, R., Marsh, P., King, J., and Boike, J.: Impact of measured and simulated tundra snowpack properties on heat transfer, The Cryosphere, 16, 4201–4222, https://doi.org/10.5194/tc-16-4201-2022, 2022.

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

Frauenfeld, O. W. and Zhang, T.: An observational 71-year history of seasonally frozen ground changes in the Eurasian high latitudes, Environmental Research Letters, 6, 044 024, https://doi.org/10.1088/1748-9326/6/4/044024, 2011.

Groenke, B., Langer, M., Nitzbon, J., Westermann, S., Gallego, G., and Boike, J.: Investigating the thermal state of permafrost with Bayesian inverse modeling of heat transfer, The Cryosphere, 17, 3505–3533, https://doi.org/10.5194/tc-17-3505-2023, 2023.

Heijmans, M. M. P. D., Magnússon, R. , Lara, M. J., Frost, G. V., Myers-Smith, I. H., Van Huissteden, J., Jorgenson, M. T., Fedorov, A. N., Epstein, H. E., Lawrence, D. M., and Limpens, J.: Tundra vegetation change and impacts on permafrost, Nature Reviews Earth & Environment, 3, 68–84, https://doi.org/10.1038/s43017-021-00233-0, 2022.

Jain, S., Scaife, A. A., Shepherd, T. G., Deser, C., Dunstone, N., Schmidt, G. A., Trenberth, K. E., and Turkington, T.: Importance of internal variability for climate model assessment, npj Climate and Atmospheric Science, 6, 68, https://doi.org/10.1038/s41612-023-00389-0, 2023.

Koven, C. D., Lawrence, D. M., and Riley, W. J.: Permafrost carbon-climate feedback is sensitive to deep soil carbon decomposability but not deep soil nitrogen dynamics, Proceedings of the National Academy of Sciences, 112, 3752–3757, https://doi.org/10.1073/pnas.1415123112, 2015.

Krinner, G., Derksen, C., Essery, R., Flanner, M., Hagemann, S., Clark, M., Hall, A., Rott, H., Brutel-Vuilmet, C., Kim, H., Ménard, C. B., Mudryk, L., Thackeray, C., Wang, L., Arduini, G., Balsamo, G., Bartlett, P., Boike, J., Boone, A., Chéruy, F., Colin, J., Cuntz, M., Dai, Y., Decharme, B., Derry, J., Ducharne, A., Dutra, E., Fang, X., Fierz, C., Ghattas, J., Gusev, Y., Haverd, V., Kontu, A., Lafaysse, M., Law, R., Lawrence, D., Li, W., Marke, T., Marks, D., Ménégoz, M., Nasonova, O., Nitta, T., Niwano, M., Pomeroy, J., Raleigh, M. S., Schaedler, G., Semenov, V., Smirnova, T. G., Stacke, T., Strasser, U., Svenson, S., Turkov, D., Wang, T., Wever, N., Yuan, H., Zhou, W., and Zhu, D.: ESM-SnowMIP: assessing snow models and quantifying snow-related climate feedbacks, Geoscientific Model Development, 11, 5027–5049, https://doi.org/10.5194/gmd-11-5027-2018, 2018.

Kuma, P., Bender, F. A., and Jönsson, A. R.: Climate Model Code Genealogy and Its Relation to Climate Feedbacks and Sensitivity, Journal of Advances in Modeling Earth Systems, 15, e2022MS003 588, https://doi.org/10.1029/2022MS003588, 2023.

Langer, M., Westermann, S., Muster, S., Piel, K., and Boike, J.: The surface energy balance of a polygonal tundra site in northern Siberia – Part 2: Winter, The Cryosphere, 5, 509–524, https://doi.org/10.5194/tc-5-509-2011, 2011a.

Langer, M., Westermann, S., Muster, S., Piel, K., and Boike, J.: The surface energy balance of a polygonal tundra site in northern Siberia – Part 1: Spring to fall, The Cryosphere, 5, 151–171, https://doi.org/10.5194/tc-5-151-2011, 2011b.

Lawrence, D. M. and Slater, A. G.: A projection of severe near-surface permafrost degradation during the 21st century, Geophysical Research Letters, 32, L24 401, https://doi.org/10.1029/2005GL025080, 2005.

Lawrence, D. M., Slater, A. G., Tomas, R. A., Holland, M. M., and Deser, C.: Accelerated Arctic land warming and permafrost degradation during rapid sea ice loss, Geophysical Research Letters, 35, 2008GL033 985, https://doi.org/10.1029/2008GL033985, 2008.

https://doi.org/10.5194/egusphere-2025-389
Preprint. Discussion started: 10 February 2025

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

Li, Q., Sun, S., and Xue, Y.: Analyses and development of a hierarchy of frozen soil models for cold region study, Journal of Geophysical Research: Atmospheres, 115, 2009JD012 530, https://doi.org/10.1029/2009JD012530, 2010.

Menard, C. B., Essery, R., Krinner, G., Arduini, G., Bartlett, P., Boone, A., Brutel-Vuilmet, C., Burke, E., Cuntz, M., Dai, Y., Decharme, B., Dutra, E., Fang, X., Fierz, C., Gusev, Y., Hagemann, S., Haverd, V., Kim, H., Lafaysse, M., Marke, T., Nasonova, O., Nitta, T., Niwano, M., Pomeroy, J., Schädler, G., Semenov, V. A., Smirnova, T., Strasser, U., Swenson, S., Turkov, D., Wever, N., and Yuan, H.: Scientific and Human Errors in a Snow Model Intercomparison, Bulletin of the American Meteorological Society, 102, E61–E79, https://doi.org/10.1175/BAMS-D-19-0329.1, 2021.

Miner, K. R., Turetsky, M. R., Malina, E., Bartsch, A., Tamminen, J., McGuire, A. D., Fix, A., Sweeney, C., Elder, C. D., and Miller, C. E.: Permafrost carbon emissions in a changing Arctic, Nature Reviews Earth & Environment, 3, 55–67, https://doi.org/10.1038/s43017-021-00230-3, 2022.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., and Thépaut, J.-N.: ERA5-Land: a state-of-the-art global reanalysis dataset for land applications, Earth System Science Data, 13, 4349–4383, https://doi.org/10.5194/essd-13-4349-2021, 2021.

Niu, G.-Y. and Yang, Z.-L.: Effects of Frozen Soil on Snowmelt Runoff and Soil Water Storage at a Continental Scale, Journal of Hydrometeorology, 7, 937–952, https://doi.org/10.1175/JHM538.1, 2006.

Obu, J., Westermann, S., Bartsch, A., Berdnikov, N., Christiansen, H. H., Dashtseren, A., Delaloye, R., Elberling, B., Etzelmüller, B., Kholodov, A., Khomutov, A., Kääb, A., Leibman, M. O., Lewkowicz, A. G., Panda, S. K., Romanovsky, V., Way, R. G., Westergaard-Nielsen, A., Wu, T., Yamkhin, J., and Zou, D.: Northern Hemisphere permafrost map based on TTOP modelling for 2000–2016 at 1 km2 scale, Earth-Science Reviews, 193, 299–316, https://doi.org/10.1016/j.earscirev.2019.04.023, 2019.

Onuma, Y. and Kim, H.: MIROC MIROC6 model output prepared for CMIP6 LS3MIP land-hist, https://doi.org/10.22033/ESGF/CMIP6.5622, 2020.

Park, H., Sherstiukov, A. B., Fedorov, A. N., Polyakov, I. V., and Walsh, J. E.: An observation-based assessment of the influences of air temperature and snow depth on soil temperature in Russia, Environmental Research Letters, 9, 064 026, https://doi.org/10.1088/1748-9326/9/6/064026, 2014.

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

Rashid, H. A.: Diverse Responses of Global-Mean Surface Temperature to External Forcings and Internal Climate Variability in Observations and CMIP6 Models, Geophysical Research Letters, 48, e2021GL093 194, https://doi.org/10.1029/2021GL093194, 2021.

https://doi.org/10.5194/egusphere-2025-389
Preprint. Discussion started: 10 February 2025

Ridley, J., Menary, M., Kuhlbrodt, T., Andrews, M., and Andrews, T.: MOHC HadGEM3-GC31-LL model output prepared for CMIP6 CMIP historical, https://doi.org/10.22033/ESGF/CMIP6.6109, 2019.

Romanovsky, V., Sazonova, T., Balobaev, V., Shender, N., and Sergueev, D.: Past and recent changes in air and permafrost temperatures in eastern Siberia, Global and Planetary Change, 56, 399–413, https://doi.org/10.1016/j.gloplacha.2006.07.022, 2007.

Rößger, N., Sachs, T., Wille, C., Boike, J., and Kutzbach, L.: Seasonal increase of methane emissions linked to warming in Siberian tundra, Nature Climate Change, 12, 1031–1036, https://doi.org/10.1038/s41558-022-01512-4, 2022.

Schwarzwald, K. and Lenssen, N.: The importance of internal climate variability in climate impact projections, Proceedings of the National Academy of Sciences, 119, e2208095 119, https://doi.org/10.1073/pnas.2208095119, 2022.

Seferian, R.: CNRM-CERFACS CNRM-ESM2-1 model output prepared for CMIP6 CMIP historical, https://doi.org/10.22033/ESGF/CMIP6.4068, 2018.

Sellar, A. A., Jones, C. G., Mulcahy, J. P., Tang, Y., Yool, A., Wiltshire, A., O'Connor, F. M., Stringer, M., Hill, R., Palmieri, J., Woodward, S., De Mora, L., Kuhlbrodt, T., Rumbold, S. T., Kelley, D. I., Ellis, R., Johnson, C. E., Walton, J., Abraham, N. L., Andrews, M. B., Andrews, T., Archibald, A. T., Berthou, S., Burke, E., Blockley, E., Carslaw, K., Dalvi, M., Edwards, J., Folberth, G. A., Gedney, N., Griffiths, P. T., Harper, A. B., Hendry, M. A., Hewitt, A. J., Johnson, B., Jones, A., Jones, C. D., Keeble, J., Liddicoat, S., Morgenstern, O., Parker, R. J., Predoi, V., Robertson, E., Siahaan, A., Smith, R. S., Swaminathan, R., Woodhouse, M. T., Zeng, G., and Zerroukat, M.: UKESM1: Description and Evaluation of the U.K. Earth System Model, Journal of Advances in Modeling Earth Systems, 11, 4513–4558, https://doi.org/10.1029/2019MS001739, 2019.

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

Smith, M. W. and Riseborough, D. W.: Climate and the limits of permafrost: a zonal analysis, Permafrost and Periglacial Processes, 13, 1–15, https://doi.org/10.1002/ppp.410, 2002.

Smith, S. L., O'Neill, H. B., Isaksen, K., Noetzli, J., and Romanovsky, V. E.: The changing thermal state of permafrost, Nature Reviews Earth & Environment, 3, 10–23, https://doi.org/10.1038/s43017-021-00240-1, 2022.

Takata, K., Emori, S., and Watanabe, T.: Development of the minimal advanced treatments of surface interaction and runoff, Global and Planetary Change, 38, 209–222, https://doi.org/10.1016/S0921-8181(03)00030-4, 2003.

Tang, Y., Rumbold, S., Ellis, R., Kelley, D., Mulcahy, J., Sellar, A., Walton, J., and Jones, C.: MOHC UKESM1.0-LL model output prepared for CMIP6 CMIP historical, https://doi.org/10.22033/ESGF/CMIP6.6113, 2019.

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

Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram, Journal of Geophysical Research: Atmospheres, 106, 7183–7192, https://doi.org/10.1029/2000JD900719, 2001.

Turetsky, M. R., Abbott, B. W., Jones, M. C., Walter Anthony, K., Olefeldt, D., Schuur, E. A. G., Koven, C., McGuire, A. D., Grosse, G., Kuhry, P., Hugelius, G., Lawrence, D. M., Gibson, C., and Sannel, A. B. K.: Permafrost collapse is accelerating carbon release, Nature, 569, 32–34, https://doi.org/10.1038/d41586-019-01313-4, 2019.

Van Den Hurk, B., Kim, H., Krinner, G., Seneviratne, S. I., Derksen, C., Oki, T., Douville, H., Colin, J., Ducharne, A., Cheruy, F., Viovy, N., Puma, M. J., Wada, Y., Li, W., Jia, B., Alessandri, A., Lawrence, D. M., Weedon, G. P., Ellis, R., Hagemann, S., Mao, J., Flanner, M. G., Zampieri, M., Materia, S., Law, R. M., and Sheffield, J.: LS3MIP (v1.0) contribution to CMIP6: the Land Surface, Snow and

https://doi.org/10.5194/egusphere-2025-389
Preprint. Discussion started: 10 February 2025

Soilmoisture Model Intercomparison Project – aims, setup and expected outcome, Geoscientific Model Development, 9, 2809–2832, https://doi.org/10.5194/gmd-9-2809-2016, 2016.

Van Kampenhout, L., Lenaerts, J. T. M., Lipscomb, W. H., Sacks, W. J., Lawrence, D. M., Slater, A. G., and Van Den Broeke, M. R.: Improving the Representation of Polar Snow and Firn in the Community Earth System Model, Journal of Advances in Modeling Earth Systems, 9, 2583–2600, https://doi.org/10.1002/2017MS000988, 2017.

Vionnet, V., Brun, E., Morin, S., Boone, A., Faroux, S., Le Moigne, P., Martin, E., and Willemet, J.-M.: The detailed snowpack scheme Crocus and its implementation in SURFEX v7.2, Geoscientific Model Development, 5, 773–791, https://doi.org/10.5194/gmd-5-773-2012, 2012.

Voldoire, A.: CMIP6 simulations of the CNRM-CERFACS based on CNRM-CM6-1 model for CMIP experiment historical, https://doi.org/10.22033/ESGF/CMIP6.4066, 2018.

Voldoire, A.: CNRM-CERFACS CNRM-ESM2-1 model output prepared for CMIP6 LS3MIP land-hist, https://doi.org/10.22033/ESGF/CMIP6.9599, 2019a.

Voldoire, A.: CNRM-CERFACS CNRM-CM6-1 model output prepared for CMIP6 LS3MIP land-hist, https://doi.org/10.22033/ESGF/CMIP6.4095, 2019b.

Walters, D., Baran, A. J., Boutle, I., Brooks, M., Earnshaw, P., Edwards, J., Furtado, K., Hill, P., Lock, A., Manners, J., Morcrette, C., Mulcahy, J., Sanchez, C., Smith, C., Stratton, R., Tennant, W., Tomassini, L., Van Weverberg, K., Vosper, S., Willett, M., Browse, J., Bushell, A., Carslaw, K., Dalvi, M., Essery, R., Gedney, N., Hardiman, S., Johnson, B., Johnson, C., Jones, A., Jones, C., Mann, G., Milton, S., Rumbold, H., Sellar, A., Ujiie, M., Whitall, M., Williams, K., and Zerroukat, M.: The Met Office Unified Model Global Atmosphere 7.0/7.1 and JULES Global Land 7.0 configurations, Geoscientific Model Development, 12, 1909–1963, https://doi.org/10.5194/gmd-12-1909-2019, 2019.

Wang, T., Ottlé, C., Boone, A., Ciais, P., Brun, E., Morin, S., Krinner, G., Piao, S., and Peng, S.: Evaluation of an improved intermediate complexity snow scheme in the ORCHIDEE land surface model, Journal of Geophysical Research: Atmospheres, 118, 6064–6079, https://doi.org/10.1002/jgrd.50395, 2013.

Wang, W., Rinke, A., Moore, J. C., Ji, D., Cui, X., Peng, S., Lawrence, D. M., McGuire, A. D., Burke, E. J., Chen, X., Decharme, B., Koven, C., MacDougall, A., Saito, K., Zhang, W., Alkama, R., Bohn, T. J., Ciais, P., Delire, C., Gouttevin, I., Hajima, T., Krinner, G., Lettenmaier, D. P., Miller, P. A., Smith, B., Sueyoshi, T., and Sherstiukov, A. B.: Evaluation of air–soil temperature relationships simulated by land surfacemodels during winter across the permafrost region, The Cryosphere, 10, 1721–1737, https://doi.org/10.5194/tc-10-1721-2016, 2016.

Wilks, D. S.: Statistical methods in the atmospheric sciences, Elsevier, Amsterdam, Netherlands ; Cambridge, MA, fourth edition edn., ISBN 978-0-12-815823-4, 2019.

Wiltshire, A., Robertson, E., Burke, E., and Liddicoat, S.: MOHC UKESM1.0-LL model output prepared for CMIP6 LS3MIP, https://doi.org/10.22033/ESGF/CMIP6.14462, 2020a.

Wiltshire, A., Robertson, E., Burke, E., and Liddicoat, S.: MOHC HadGEM3-GC31-LL model output prepared for CMIP6 LS3MIP, https://doi.org/10.22033/ESGF/CMIP6.14460, 2020b.

Wiltshire, A. J., Duran Rojas, M. C., Edwards, J. M., Gedney, N., Harper, A. B., Hartley, A. J., Hendry, M. A., Robertson, E., and Smout-Day, K.: JULES-GL7: the Global Land configuration of the Joint UK Land Environment Simulator version 7.0 and 7.2, Geoscientific Model Development, 13, 483–505, https://doi.org/10.5194/gmd-13-483-2020, 2020c.

Ye, K.: Inter-model spread in the wintertime Arctic amplification in the CMIP6 models and the important role of internal climate variability, Global and Planetary Change, 2021.

https://doi.org/10.5194/egusphere-2025-389
Preprint. Discussion started: 10 February 2025

You, Q., Cai, Z., Pepin, N., Chen, D., Ahrens, B., Jiang, Z., Wu, F., Kang, S., Zhang, R., Wu, T., Wang, P., Li, M., Zuo, Z., Gao, Y., Zhai, P., and Zhang, Y.: Warming amplification over the Arctic Pole and Third Pole: Trends, mechanisms and consequences, Earth-Science Reviews, 217, 103 625, https://doi.org/10.1016/j.earscirev.2021.103625, 2021.

595 Zhang, T.: Influence of the seasonal snow cover on the ground thermal regime: An overview, Reviews of Geophysics, 43, 2004RG000 157, https://doi.org/10.1029/2004RG000157, 2005.

Zhang, Y., Sherstiukov, A. B., Qian, B., Kokelj, S. V., and Lantz, T. C.: Impacts of snow on soil temperature observed across the circumpolar north, Environmental Research Letters, 13, 044 012, https://doi.org/10.1088/1748-9326/aab1e7, 2018.

Åkerman, H. J. and Johansson, M.: Thawing permafrost and thicker active layers in sub-arctic Sweden, Permafrost and Periglacial Processes,

600 19, 279–292, https://doi.org/10.1002/ppp.626, 2008.