

# Permafrost sensitivity to soil hydro-thermodynamics in historical and scenario simulations with the MPI-ESM

Félix García-Pereira<sup>1,2</sup>, Jesús Fidel González-Rouco<sup>1</sup>, Nagore Meabe-Yanguas<sup>1</sup>, Philipp de Vrese<sup>3</sup>, Norman Julius Steinert<sup>4</sup>, Johann Jungclaus<sup>3</sup>, and Stephan Lorenz<sup>3</sup>

<sup>1</sup>Geosciences Institute, IGEO (CSIC-UCM), Madrid, Spain

<sup>2</sup>Faculty of Physical Sciences, Complutense University of Madrid (UCM), Madrid, Spain

<sup>3</sup>Max Planck Institute for Meteorology, Hamburg, Germany

<sup>4</sup>CICERO - Center for International Climate Research, Oslo, Norway

**Correspondence:** Félix García-Pereira (felgar03@ucm.es)

## Abstract.

Global warming is particularly intense in the Arctic, where temperature trends are up to four times higher than those observed globally. Among its multiple consequences, this enhanced warming is especially concerning because it triggers the degradation of permafrost. However, the increasing temperature is not the only factor conditioning permafrost thaw. Changes in water availability induced by disturbances in the permafrost water table considerably affect Arctic soils moisture and ice presence and thermal structure in permafrost regions. Our understanding of the interactions between hydrology and thermodynamics is still limited as well as its representation in the majority of state-of-the-art land surface models (LSMs), mainly in terms of the processes included, LSM depth, spatial resolution, vertical discretization, and hydro-thermodynamic coupling.

This work explores the response of the Max Planck Institute Earth System Model (MPI-ESM) to changes in the soil hydrology and thermodynamics in permafrost-affected regions. An ensemble of fully-coupled historical and climate change scenario simulations was performed under three configurations: the standard model that will be taken as a reference, and two variants that generate rather *dry* or *wet* conditions across permafrost areas. Enhanced soil depth and vertical resolution within the LSM, JSBACH, were also incorporated globally to capture fine-scale thermodynamics and allow for deeper heat propagation. Results show that deepening the LSM reduces the intensity of near-surface soil warming by  $0.1^{\circ}\text{C dec}^{-1}$  in high radiative forcing scenarios, reducing permafrost retreat by up to  $1.9\text{--}3.1 \cdot 10^6 \text{ km}^2$  by the end of the 21st century. However, the greatest influences are produced by the *dry* and *wet* configurations, which lead to distinct initial states, historical, and future evolution for permafrost temperatures (offset of  $3^{\circ}\text{C}$ ), active layer thickness (1–2 m) and permafrost extent ( $5 \cdot 10^6 \text{ km}^2$ ). This work underseores the importance of refining hydrological and thermodynamic processes in ESMs to improve projections of permafrost responses under climate change scenarios, with implications for the assessment of risks related to carbon feedbacks and infrastructure vulnerabilities in Arctic regions.

The limited representation of soil processes in Land Surface Models (LSMs) contributes to the uncertainty in current state and future projections of permafrost change. In particular, insufficient LSM depths, coarse vertical discretizations, and the omission of hydro-thermodynamic coupling can strongly affect subsurface temperatures, active layer thickness (ALT), and permafrost extent. This work aims to gain knowledge on permafrost sensitivity to changes in the soil hydrology and thermody-

namics in permafrost-affected regions. We explore the response of the Max Planck Institute Earth System Model (MPI-ESM) to historical and climate change scenario forcing using an ensemble of fully-coupled simulations under three configurations of permafrost hydrology: the standard model that will be taken as a *reference*, and two variants that generate rather *dry* or *wet* conditions across permafrost areas. Enhanced soil depth and vertical resolution within the LSM, JSBACH, were also incorporated globally to capture fine-scale thermodynamics and allow for deeper heat propagation.

Deepening the LSM reduces near-surface soil warming by about  $0.1\text{ }^{\circ}\text{C dec}^{-1}$  in high radiative forcing scenarios, reducing permafrost retreat by up to  $1.9\text{--}3.1\text{ }10^6\text{ km}^2$  by the end of the 21st century. However, the greatest influences are produced by the *dry* and *wet* configurations, which lead to distinct initial states, historical, and future evolution for permafrost temperatures (offset of  $3\text{ }^{\circ}\text{C}$ ), active layer thickness ( $1\text{--}2\text{ m}$ ) and permafrost extent ( $5\text{ }10^6\text{ km}^2$ ). These results indicate that inter-model spread in permafrost responses to climate change can be partly due to differences in the representation of soil physics. This underscores the importance of refining LSM hydrological and thermodynamic processes in ESMs, with implications for the assessment of risks related to carbon feedbacks and infrastructure vulnerabilities in Arctic regions.

## 1 Introduction

Global mean near-surface air temperature has increased by more than  $1\text{ }^{\circ}\text{C}$  since the end of the 19th century due to an enhanced anthropogenic greenhouse effect (Chen et al., 2021; Gulev et al., 2021). This warming is *nearly* four times more intense in northern high-latitude areas due to Arctic amplification (Rantanen et al., 2022). As most of the arctic and subarctic land areas are underlain by permafrost (about  $20.8\text{ }10^6\text{ km}^2$ ; Obu, 2021), i.e., perennially frozen soils, the intense and continuous surface warming is propagated within the soil degrading the permafrost. As a result, an increasing amount of soil organic carbon that has accumulated in the region since the last glaciation (ca.  $1035 \pm 150\text{ PgC}$ ; Hugelius et al., 2014) can be decomposed by microbes in the soil (Hugelius et al., 2014; Turetsky et al., 2019). This, in turn, releases greenhouse gases into the atmosphere, amplifying the warming signal, i.e., the permafrost carbon feedback (Schuur et al., 2022). However, permafrost thaw does not only have implications at a global scale, but also has substantial regional impacts. Permafrost degradation causes soil compression and subsidence and the development of surface water bodies (Vonk et al., 2015), posing structural risks for infrastructures in the Arctic (Hjort et al., 2022). Moreover, thawing triggers contamination problems from mercury release to the atmosphere and ground waters (Schaefer et al., 2020; Langer et al., 2023). A good understanding of soil dynamics of permafrost areas is therefore crucial to foreseeing the global consequences of their degradation and evaluating the risks for Arctic communities and ecosystems (Parmesan et al., 2023).

The state of permafrost is monitored on relatively scarce and heterogeneously distributed sites across the Arctic. These sparse measures have been aggregated to create observational products that allow for analyzing the state of permafrost at a global scale. For instance, the Circumpolar Active Layer Monitoring Network (CALM; Brown et al., 2000) tracks the long-term response of the active layer thickness (ALT), i.e., the annual maximum permafrost thaw depth, while the Global Terrestrial Network for Permafrost (GTN-P; Biskaborn et al., 2015) also monitors permafrost temperatures at different levels. Despite their great value, these observational products offer a limited representation of permafrost areas due to their poor temporal

and spatial resolution (Brown et al., 2002; Biskaborn et al., 2019; Obu et al., 2019), and the uncertainties associated with the thermal and hydrological properties of the terrain (Heuvelink, 2018).

State-of-the-art Earth System Models (ESMs) avert the spatial and temporal heterogeneity of the observational permafrost measurements. They allow for assessing the thermodynamic and hydrological responses of permafrost regions to the combined influence of natural and anthropogenic forcing (Andresen et al., 2020; Burke et al., 2020; Steinert et al., 2023). Near-surface temperatures and soil moisture resolved by Land Surface Models (LSMs) within ESMs can be subsequently used to derive estimates of ~~PE~~permafrost extent (PE) and ALT evolutions (Burke et al., 2020; Steinert et al., 2023). ESM simulations of the industrial period stemming from different models of the Coupled Model Intercomparison Project Phases 5 (CMIP5; Taylor et al., 2012) and 6 (CMIP6; Eyring et al., 2016) agree on a severe decrease of PE and deepening of ALT in the industrial era (Koven et al., 2013; Burke et al., 2020; Steinert et al., 2023). However, there is considerable inter-model variability in the results, with differences in PE decrease of 10 million km<sup>2</sup> (Steinert et al., 2023) and mean ALT deepening of 2 m (Burke et al., 2020) by the end of the 21st century. These discrepancies are mainly due to different climate sensitivity of the ESMs and the different modeling of soil hydrology and thermodynamic processes by their LSMs, which are especially sensitive for Arctic regions (Slater and Lawrence, 2013). The representation of different processes related to surface and soil hydrology, such as snow insulation (Mudryk et al., 2020; Menard et al., 2021; Zhu et al., 2021), inclusion/omission of organic matter and moss insulation effects (Walvoord and Kurylyk, 2016), and vertical drainage impedance in the presence of ice (Andresen et al., 2020) can lead to relatively wetter or drier states in Arctic soils (de Vrese et al., 2023), with feedbacks for soil moisture availability and water cycle evolution under future warming Shared Socioeconomic Pathway (SSP) scenarios (Andresen et al., 2020). Moreover, as shown by Chadburn et al. (2015), a more realistic representation of these processes within LSMs can lead to estimates of ALT and PE that are closer to the latest observation-constrained estimates.

On the other hand, the representation of subsurface thermodynamics in ESMs is influenced by the depth of the LSM. This LSM depth for the heat is defined by a zero-flux bottom boundary condition imposed at the deepest layer, which ensures energy preservation. Most ESMs include LSMs that are too shallow (Steinert et al., 2021a) to realistically accommodate long-term soil warming trends (González-Rouco et al., 2009, 2021) and heat storage (Cuesta-Valero et al., 2016; García-Pereira et al., 2024a; Steinert et al., 2024) under long-term anthropogenic warming. Including too shallow LSMs leads to overestimated soil temperature trends and variability, both globally (González-Rouco et al., 2021; Steinert et al., 2021b) and over permafrost regions (Alexeev et al., 2007; Hermoso de Mendoza et al., 2020). These temperature biases yield an overestimation of ALT deepening and PE loss with warming, being more pronounced for deep permafrost (Hermoso de Mendoza et al., 2020). This is especially sensitive for assessing the vulnerability of vast areas in Alaska and Siberia underlain by ~~yedoma~~Yedoma, i.e., the deep organic-rich permafrost formed in the Pleistocene that contains around 400 PgC (Schuur et al., 2022), additional to the ca. 1100 PgC allocated in near-surface permafrost (Hugelius et al., 2014).

Previous studies using the JULES (Chadburn et al., 2015), the MATSIRO (Yokohata et al., 2020), the CLM4.5 (Hermoso de Mendoza et al., 2020), ~~and~~ the JSBACH (Ekici et al., 2014; Steinert et al., 2021b), ~~and the~~ QUINCY (Lacroix et al., 2022) LSMs have explored the impacts of a better representation of thermodynamic and hydrological permafrost physics in standalone LSM simulations. ~~This work examines~~Here we examine for the first time the potential role of Arctic soil wetness, LSM depth, and

layer discretization in driving changes in PE loss and ALT deepening in the historical and SSP scenario periods using fully-coupled ESM simulations. For that purpose, an ensemble of simulations with a modified version of the Max Planck Institute Earth System Model (MPI-ESM) was run. These modifications entail an increased LSM depth with a finer layer discretization and more realistic features in the permafrost soil hydrology of the MPI-ESM LSM component, JSBACH. ~~Some of these features consist of a multi-layer snow scheme, the dynamic response of soil thermal properties to hydrological changes, the occurrence of phase changes within the soil, the representation of the thermodynamic and hydrological effects of the organic layer, and the inclusion of a wetland scheme (de Vrese et al., 2023). As the relative contribution of the different soil hydrological processes in permafrost soils, such as evaporation, runoff, or drainage, can produce diverging states of its soil moisture content under similar atmospheric conditions (Andresen et al., 2021), these processes can be numerically tuned to generate different states of the Arctic water balance.~~ In this paper, the sensitivity of permafrost dynamics to changes in soil hydrology is assessed under three configurations of the Arctic water-cycle. ~~One is the reference configuration (REF, hereafter), which corresponds to the configuration of the MPI-ESM used in the CMIP6 (Mauritsen et al., 2019). The other two configurations aim at representing two extreme hydrological states of the Arctic covering the range of uncertainty shown by the latest ESMs in climate scenario simulations (de Vrese et al., 2023): one with relatively higher and the other one with relatively lower soil wetness, coined as WET and DRY hereafter.~~ In order to assess how extreme these hydrologically divergent configurations are in representing PE and ALT, the results stemming from the ensemble are compared to CMIP6 PE estimates ~~in~~from Steinert et al. (2023). This comparison is extended to the historical PE using observation-constrained data coming from the European Space Agency Climate Change Initiative (Obu et al., 2021).

The article is structured as follows: an explanation on the physics of the modified version of JSBACH and MPI-ESM and the design and configuration of the experiments analyzed in this work is provided by Section 2, as well as of the metrics and data sets used for comparison; the response of the different configurations and vertical discretizations of JSBACH in terms of near-surface temperature, ALT, and PE are described and compared to observational estimates in Section 3; a discussion of the results and concluding remarks are presented in Section 4.

## 2 Data and methods

### 2.1 Model description

For this work, ~~an~~An ensemble of fully-coupled simulations ~~conducted with~~has been developed, using both the standard version of the ~~MPI-ESM1.2-LR (Mauritsen et al., 2019)~~MPI-ESM (MPI-ESM1.2-LR, Mauritsen et al., 2019) and a modified version that accounts for an improved representation of permafrost physics ~~has been developed for this work. The ensemble is coined as the MPI-ESM Permafrost Physics Ensemble (MPIESM-PePE) hereafter.~~ MPI-ESM1.2 consists of the atmosphere model ECHAM6.3 (Stevens et al., 2013), and the ocean model MPIOM1.6, including the ocean biogeochemistry module HAMOCC (Jungclaus et al., 2013). In the low-resolution LR setup, the ECHAM resolution is T63/L95 and MPIOM is GR1.5/L40, which corresponds to a  $1.875^\circ$  ( $1.5^\circ$ ) grid cell width over land (ocean). The LSM is JSBACH3.2 (version 3.2, Reick et al., 2021), which includes the simulation of the surface and subsurface thermodynamics and hydrology as well as vegetation dynamics

(Giorgetta et al., 2013). The standard version of JSBACH3.2 (~~JSBACH hereafter~~; Fig. 1) does not include any coupling processes between soil thermodynamics and hydrology, which prevents soil ice from forming under freezing conditions. This limitation was addressed by Ekici et al. (2014) and Steinert et al. (2021b) by incorporating the occurrence of water phase changes with temperature, and the dynamic calculation of soil thermal properties in previous versions of JSBACH. However, these changes were run only in standalone LSM simulations. A modified version of JSBACH3.2 allowing for hydro-thermodynamic soil coupling in permafrost regions in fully-coupled ESM simulations (JSBACH-HTCp hereafter, Fig. 2b) was introduced by de Vrese and Brovkin (2021) and de Vrese et al. (2023) and is used in this work. This version is only active across the areas covered by the mask shown in Fig. 2a (Hugelius et al., 2013, 2014), which represents the observed early 21st century permafrost extent. This allows for attributing the simulated changes in global climate exclusively to differences in the representation of processes in the northern permafrost domain. ~~Furthermore, the JSBACH3.2 standard vertical scheme of five layers (5L) with a LSM depth of 9.83 m was deepened and refined near the surface. An enhanced vertical resolution accounts for a better representation of hydro-thermodynamic processes near the surface (Chadburn et al., 2015). Further, a more realistic deeper LSM depth permits an unbiased representation of subsurface temperature variability and land heat storage (Gonzalez-Rouco et al., 2021; Steinert et al., 2024; Garcia-Pereira et al., 2024). Thus, two new vertical discretizations are introduced: an 11-layer (11L) scheme with a LSM depth of 9.98 m, and a deepened 18-layer (18L) scheme with a LSM depth of 1391.48 m (Fig. 1). The 11L scheme substitutes the original 5L discretization by enhancing the vertical resolution while keeping an almost identical LSM depth. The 18L scheme shares the top 11 layers with the 11L scheme and adds seven additional layers to reach a depth of 1391.48 m. Imposing such a deep zero-flux bottom boundary condition guarantees that the bottom layer temperature is detached from surface temperature variations at multi-centennial and millennial time scales (Garcia-Pereira et al., 2024). A detailed description of the mid-layer ( $z_i$ ) and LSM depths and thickness ( $dz_i$ ) values of the 5L, 11L, and 18L schemes used in MPIESM-PePe can be found in Table 1 and Fig. 1. The design of the layering follows an exponential function of the form  $dz_i = a_1 \exp[a_2(i - a_2) + a_3] + a_4$  (Oleson et al., 2010; Gonzalez-Rouco et al., 2021). In this work, the coefficients  $a_1$ ,  $a_2$ ,  $a_3$ , and  $a_4$  were set to 0.00198, 0.743, 83.77, and 0.0585, respectively. These values for  $a_{1-4}$  aimed at attaining virtually identical LSM depths for the 5L and 11L configurations, while maintaining a thickness of 0.065 m for the first soil layer in the three vertical schemes. The same  $a_{1-4}$  values were used to estimate the thickness of layers 12–18 for the 18L configuration (Table 1).~~

A more realistic feature in JSBACH-HTCp in comparison with JSBACH is that both the soil water phase and thermal properties vary with temperature and moisture content, respectively. This allows for the existence of ice in Arctic soils, which is not represented by JSBACH. Other new hydrological features included with the JSBACH-HTCp comprise adapted interactions between soil hydrology and vegetation, with the inclusion of an organic topsoil layer, the possibility for supercooled water, an improved representation of percolation and drainage by including the effect of soil ice impedance, and the implementation of a simple wetland and a new multi-layer snow ~~and a simple wetland~~ scheme. ~~This multi-layer scheme explicitly resolves thermal conductivity within the snow cover.~~ By contrast, JSBACH uses a single-layer scheme that does not resolve internal snow temperature gradients, but only affects the computation of surface albedo and the thermal properties of the top soil layers. A thorough description of ~~these~~all changes incorporated in JSBACH-HTCp can be found in de Vrese et al. (2023). Moreover,

the JSBACH-HTCp allows for controlling some of the hydrological processes involved in the surface and near-surface soil water cycle within the permafrost areas (Fig. 2a). Two combinations of parameters (hereafter referred to as MPIESM-PePE configurations) based on de Vrese et al. (2023) were used to generate two subsets of fully-coupled experiments with relatively drier or wetter states of the Arctic soil, DRY and WET, respectively. The DRY configuration (red labels and arrows in Fig. 2b) exhibits weak local moisture recycling due to low infiltration rates and drainage resistance. Consequently, the near-surface permafrost degradation results in progressively drier soils under this configuration. In contrast, the WET configuration (blue, Fig. 2b) assumes favorable infiltration coupled with high drainage resistance, leading to moist soils. An intense local moisture recycling is also fostered by imposing a low resistance to evapotranspiration. Apart from the two subsets of three DRY and WET simulations with different vertical discretizations, a third subset of experiments was run ~~with~~using the ~~standard~~ physics of the original version of JSBACH (REF configuration hereafter) imposed globally as a benchmark to assess the impacts of implementing the WET and DRY configurations across the permafrost domain. Outside of the permafrost mask (Fig. 2a), the soil physics are the same as included in REF (Fig. 1) in all the experiments.

Furthermore, the JSBACH standard vertical scheme of five layers (5L) with a LSM depth of 9.83 m was deepened and refined near the surface. An enhanced vertical resolution accounts for a better representation of hydro-thermodynamic processes near the surface (Chadburn et al., 2015, see Table 1). Further, a more realistic deeper LSM depth permits an unbiased representation of subsurface temperature variability and land heat storage (González-Rouco et al., 2021; Steinert et al., 2021a; García-Pereira et al., 2024a). Thus, two new vertical discretizations are introduced: an 11-layer (11L) scheme with a LSM depth of 9.98 m, and a deepened 18-layer (18L) scheme with a LSM depth of 1391.48 m (Fig. 1). The 11L scheme substitutes the original 5L discretization by enhancing the vertical resolution while keeping an almost identical LSM depth. The 18L scheme shares the top 11 layers with the 11L scheme and adds seven additional layers to reach a depth of 1391.48 m. Imposing such a deep zero-flux bottom boundary condition guarantees that the bottom layer temperature is detached from surface temperature variations at multi-centennial and millennial time scales (García-Pereira et al., 2024a). A detailed description of the mid-layer ( $z_i$ ) and LSM depths and thickness ( $dz_i$ ) values of the 5L, 11L, and 18L schemes used in this work can be found in Table 1 and Fig. 1. The design of the layering follows an exponential function of the form  $dz_i = a_1 \exp[a_2(i - a_2) + a_3] + a_4$  (Oleson et al., 2010; González-Rouco et al., 2021). In this work, the coefficients  $a_1$ ,  $a_2$ ,  $a_3$ , and  $a_4$  were set to 0.00198, 0.743, 83.77, and 0.0585, respectively. These values for  $a_{1-4}$  aimed at attaining virtually identical LSM depths for the 5L and 11L configurations, while maintaining a thickness of 0.065 m for the first soil layer in the three vertical schemes. The same  $a_{1-4}$  values were used to estimate the thickness of layers 12–18 for the 18L configuration (Table 1).

## 2.2 Experimental setup: the MPIESM-PePE

Combining the three configurations of permafrost hydro-thermodynamics (WET, DRY, and REF) and the three vertical discretizations of the subsurface (5L, 11L, and 18L) yields a total number of nine experiments integrating the ensemble presented in this work, the MPI-ESM Permafrost Physics Ensemble (MPIESM-PePE hereafter). The MPIESM-PePE extends from 1850 to 2100. The study focuses on assessing the combined impact of the different configurations and vertical discretizations on permafrost-related variables under industrial forcing conditions. The experiment setup follows the scheme in Fig. 3. In all

**Table 1.** Layer number ( $i$ ), mid-layer depth ( $z_i$ ), LSM depth, and thickness ( $dz_i$ ) of each layer in the three MPIESM-PePE vertical discretizations used herein. Note that 5L corresponds to the vertical structure in the standard JSBACH version, and 11L and 18L correspond to the vertical resolution enhancement and extension introduced in this work (see also Fig. 1).

| $i$              | $z_i$ (m) | LSM depth (m) | $dz_i$ (m) |
|------------------|-----------|---------------|------------|
| <b>5L</b>        |           |               |            |
| 1                | 0.03      | 0.06          | 0.06       |
| 2                | 0.19      | 0.32          | 0.26       |
| 3                | 0.78      | 1.32          | 1.00       |
| 4                | 2.68      | 4.13          | 2.81       |
| 5                | 6.98      | 9.83          | 5.70       |
| <b>11L / 18L</b> |           |               |            |
| 1                | 0.03      | 0.06          | 0.06       |
| 2                | 0.18      | 0.30          | 0.24       |
| 3                | 0.42      | 0.54          | 0.24       |
| 4                | 0.66      | 0.78          | 0.24       |
| 5                | 0.92      | 1.06          | 0.28       |
| 6                | 1.21      | 1.36          | 0.30       |
| 7                | 1.59      | 1.82          | 0.46       |
| 8                | 2.14      | 2.46          | 0.64       |
| 9                | 3.03      | 3.60          | 1.14       |
| 10               | 4.67      | 5.74          | 2.14       |
| 11               | 7.86      | 9.98          | 4.24       |
| 12               | 14.33     | 18.68         | 8.70       |
| 13               | 27.66     | 36.64         | 17.96      |
| 14               | 55.42     | 74.20         | 37.56      |
| 15               | 113.49    | 152.78        | 78.58      |
| 16               | 235.28    | 317.78        | 165.00     |
| 17               | 490.91    | 664.04        | 346.26     |
| 18               | 1027.76   | 1391.48       | 727.44     |

195 ~~the~~ simulations the ocean component was initialized using a restart from a long-term ocean simulation ~~restart that had been~~ stabilized to pre-industrial conditions. Each 5L and 11L experiment was run first in a piControl (PIC) phase of 50 years, which simulates a climate state compatible with the external forcing in 1850 and provides initial conditions for the historical runs (HIS; 1850 to 2014). These runs are subsequently continued into the 21st climate change stage under different radiative forcing scenarios (SSP1-1.9, SSP2-4.5, and SSP5-8.5; Riahi et al., 2017), following the standard forcings of CMIP6 (Eyring et al.,

200 2016). However, subsurface temperatures at the deepest layers of the 18L simulations do not have enough time to converge to the near-surface temperature equilibrium state within a 50-year-long PIC (González-Rouco et al., 2021). Therefore, a pre-piControl (prePIC) phase of 100 years is run before PIC for the 18L experiments to equilibrate the subsurface vertical structure and depart from vertically uniform temperatures. To speed up temperature convergence, the mean subsurface temperature of the last 50 years of prePIC of layer 12 (ST12) is imposed as the initial condition for the deepest layers (ST13 to ST18) in the  
205 18L PIC simulations. The rest of the STs are continued from the last timestep in prePIC.

## 2.3 Permafrost evaluation metrics

### 2.3.1 Permafrost temperature variability

Surface and subsurface permafrost temperature variability is characterized by the mean annual air temperature at 2 m (MAAT), mean annual ground surface temperature (MAGST), which is here considered as the near-surface temperature at a depth of 0.1  
210 m, and mean annual subsurface temperature at a depth of 5 m (MAST 5 m). The MAGST value is taken at 0.1 m depth for two reasons: to state a common depth for comparison between the 5L, 11L and 18L simulations, which have different mid-layer depths from layer 2 downwards (Table 1); and to fully account for the change in soil properties when there ~~exists~~ ~~is~~ snow cover. The latter is due to the fact that REF considers that snow gradually replaces the top layers of the soil when calculating the heat capacity and thermal diffusivity (Reick et al., 2021), thus modifying near-surface heat transfer and temperature variability.  
215 To address the temporal and spatial evolutions of MAAT, MAGST, and MAST 5 m, areal boxplots summarizing the distribution of these variables over the permafrost domain (permafrost domain boxplots herein, see permafrost mask in Fig. 2a) are used for all the members of the JSBACH-HTCp ensemble in different periods. Furthermore, to assess the insulation effects of the refined multiple-layer snow scheme and organic layer introduced by Ekici et al. (2014), the winter and summer temperature offsets are assessed, respectively. In this paper, the winter (summer) offset is computed as the ground surface temperature at 0.1  
220 m depth (GST) vs. surface air temperature (SAT) differences (GST-SAT) in December-January-February (June-July-August), following a similar approach to Burke et al. (2020). *Winter and summer offsets govern seasonal ground–air coupling, so evaluating them helps identify model biases in the representation of present-day and future projections of permafrost temperature and ALT.*

### 2.3.2 Active layer thickness (ALT)

225 ALT is defined as the annual maximum thaw depth, which in northern high latitudes occurs in summer. This variable is computed by JSBACH as the annual maximum subsurface depth where the temperature exceeds the water freezing/melting point, i.e., 273.15 K. ~~In case this depth does not specifically match a certain mid-layer depth value, JSBACH yields ALT using linear interpolation~~ If the depth of the 273.15 K isotherm does not precisely coincide with a given mid-layer depth, JSBACH computes ALT by linearly interpolating between the temperatures of the two adjacent layers that are immediately warmer or  
230 cooler than this isotherm. For all vertical discretizations, the maximum physically resolvable ALT corresponds to the mid-layer depth of its bottom layer. To ensure comparability between schemes, ALT was capped to 7.86 m; any value above this threshold

was set to 7.86 m. This limit matches the mid-layer depth of the bottom layer in the 11L configuration and thus represents the deepest meaningful ALT for that setup. To evaluate the temporal and spatial variability of ALT, the same type of boxplot analysis as described for MAAT, MAGST, and MAST 5 m in Section 2.3.1 was used.

235 Furthermore, the spatial progression of ALT deepening under industrial warming in the MPIESM-PePE is analyzed by evaluating the fraction of grid cells within the initial permafrost domain (Fig. 2a) that exhibit an ALT shallower than a given threshold depth. This approach yields vertical profiles with values ranging from 0 (no gridcells) to 1 (all gridcells comply with the condition  $ALT < \text{depth}$ ). As a result of permafrost thawing and subsequent ALT deepening with warming, this profile is expected to displace leftwards, i.e., values become closer to 0; the opposite happens under permafrost formation with cooling  
240 conditions. Finally, near-surface and deep permafrost degradation is evaluated in terms of ALT deepening assessing the time evolution of maps with gridcells having an ALT above 3 m and an ALT equal to the bottom mid-layer depth, respectively.

### 2.3.3 Permafrost extent (PE)

Permafrost is defined as the ground that remains perennially frozen for two or more consecutive years (Obu, 2021). However, there is a wide variety of metrics to evaluate whether the soil is frozen or not and subsequently determine PE based on  
245 different criteria that consider soil thermodynamics, hydrology, or soil coupling to the atmosphere (Steinert et al., 2023). Air-ground temperature coupling and some of the thermodynamic-based definitions capture near-surface PE variations, whilst others represent long-term deep PE changes. Hence, as the diverse definitions of PE target different aspects of permafrost physics, they can yield considerably different estimates. The PE differences between the different definitions are even larger than the inter-model variability of the CMIP6 ensemble (Steinert et al., 2023).

250 Since the WET and DRY configurations represent extreme states of the hydrological cycle in the permafrost domain, encompassing a substantial portion of the uncertainty range within the ensemble of state-of-the-art Earth System Models (de Vrese et al., 2023), it is of interest to assess whether they also capture the range of variability in PE responses. To address this question, ~~trying, we try~~ to minimize the uncertainty associated with the PE definition selection, ~~by estimating~~ the PE evolution for the 9 ~~historical + SSP5-8.5 simulations of the MPIESM-PePE is estimated~~ using two of these metrics: i) TTOP, which defines  
255 permafrost based on temperatures at the top of the permafrost layer being below 0 °C (Obu et al., 2019); and ii) ZAA, which considers permafrost existence when the temperature at the depth of zero annual amplitude is below freezing. TTOP has proven to be a comparatively better indicator of surface permafrost presence (Obu et al., 2019; Steinert et al., 2023), while ZAA is a more appropriate descriptor of deep permafrost occurrence than TTOP (Slater and Lawrence, 2013; Burke et al., 2020). The zero annual amplitude layer, defined as the subsurface layer where the annual temperature varies less than 0.1 °C, is generally  
260 located below 15 m, depending on the thermal diffusivity and the annual GST amplitude of the site (Alexeev et al., 2007; Cermak et al., 2014; Burke et al., 2020; García-Pereira et al., 2024b). Therefore, as the 5L and 11L subsurface temperature bottom layers lie above this depth (at 6.98 and 7.86 m, respectively; Table 1), the ground temperatures at the deepest layer for 5L, and 11L, and at layer 11 for the 18L configuration are selected. To achieve a seasonal temperature amplitude that is close to zero, the remaining annual cycle at these depths is filtered out by subtracting the annual cycle of monthly mean anomalies  
265 for 1850–1900. Subsequently, permafrost detection using ZAA is applied to the data.

In order to illustrate the magnitude of the differences between the variety of configurations and vertical discretizations used in this work, the results for the MPIESM-PePE are compared with the ZAA and TTOP PE estimates stemming from Steinert et al. (2023) for an ensemble of 34 CMIP6 HIS and SSP5-8.5 experiments. Also, data provided by the Permafrost Climate Research Data Package version 3 product of the European Space Agency Climate Change Initiative database (ESApCCIV3) are used as an illustrative reference. ESApCCIV3 delivers mid- to high-latitude Northern Hemisphere (north of 30°) gridded data of permafrost temperature and presence at yearly resolution for the period 1997–2018. The database is generated by using the CryoGrid permafrost model (Westermann et al., 2023) driven and constrained by satellite and reanalysis land surface temperature data. ESApCCIV3 is computed using ensemble modeling procedures to enhance accuracy and provide estimates of uncertainty (see Obu et al., 2021). In this study, permafrost temperatures at multiple depths from ESApCCIV3 within the mask in Fig. 2a are used to estimate the temperature at the top of the permafrost layer, from which the TTOP PE is derived. For the ZAA PE, the annual mean temperature at a depth of 10 m is assumed to represent the zero annual amplitude layer and is used to compute the corresponding PE. It is important to remark that, while ESApCCIV3 provides a valuable pan-Arctic benchmark of permafrost state, it remains a observation-constrained model product with known biases in the representation of permafrost temperature, ALT, and PE (Obu et al., 2021). For this reason, it is used as a reference to illustrate the order of magnitude of PE differences between the different hydrological configurations and vertical discretizations presented in this study. However, this comparison does not attempt to provide an assessment of whether any of the MPIESM-PePE simulations are better or closer to the real Arctic permafrost state.

### 3 Results and discussion

#### 3.1 Temperature variability

The surface and subsurface temperature variability across the permafrost domain (Fig. 2a) in the MPIESM-PePE is shown in Fig. 4. The median values for MAAT, MAGST, and MAST 5 m (Fig. 4a) remain stable during the PIC period for all simulations except 5-layer WET (hereafter WET5) and 11-layer WET (WET11), which start from warmer initial conditions that gradually converge toward the colder state of the 18-layer simulation (WET18) by the end of the PIC. Note that the 5L and 11L simulations have a PIC duration of 50 years, whereas the 18L simulations are first equilibrated using a 100-year-long prePIC simulation and incorporating subsurface temperature corrections (Section 2.2). This strategy guarantees the departure from vertically uniform initial conditions for temperature at the beginning of the PIC and stable PIC temperatures for the deepest layers (not shown). For the near-surface temperatures, 18L simulations also show a stable PIC evolution and transition from the PIC to the HIS period. The thermal state established during the PIC period for all the-simulations remains largely unchanged until the second half of the 20th century, when global warming begins to intensify. Consequently, the early-historical (1850–1900) period serves as a reliable representation of the pre-industrial thermal conditions, and the late historical (1995–2014), mid-21st century (2041–2060), and late 21st century (2081–2100) conveniently represent the progressive temperature increase and subsequent warming standstill (SSP1-1.9), stabilization (SSP2-4.5), or acceleration (SSP5-8.5) depending on the SSP scenario (Fig. 4a).

Early-historical temperature differences between WET and DRY simulations are of about 3 °C for the permafrost domain median MAAT (Fig. 4b), MAGST (Fig. 4d), and MAST 5 m (Fig. 4f) values. This agrees with previous results by de Vrese et al. (2023), who demonstrated that WET permafrost soils exhibit higher latent heat flux values than DRY, enhancing evapotranspiration and subsequently increasing Arctic cloud cover. This, in turn, reduces the incoming solar radiation, resulting in colder temperatures for WET than DRY. Thus, the DRY configuration depicts the warmest state, with permafrost domain median (10th, 90th percentile) values of -6.0 (-11.2, 0.0) °C, -3.5 (-9.4, 2.4) °C, and -3.9 (-9.9, 2.3) °C for MAAT, MAGST, and MAST 5 m, respectively. The WET simulations show colder MAATs than REF, with a permafrost median (10th, 90th percentile) MAAT of -9.3 (-14.9, -2.4) °C for WET vs. -8.0 (-13.7, -0.7) °C for REF. However, these differences are reversed at the ground surface and subsurface, with WET MAGST and MAST 5 m being on average 1.4 and 0.3 °C warmer than REF, respectively. This is due to the enhanced snow insulation effect in winter in the WET and DRY simulations, as a result of the multi-layer snow scheme that superseded the more simplistic single-bucket scheme included in REF (see Section 2.1). Temperature differences between the 5L, 11L, and 18L vertical schemes are at most around 0.4 °C. Therefore, there is a noticeable effect of changing the hydrological state with the DRY/WET configurations but no evidence of the vertical discretization affecting surface temperature variability, as reported in previous studies with the MPI-ESM by González-Rouco et al. (2021), and García-Pereira et al. (2024a).

Permafrost warming increases during the late historical period to reach median MAATs 1.3–2.2 °C higher than in the early-historical (Fig. 4b), depending on the MPIESM-PePE member. Overall, the temperature increase for MAAT, MAGST, and MAST 5 m is more intense for the WET simulations, with median values being significantly higher in 1995–2014 than in 1850–1900 (1.5–2.2 °C,  $p < 0.05$ ). This increment is slightly smaller for DRY, but also significant (1.3–1.9 °C). The combined effect of the 20th and early 21st temperature changes translates into decadal historical trends during 1850–2014 of about 0.1 °C dec<sup>-1</sup> for MAAT, MAGST, and MAST 5 m (Fig. 4c,e,g), which are greater for WET and REF than for DRY in every case. The historical warming is followed for all the MPIESM-PePE simulations by an overall permafrost median warming trend of around 0.8 °C dec<sup>-1</sup> in the 21st century for MAAT under SSP5-8.5, 0.4 °C dec<sup>-1</sup> under SSP2-4.5, and non-significant MAAT trends for SSP1-1.9 (Fig. 4c). Moreover, temperature variability increases with the intensity of the SSP scenario, with the interdecile range being around 0.1 °C dec<sup>-1</sup> for all simulations in the SSP1-1.9 scenario, and ranging from 0.3 to 0.4 °C dec<sup>-1</sup> for SSP5-8.5 simulations. Temperature trends are slightly smaller and spatial variability is enhanced at the ground surface. The increase in MAGST median trends and spatial variability varies with the configuration considered. For instance, while the MAGST trend values for the WET SSP5-8.5 simulations show a median (10th, 90th percentile) value of 0.70 °C dec<sup>-1</sup> (0.42–1.13 °C dec<sup>-1</sup>), the DRY simulations have a median of 0.76 °C dec<sup>-1</sup> (0.54–1.04 °C dec<sup>-1</sup>), and the REF median values increase to 0.83 °C dec<sup>-1</sup> (0.66–0.99 °C dec<sup>-1</sup>; Fig. 4e,f). The trends in the SSP5-8.5 experiments yield MAGST median values of permafrost by the end of the century that are the highest for the DRY configuration, with at least 90 % of the permafrost domain (Fig. 2a) experiencing MAGST values that are above the melting point in 2081–2100 (Fig. 4d).

Regarding the effects of having different vertical discretizations, ~~there is a noticeable impact of using the 18L scheme across all configurations in the case of MAST at 5 m under SSP5-8.5 warming~~ differences between the 5L, 11L, and 18L schemes can be identified across several configurations and scenarios. However, the only signal that is robust and consistent across REF,

WET, and DRY is the attenuation of the MAST 5 m warming trend under the higher radiative forcing scenarios (SSP2-4.5 and SSP5-8.5). Deepening JSBACH ~~leads to an overall reduction in the SSP5-8.5 warming trend of~~ systematically reduces the SSP5-8.5 warming trend at 5 m depth by approximately  $0.1\text{ }^{\circ}\text{C dec}^{-1}$  at 5 m depth, indicating that a deeper LSM dampens subsurface temperature responses to surface warming. This result is consistent with previous findings by González-Rouco et al. (2021) with the JSBACH in standalone experiments, who showed that increasing LSM depth results in weaker subsurface warming. However, the impact of changing LSM depth is not significant for MAAT or MAGST. Overall, there is a clear effect of changing the hydrological state in the DRY/WET configurations, but no evidence that vertical discretization affects surface temperature variability. This agrees with earlier conclusions from González-Rouco et al. (2021) and García-Pereira et al. (2024a).

A considerable reduction in permafrost temperature variability, especially noticeable for MAGST and MAST 5 m (Fig. 4d,f), starts at the beginning of the 21st century for the SSP5-8.5 scenario. Interdecile ranges are reduced by a factor of 1.08–1.21 for 2041–2060, and 1.14–1.49 for 2081–2100 with respect to 1995–2014, depending on the MPIESM-PePE configuration. This reduction is likely produced by a decrease in spatial near-surface temperature heterogeneity associated with the retreat of the snow cover area in winter. Fig. 5a explores this by showing the variability of the winter offset (see Section 2.3.1, boxplots) and snow depth (dots) in the permafrost domain for the same periods and scenarios as in Fig. 4. It can be seen that both snow cover and winter offset barely change in the historical, SSP1-1.9, and SSP2-4.5 periods, with very stable snow depth median values of 7.7 (REF), 7.2 (DRY), and 7.9 (WET) cm. A small decrease of around 1 cm is shown for SSP5-8.5 by the late 21st century. Even though snow depth values are quite similar between the three MPIESM-PePE configurations, winter offset shows remarkable differences between the REF, and the WET and DRY experiments. The permafrost domain median winter offset reaches 10–10.5  $^{\circ}\text{C}$  for WET and DRY, but is 8–9  $^{\circ}\text{C}$  lower in REF. This contrast extends to the spatial variability: the interdecile range (box width in Fig. 5a) ranges from 6–8  $^{\circ}\text{C}$  in WET and DRY, but is only of about 1  $^{\circ}\text{C}$  for REF. ~~Both results reflect weaker insulation in the standard snow model.~~ Moreover, WET winter offset values are systematically about 0.5  $^{\circ}\text{C}$  higher than DRY, which is consistent with WET being colder (Fig. 4) and having more snow cover. For the three configurations, thinning snow depth leads to less winter offset, being of 9.1  $^{\circ}\text{C}$  for WET and 8.4  $^{\circ}\text{C}$  for DRY by the end of the SSP5-8.5 period.

To better assess the dependency of the winter offset on snow depth for every simulation, a composite is computed for the December-January-February season by aggregating values of snow depth vs. winter offset for every grid cell and year during 1850–2100, including information of the historical period and the three SSP scenarios (Fig. 5b). Snow depth was binned at 0.2 cm intervals for depths below 2 cm and 1 cm intervals above, with the winter offset averaged within each bin. This approach allows for reducing the heterogeneity in snow depth values associated with its temporal and spatial variability. The WET and DRY simulations show a strong relationship between winter offset and snow depth, starting at 5  $^{\circ}\text{C}$  for a 2 cm snow cover and rising by 1  $^{\circ}\text{C cm}^{-1}$  up to 10 cm of snow, followed by a slower rate of 0.5  $^{\circ}\text{C cm}^{-1}$  for deeper snow cover values. In contrast, winter offset in REF remains significantly lower than in the WET and DRY simulations, with a maximum of only 1.8  $^{\circ}\text{C}$  at 7–8 cm of snow depth and an increase of just 0.25  $^{\circ}\text{C cm}^{-1}$ . For snow depth values greater than 8 cm, the offset steadily declines at a rate of 0.07  $^{\circ}\text{C cm}^{-1}$ . These differences are primarily due to the implementation of a multi-layer snow scheme in WET and

DRY (Section 2.1), which enables more realistic thermal insulation. As the snow depth increases, new layers are added every 5 cm up to a maximum of five, beyond which the top layer continues to accumulate snow (Ekici et al., 2014). Each layer's temperature is explicitly computed, and energy transfer between layers occurs via conduction. ~~Since heat conductivity of snow is very low (in the range of  $0.25 \text{ WK}^{-1}\text{m}^{-1}$ ) but not so the volumetric heat capacity (around  $6 \times 10^5 \text{ Jm}^{-3}\text{K}^{-1}$ ), the resulting thermal diffusivity is low, hindering heat transfer and effectively insulating the soil beneath the snow cover.~~ As the thermal diffusivity of snow is low due to its low heat conductivity (in the range of  $0.25 \text{ WK}^{-1}\text{m}^{-1}$ ) and relatively high volumetric heat capacity (around  $6 \times 10^5 \text{ Jm}^{-3}\text{K}^{-1}$ ), it hinders heat transfer and effectively insulates the soil beneath the snow cover. REF lacks this vertical layering. Instead, snow depth only modifies the surface albedo and thermal properties of the uppermost soil layers (Reick et al., 2021). As a result, REF winter offset values remain small and less responsive to snow depth, as opposed to WET and DRY. These contrasting behaviors and the importance of layered snow schemes for capturing snow insulation effects have also been highlighted in standalone JSBACH simulations (Ekici et al., 2014; Steinert et al., 2021b).

In addition to the relationship between snow depth and winter temperature offset in the WET and DRY simulations, Fig. 5b shows that a temperature offset also ~~exists-even occurs~~ in the absence of snow for these configurations. This offset, which is not present in REF, results from the insulation effect of the organic layer in the MPIESM-PePE simulations (see Section 2.1; de Vrese et al., 2023). The organic layer helps to smooth GST variability by dampening the response to SAT changes (Fig. 6). Specifically, it enhances the snow insulation effect in winter while mitigating GST increases in summer (Fig. 6a). As a result, the combined insulation effects of both the snow and the organic layer, along with latent heat exchanges from permafrost thawing and refreezing in summer and winter, lead to a reduced amplitude of the annual GST cycle in WET and DRY compared to REF (Fig. 6b). Thus, the permafrost domain median (10th, 90th percentiles) GST–SAT amplitude ratio in 1850–1900 is of 62 (50–72) % for WET11 and WET18, 65 (52–74) % for WET5, DRY11, and DRY18, 66 (53–75) %, but 95 (93–97) % for REF (Fig. 6c). The differences in the GST annual amplitude can influence ALT (Chadburn et al., 2015) and PE (Burke et al., 2020) variability.

### 3.2 Active layer thickness

The spatial and temporal changes for ALT in the MPIESM-PePE simulations are represented in Fig. 7. Both the ALT permafrost domain median, quartiles, and 10th and 90th percentiles are the smallest for WET during the historical and scenario periods. This is due to the permafrost temperatures being on average  $3^\circ\text{C}$  and  $1.5^\circ\text{C}$  colder than for the DRY and REF simulations (Fig. 4), respectively. However, despite REF having colder MAAT, MAGST, and MAST 5 m than DRY (Fig. 4b,d,f, respectively), REF median ALT is approximately 1 m deeper than DRY in the historical period. Figure 7b,c further shows ALT for REF is greater than for DRY for ca. 70 % of the gridcells in 1850–1900, and 60 % in 1995–2014. The higher values of ALT in the REF simulations can be explained by the combination of two factors. Although REF MAGST is lower than DRY in 1850–1900 (Fig. 4b), summer GST is considerably higher (Fig. 6b). This is due to the lack of organic layer insulation and the absence of water phase changes with temperature in this configuration. ~~In the WET and DRY simulations, when~~ When soil temperatures rise above  $0^\circ\text{C}$  further warming is partially prevented, since an amount of heat is absorbed by the ice melt as latent heat. This phenomenon is known as the zero-curtain effect (Outcalt et al., 1990; Mottaghy and Rath, 2006). Since the REF simulations do

not include phase changes, summer warming is capable of penetrating deeper into the soil. The DRY configuration, in contrast, counterbalances the stronger temperature increase in summer with the zero-curtain effect, yielding lower ALT values than REF. Still, around 30 % of gridcells for DRY have a thicker ALT than for REF in 1850–1900, increasing to 40 % in 1995–2014 (note the value of the x-axis when DRY and REF curves intersect). This indicates the zero-curtain effect does not outweigh the greater heat propagation from higher MAGST at depths below 4 (1850–1900, Fig. 7b) to 4.5 m (1995–2014, Fig. 7c), a depth at which the annual cycle is very attenuated (Fig. 6b). Moreover, permafrost soils are very shallow across large areas of Siberia in JSBACH (Steinert et al., 2021b), so the absence of deep soil ice in these areas, and therefore of deep zero-curtain effect, promotes further ALT deepening as a response to the relative higher summer temperatures (Fig. 6b) in DRY than REF. The WET simulations, which yield colder permafrost temperatures (Fig. 4b,c,e) and more soil moisture (de Vrese et al., 2023), shows the highest fractions of gridcells holding permafrost for any value of ALT.

As permafrost warming intensifies in the 21st century under all scenarios, especially the SSP5-8.5, ALT depth increases for all the simulations. That is evidenced from an extreme deepening of the ALT in 2081–2100 under the SSP5-8.5 scenario compared to 1995–2014, with permafrost domain median differences of 4.1 for REF, 5.8 for WET, and 4.4 m for DRY for all configurations (Fig. 7a, Table 2). These strong ALT variations are also illustrated by long-term changes in the vertical distribution of permafrost (Fig. 7d,e), including near-surface (i.e., shallower than 3 m, Fig. 8), and deep permafrost (Fig. 9). Even though most of the gridcells have a deeper ALT for REF than for DRY up to the late-historical period 1995–2014, the DRY simulations exhibit a greater deepening of ALT during the 21st century. That entails that only 36 % of the gridcells for DRY5, 36 % for DRY11, and 44 % for DRY18 still hold permafrost at the summer maximum thawing by 2081–2100 (Fig. 7e); these values are 18–20 % lower than those of REF, with a proportion of gridcells that still hold permafrost in 2081–2100 for DRY around 20 % lower than for REF (Fig. 7e, Table 2).

**Table 2.** Permafrost domain median ALT values (in m) in 1850–1900, 1995–2014, and 2081–2100 in the SSP1-1.9, SSP2-4.5, and SSP5-8.5 scenarios for the nine simulations of the MPIESM-PePE. Fraction of gridcells that have permafrost in 1850–1900, 1995–2014, and 2081–2100 under the SSP5-8.5 scenario (see Fig. 7e) and its equivalence in deep PE in 2100 (in  $10^6 \text{ km}^2$ , see Fig. 9) is also provided.

|       | Median ALT (m) |           |           |          |          | Permafrost gridcells (%) |           |           | Deep PE ( $10^6 \text{ km}^2$ ) |
|-------|----------------|-----------|-----------|----------|----------|--------------------------|-----------|-----------|---------------------------------|
|       | HIS            |           | SSP1-1.9  | SSP2-4.5 | SSP5-8.5 | HIS                      |           | SSP5-8.5  | SSP5-8.5                        |
|       | 1850–1900      | 1995–2014 | 2081–2100 |          |          | 1850–1900                | 1995–2014 | 2081–2100 | 2100                            |
| REF5  | 2.68           | 3.62      | 4.13      | 5.79     | 6.98     | 96                       | 84        | 54        | 5.18                            |
| REF11 | 3.01           | 3.73      | 4.35      | 6.51     | 7.86     | 96                       | 85        | 57        | 4.46                            |
| REF18 | 2.93           | 3.58      | 4.19      | 5.83     | 7.72     | 96                       | 86        | 62        | 7.88                            |
| WET5  | 0.78           | 0.85      | 1.05      | 2.68     | 6.64     | 99                       | 93        | 64        | 7.69                            |
| WET11 | 0.82           | 1.20      | 1.58      | 3.78     | 7.80     | 100                      | 92        | 63        | 6.11                            |
| WET18 | 0.81           | 1.23      | 1.33      | 2.89     | 6.35     | 100                      | 92        | 70        | 9.09                            |
| DRY5  | 2.57           | 2.68      | 3.83      | 6.65     | 6.98     | 79                       | 73        | 36        | 1.01                            |
| DRY11 | 2.52           | 3.67      | 4.94      | 7.86     | 7.86     | 79                       | 72        | 36        | 1.45                            |
| DRY18 | 2.17           | 3.26      | 4.46      | 7.63     | 7.86     | 81                       | 74        | 44        | 2.93                            |

The rapid ALT deepening for DRY is ~~portrayed by maps shown~~ in Fig. 8. While pre-2000 deeper near-surface permafrost thaw in summer (purple to red) extends over a greater area for the REF simulations, DRY is dominated by a more intense near-surface and deep permafrost thaw during the 21st century (orange). In fact, Fig. 9 shows that in 2100, a much greater area still holds deep permafrost in REF than in DRY, ~~with values of 5.18 (REF5), 4.46 (REF11), 7.88 (REF18) versus 1.01 (DRY5), 1.45 (DRY11), and 2.93 (DRY18)  $10^6$  km<sup>2</sup>~~ (see Table 2). The WET configuration yields shallower ALT and greater gridcell fraction values than REF and DRY for the whole duration of the simulation. The median WET ALT difference with respect to DRY increases during the 21st century (Fig. 7a), which agrees with higher MAGST (Fig. 4e) and MAST 5 m (Fig. 4g) trends, of 0.1–0.2 °C dec<sup>-1</sup>, for DRY than WET for the SSP5-8.5 scenario. The relatively colder temperatures results in ~~WET5, WET11, and WET18~~ the WET configuration still having permafrost in summer for ~~64, 63, and~~ around 70 % of the gridcells in 2081–2100, ~~which corresponds to a deep permafrost area of 7.69, 6.11, and 9.09  $10^6$  km<sup>2</sup> by 2100, respectively.~~ However, the SSP5-8.5 represents the most extreme warming scenario, and therefore the biggest response in ALT. The SSP1-1.9 permafrost ALT response is clearly weaker, with median values ~~from 0.2 (WET) to 1.2 (DRY)~~ 0.2–1.2 m deeper in 2081–2100 than ~~by in~~ 1995–2014 (Fig. 7a, Table 2). In turn, the SSP2-4.5 depicts a stronger ALT deepening occurrence than SSP1-1.9 in the 21st century, with an increase in ALT ranging from 2.0 (WET) to 4.2 (DRY) m (Table 2). In both SSP1-1.9 and SSP2-4.5 scenarios, the differences between the different MPIESM-PePE configurations are equivalent to those already mentioned for SSP5-8.5, but comparatively smaller.

Apart from the differences in ALT due to the MPIESM-PePE configuration, there are also notable differences produced by the LSM depth. All configurations show increasing ALT differences during the 21st century among the 5L, 11L, and 18L schemes. This is depicted by a greater increment in the ALT values for the shallow (5L, 11L) than the 18L simulations in 2041–2060, and 2081–2100 for the SSP5-8.5 scenario (Fig. 7a). The effect of the LSM depth is also evidenced by a growing offset in the gridcell fraction profiles during 2041–2060 (Fig. 7d), and 2081–2100 (Fig. 7e) between the shallow and 18L discretizations. Deepening the LSM also produces differences in deep permafrost that range from 1.4 to 3.4  $10^6$  km<sup>2</sup> in 2100, depending on the MPIESM-PePE configuration (Fig. 9, Table 2). The shallow vs. deep vertical scheme differences are greater when comparing 11L with 18L than between 5L and 18L, which indicates that using a coarse vertical resolution reduces the ALT bias produced by using a too shallow LSM depth (Chadburn et al., 2015). ~~Even so~~ Nevertheless, it is ~~expanding~~ increasing the LSM depth when resolving vertical heat diffusion the ~~factor that prevails~~ prevailing factor in damping long-term warming (González-Rouco et al., 2021; Steinert et al., 2021a; García-Pereira et al., 2024a), producing a relative cooling of 18L vs. 5L and 11L simulations.

Although LSM depth impacts the representation of deep permafrost, its near-surface effect is negligible. Fig 9b-e show no consistent differences in permafrost fraction of gridcells for historical ALT, and very small differences for gridcells having an ALT above 5 (2) m in 2041–2060 (2081–2100) under the SSP5-8.5 scenario. These small differences are also very hard to distinguish in terms of the evolution of near-surface permafrost existence, as shown in Fig. 8.

Regarding the spatial retreat pattern of permafrost in the MPIESM-PePE, both Figs. 8 and 9 show high agreement for the nine MPIESM-PePE simulations. All of them portray little to no presence of near-surface permafrost in Southern Siberia, Southern Canada, the Labrador Peninsula, and Scandinavia. The warming in the 20th and the first-half of the 21st century under the

SSP5-8.5 scenario affects much more near-surface than deep permafrost, yielding its retreat in Southern Alaska, Western, and Southeastern Siberia. [These regions lie near the southern margin of the permafrost zone, where comparatively warmer surface conditions make them more sensitive to permafrost retreat and active layer thickening under climate change \(e.g., Biskaborn et al., 2019; Rantanen et al., 2022\).](#) By 2081–2100 for the SSP5-8.5 scenario, near-surface permafrost only remains in Northern Siberia for the WET simulations. In contrast, deep permafrost extent varies considerably with the MPIESM-PePE configuration and vertical discretization, being maximum for WET18, where permafrost in 2100 is still present in Northern Canada and a vast area across Central, Eastern, and Northern Siberia, and minimum for DRY5 and DRY11, where only a small patch of deep permafrost remains in the Central Siberian highlands.

### 3.3 Permafrost extent estimates: multi-model comparison

**Table 3.** Initial state of permafrost extent ( $PE_{1850-1900}$ , in  $10^6 \text{ km}^2$ ), and trends in the SSP1-1.9, SSP2-4.5, and SSP5-8.5 scenarios (in  $10^6 \text{ km}^2 \text{ dec}^{-1}$ ) for the nine simulations of the MPIESM-PePE and the 10th, 25th, 50th, 75th, and 90th percentiles of the CMIP6 ensemble stemming from Steinert et al. (2023). Values for both ZAA and TTOP definitions. Non-significant ( $p > 0.05$ ) SSP trends are marked in bold.

|                      | $PE_{1850-1900}$      |       | Trend $PE_{SSP1-1.9}$                  |              | Trend $PE_{SSP2-4.5}$                  |       | Trend $PE_{SSP5-8.5}$                  |       |
|----------------------|-----------------------|-------|----------------------------------------|--------------|----------------------------------------|-------|----------------------------------------|-------|
|                      | $(10^6 \text{ km}^2)$ |       | $(10^6 \text{ km}^2 \text{ dec}^{-1})$ |              | $(10^6 \text{ km}^2 \text{ dec}^{-1})$ |       | $(10^6 \text{ km}^2 \text{ dec}^{-1})$ |       |
|                      | ZAA                   | TTOP  | ZAA                                    | TTOP         | ZAA                                    | TTOP  | ZAA                                    | TTOP  |
| WET5                 | 19.92                 | 20.33 | -0.11                                  | -0.16        | -0.71                                  | -0.68 | -1.49                                  | -1.34 |
| WET11                | 20.29                 | 20.43 | -0.10                                  | -0.08        | -0.75                                  | -0.72 | -1.42                                  | -1.39 |
| WET18                | 20.32                 | 20.37 | -0.09                                  | <b>-0.01</b> | -0.64                                  | -0.63 | -1.15                                  | -1.35 |
| DRY5                 | 15.22                 | 15.49 | -0.15                                  | -0.10        | -0.84                                  | -0.90 | -1.62                                  | -1.65 |
| DRY11                | 15.12                 | 15.30 | -0.13                                  | -0.10        | -0.73                                  | -0.71 | -1.52                                  | -1.57 |
| DRY18                | 15.87                 | 16.05 | -0.15                                  | -0.11        | -0.67                                  | -0.70 | -1.38                                  | -1.58 |
| REF5                 | 20.52                 | 20.78 | <b>-0.04</b>                           | -0.06        | -0.61                                  | -0.61 | -1.43                                  | -1.40 |
| REF11                | 19.73                 | 20.01 | -0.08                                  | -0.08        | -0.59                                  | -0.57 | -1.45                                  | -1.44 |
| REF18                | 20.03                 | 20.26 | -0.16                                  | -0.15        | -0.55                                  | -0.62 | -1.09                                  | -1.36 |
| CMIP6 <sub>P10</sub> | 6.68                  | 13.47 |                                        |              |                                        |       | -0.39                                  | -1.21 |
| CMIP6 <sub>P25</sub> | 12.91                 | 16.07 |                                        |              |                                        |       | -1.33                                  | -1.60 |
| CMIP6 <sub>P50</sub> | 17.20                 | 20.22 |                                        |              |                                        |       | -1.79                                  | -1.57 |
| CMIP6 <sub>P75</sub> | 21.01                 | 24.75 |                                        |              |                                        |       | -1.24                                  | -1.49 |
| CMIP6 <sub>P90</sub> | 27.66                 | 28.80 |                                        |              |                                        |       | -1.25                                  | -1.36 |

The time evolution of PE for the nine simulations of the MPIESM-PePE in the three SSP scenarios are compared with the PE values derived from the ESApCCIv3 and CMIP6 PE estimates provided by Steinert et al. (2023) in Fig. 10. Both the MPIESM-PePE and the CMIP6 ensemble show a mostly constant permafrost area since 1850–1900 (see values in Table 3) until the last decades of the 20th century. WET and REF depict very similar early-historical PE values for both ZAA (Fig. 10a)

and TTOP (Fig. 10b), in the range of  $19.9\text{--}20.4 \times 10^6 \text{ km}^2$  and  $19.7\text{--}20.8 \times 10^6 \text{ km}^2$ , respectively (Table 3). However, REF5 and REF11 ZAA, and REF11 TTOP estimates show more variability, with the periods of the largest PE values coinciding with the strongest volcanic cooling events in the historical period (e.g., an abrupt PE rise in the 1880s as a response to the Krakatoa eruption). In contrast, the DRY simulations depict an area around  $5 \times 10^6 \text{ km}^2$  smaller than the REF and WET ones for both ZAA and TTOP (Table 3) and less temporal variability.

In general, CMIP6 PE median values lie in between WET and DRY when the ZAA permafrost definition is used, and agree with WET and REF for the TTOP definition. As ZAA and TTOP definitions physically target deep and near-surface permafrost degradation, respectively, this result indicates that the CMIP6 yields larger near-surface than deep permafrost estimates, which are approximately  $20$  and  $17 \times 10^6 \text{ km}^2$ , respectively. The interquartile range in CMIP6 PE is similar for ZAA and TTOP, of around  $8 \times 10^6 \text{ km}^2$ , but greater for the deep permafrost (ZAA) when considering the interdecile range, of around  $20\text{--}22 \times 10^6 \text{ km}^2$ . Considering the CMIP6 PE interdecile range as representative of the CMIP6 ensemble spread, the differences in the PE early-historical equilibrium state associated with imposing the WET or DRY conditions in northern high latitudes account for a quarter of it. This suggests that the representation of soil hydrology plays an important role in determining the current and future extent of large-scale permafrost. In fact, state-of-the-art ESMs have been claimed to have difficulties in reproducing nowadays PE, due to their limitations in representing land surface processes (Chadburn et al., 2015; Burke et al., 2020; Steinert et al., 2023), the non-inclusion of specific processes accelerating permafrost degradation at local scales, such as soil subsidence (Andresen et al., 2020), talik formation or thermokarst (Farquharson et al., 2019; Nitzbon et al., 2024), or the inability to account for the permafrost-carbon warming feedback (Burke et al., 2013). These observation-simulation discrepancies in PE are suggested in Fig. 10, which exhibits that PE values from ESAPCCiv3 in 1997–2019 are not only lower than the CMIP6 median, but also lower than all the MPIESM-PePE experiments, both for ZAA and TTOP. The differences are especially large for ZAA, which might be more sensitive to the aforementioned degradation processes enhancing deep permafrost losses that are not implemented in current LSMs, including JSBACH. However, the observation–simulation comparison presented here must be interpreted with caution. The ESAPCCiv3 product is not a purely observational dataset but an assimilation product, and it has documented biases in the representation of permafrost temperature, ALT, and PE (Obu et al., 2021). Nonetheless, the comparison with ESAPCCiv3 depicts the scale of changes in our different simulated states with respect to observation-constrained estimates. Also, the broad spread in PE values across ESAPCCiv3, CMIP6 models, and MPIESM-PePE experiments illustrates the substantial uncertainty that still surrounds our knowledge of the current permafrost state, of which soil hydrology is only one contributing factor.

The steady state of PE in the 19th and most of the 20th century is followed by a decline that starts in the last quarter of the 20th century. The severity of the PE retreat mainly depends on the intensity of the SSP forcing (Fig. 10, Table 3). Hence, the low forcing SSP1-1.9 scenario shows a stabilization of PE after 2030 for the nine simulations of the MPIESM-PePE, with similar values for ZAA and TTOP DRY of around  $12 \times 10^6 \text{ km}^2$  and of  $16 \times 10^6 \text{ km}^2$  for WET and REF in their three vertical discretizations. SSP2-4.5 depicts a moderate loss until 2070 and a subsequent stabilization, with near-surface permafrost area values (TTOP, Fig. 10b) of  $8 \times 10^6 \text{ km}^2$  for DRY and  $12 \times 10^6 \text{ km}^2$  for WET and REF. For this scenario, there are moderate differences among 5L, 11L, and 18L simulations for ZAA, with PE values that differ by  $1\text{--}2 \times 10^6 \text{ km}^2$  by the end of the 21st

century. Thus, DRY18 and WET18 PE are  $8.6$  and  $13.1 \cdot 10^6 \text{ km}^2$  in 2081–2100, but  $7.3$  and  $11.1 \cdot 10^6 \text{ km}^2$  for DRY11 and WET11, respectively. These differences in the deep permafrost are due to the increase in LSM depth better accommodating for the subsurface propagation of the 21st century warming (González-Rouco et al., 2021; Steinert et al., 2021a; García-Pereira et al., 2024a). ~~Moreover, the PE loss shows no stabilization either for MPIESM-PePE or for the CMIP6 simulations under the SSP5-8.5 scenario~~ Under the SSP5-8.5 scenario, the PE loss shows no signs of stabilising, either for the MPIESM-PePE or the CMIP6 simulations. The deep permafrost (Fig. 10a) is intensely degraded for the 5L and 11L simulations, with PE around  $2.1$ ,  $6.4$ , and  $6.9 \cdot 10^6 \text{ km}^2$  in 2081–2100, for DRY, WET, and REF, respectively. The degradation is less pronounced for the 18L experiments, with PE values of  $4.0$ ,  $9.5$ , and  $9.8 \cdot 10^6 \text{ km}^2$ , respectively. These PE differences by the late 21st century are consistent with differences in trends, which are between  $0.14$  and  $0.36 \cdot 10^6 \text{ km}^2 \text{ dec}^{-1}$  smaller for the 18L than for the 5L and 11L simulations (Table 3). In contrast, near-surface permafrost degradation, although also intense, does not show notable differences between the 5L, 11L, and 18L schemes, with DRY TTOP by 2081–2100 being very similar for the three discretizations ( $2.4$ – $2.9 \cdot 10^6 \text{ km}^2$ ), WET and REF being in the range of  $6.9$ – $8.7 \cdot 10^6 \text{ km}^2$ , and the CMIP6 median PE being slightly lower than WET and REF ( $6.1 \cdot 10^6 \text{ km}^2$ ). Both the 10th and 25th percentiles of the CMIP6 ensemble ZAA and TTOP PE are below than  $1 \cdot 10^6 \text{ km}^2$  for the last three decades of the 21st century, which entails that near-surface permafrost is almost fully thawed for at least a quarter of CMIP6 models by the end of the century for the highest radiative forcing scenario.

Figure 11 shows the spatial distribution of PE depicted in Fig. 10 for the historical and SSP5-8.5 scenario periods. At least 90 (75) % of CMIP6 models, all MPIESM-PePE experiments, and the observations display near-surface (deep) permafrost in Fig. 11b (Fig. 11f) across central, western, and eastern Siberia, southern Siberia highlands, northern Alaska and Canada, and Nunavut mainland and island territories by the end of the historical period (1995–2014). There is a notable decline in PE since 1995–2014 for both definitions under the SSP5-8.5 scenario (Fig. 11c,d,g,h). Permafrost degradation is more intense in southern Siberia, Fennoscandia, and the Labrador Peninsula, ~~warmer areas lying near the southern margin of the permafrost zone. By 2041–2060, with~~ only a quarter of the CMIP6 members and REF18, WET11, and WET18 still showing some permafrost patches in southern Siberia ~~by 2041–2060~~ (Fig. 11c,g). By the end of the century (2081–2100, Fig. 11d,h), the existence of permafrost is predominantly reduced to central and northern Siberia, the highlands in southern Siberia, northern Alaska, and the island territories of Nunavut. In these regions, at least half of CMIP6 models, and WET, REF, and DRY18 maintain some deep permafrost. In contrast, in DRY5 and DRY11 simulations near-surface and deep permafrost existence is reduced to Northern Siberia. In the WET18 and REF18 simulations, deep permafrost covers a much larger area in central Siberia and northern Canada than near-surface permafrost. Likewise, the DRY18 deep permafrost extends across vast areas in western Siberia, while near-surface permafrost is confined to eastern Siberia. Overall, the retreat pattern of PE for ZAA for 5L and 11L simulations is similar to TTOP, which indicates that shallow LSMs yield similar near-surface and deep permafrost degradation, whilst deep LSMs delay deep permafrost thawing as a response to intense surface warming (see also Fig. 10b).

### 3.4 Permafrost extent and LSM depth

To better disentangle the potential role of the LSM depth from soil hydrology in producing differences in near-surface and deep permafrost estimates, the ZAA and TTOP definitions are compared in Fig. 12. As already exhibited in Fig. 10, the WET and

540 REF simulations depart from a higher PE value in 1850–1900 than the DRY for both PE definitions. CMIP6 TTOP PE values are very similar to WET and REF, and ZAA PE closer to DRY. In general, the CMIP6 median always shows larger TTOP than ZAA PE values (golden line in Fig. 12a being above the grey diagonal). In contrast, all the MPIESM-PePE 18L simulations show negative values of ZAA - TTOP. Even though the DRY18 simulation starts from a smaller PE than WET18 and REF18, and further diverges by 2100, the evolution for the all three is quite robust in yielding a larger ZAA PE. However, this is not  
 545 the case for the 5L and 11L simulations, which show a very close agreement in the evolution of TTOP and ZAA permafrost for REF simulations, and slightly greater TTOP than ZAA PE values by the end of the 21st century for WET and DRY. This entails that the MPIESM-PePE simulations with a shallow LSM (see Table 1 or Figs. 1,2b) experience a more intense deep permafrost thawing, as previously noted in Figs. 10b and 11h. This factor also explains the asymmetry observed in the median and spread of TTOP and ZAA PE estimates from CMIP6 simulations (Figs. 10,12), as previously remarked by Steinert et al.  
 550 (2023).

In fact, Burke et al. (2020) stated that having too shallow LSMs could result in an overestimation of deep permafrost thawing (i.e., an underestimation of PE with warming) using the ZAA metric. Since the zero-annual temperature amplitude layer is usually found below 15 m (e.g., García-Pereira et al., 2024b), LSMs imposing zero-flux bottom boundary conditions above this depth would not experience a full attenuation of the annual cycle at any level. For that reason, we apply a filter to remove the  
 555 remaining annual temperature cycle in the lowest soil layer in the 5L and 11L simulations (see Section 2.3.3). However, this is not the only caveat about shallow LSMs. An insufficient LSM depth produces an overestimation of the temperature variability with depth, as shown in González-Rouco et al. (2021) with the same modeling framework. This results in the amplification of the subsurface temperature trends associated with global warming. A LSM depth of at least 170 m has to be imposed to ensure the full thermal decoupling between the ground surface and the bottom layer response in simulations spanning the historical +  
 560 SSP period (1850–2100). This is not the case for any LSM incorporated in CMIP6 ESMs, as represented by Fig. 12b, which depicts the differences in deep vs. near-surface PE. There is an apparent bias in ZAA - TTOP for CMIP6 models having a very shallow land component, i.e., with a LSM depth at 3–5 m ( $-3.1 \pm 3.7 \cdot 10^6$  for 1995–2014, and  $-6.7 \pm 3.6 \cdot 10^6 \text{ km}^2$  for 2081–2100). This bias still exists for a subgroup of models having LSM depths of around 10–15 m, which includes MPIESM-PePE 5L and 11L simulations ( $-0.4 \pm 3.9 \cdot 10^6$  for 1995–2014, and  $-0.3 \pm 1.4 \cdot 10^6 \text{ km}^2$  for 2081–2100). ESMs having relatively deep  
 565 LSM components ( $> 40 \text{ m}$ ) exhibit negligible ZAA - TTOP differences and dispersion ( $-0.2 \pm 0.4 \cdot 10^6$  for 1995–2014, and  $0.3 \pm 0.7 \cdot 10^6 \text{ km}^2$  for 2081–2100), excluding IPSL-CM6A-LR (LSM depth of 90 m) in 2081–2100, which is an evident outlier ( $-7.3 \cdot 10^6 \text{ km}^2$ ). This deep cluster also includes WET18, DRY18, and REF18 simulations, which have deep vs. near-surface PE differences of 0.3, 0.2, and 0.0 for 1995–2014, and 1.5, 1.5, and  $1.8 \cdot 10^6 \text{ km}^2$  for 2081–2100, respectively. This reduction in ZAA–TTOP bias with increasing LSM depth is consistent with Steinert et al. (2023), who reported greater agreement in PE  
 570 estimates by the end of the historical period across different permafrost definitions for CMIP6 models with deep LSMs. The little impact observed in ZAA - TTOP PE when including a deep LSM in a set of simulations with varying conditions of Arctic soil wetness confirms the relevance of LSM depth to yield unbiased PE estimates under global warming conditions.

## 4 Summary and conclusions

Permafrost degradation is a critical element for the sensitivity of the Earth system to climate change because of its significant impact on the terrestrial carbon cycle, society, and ecosystems. Despite observational data being important for monitoring the ongoing state of permafrost soils, their coverage is often spatially sparse and uneven in time at a global scale. ESMs can overcome these spatial and temporal limitations, as well as provide comprehensive insights into permafrost dynamics. However, state-of-the-art ESMs exhibit considerable inter-model variability in their projections of PE and ALT (Burke et al., 2020; Steinert et al., 2023). These differences are mainly due to their different representation of soil thermodynamics and hydrology (Andresen et al., 2020). This study assesses the impacts of changing both soil thermodynamics and hydrology on the Arctic permafrost dynamics for the first time in a fully-coupled ESM. This is achieved by modifying the LSM, including progressively refined and deeper vertical discretizations of the LSM (5L, 11L, and 18L) and generating comparatively wetter (WET) or dryer (DRY) conditions to quantitatively capture the permafrost conditions induced by changes in the LSM hydrological state (de Vrese et al., 2023).

Our results highlight the relevance of the representation of soil hydro-thermodynamic processes in determining permafrost temperature, active layer thickness, and permafrost extent. Differences in the permafrost water table can produce very different thermal states, with MAAT differences of more than 3 °C between WET and DRY configurations. However, differences in the GST seasonal variability are mainly produced by an improved representation of the snow and organic layer insulation effects, included in WET and DRY but not represented in REF. Higher summer temperatures, together with the omission of latent heat absorption due to permafrost thawing, leads to REF ALT being on average 1 to 2 m deeper in 1850–1900 than DRY and WET, respectively. These differences decrease with warming, with the DRY simulation experiencing the greatest ALT values by the late 21st century under the SSP2-4.5 and SSP5-8.5 scenarios. Apart from the already mentioned factors, LSM depth plays also a substantial role in determining future ALT. Deepening the LSM, i.e., switching from the 5L or 11L to the 18L discretization, prevents the full degradation of the permafrost profile for  $\sim 6\text{--}8\%$  of the gridcells within the permafrost mask, which is equivalent to an area of  $1.5\text{--}2 \times 10^6 \text{ km}^2$ .

The implications of soil hydrology and LSM depth are also remarkable for the evolution of the PE. Both TTOP, which physically represents near-surface PE, and ZAA, which tracks deep PE evolution, yield an initial PE of around  $15 \times 10^6 \text{ km}^2$  for DRY and  $20 \times 10^6 \text{ km}^2$  for REF and WET, respectively. These differences explain up to 76 % of the interquartile range for CMIP6 deep PE, 54 % for near-surface PE, and about a quarter of the interdecile range for both, highlighting the uncertainty in current PE estimates associated with the representation of soil hydrology. [Considering that permafrost region contains  \$1035 \pm 150 \text{ PgC}\$  over  \$17.6 \times 10^6 \text{ km}^2\$ , which corresponds to an areal mean of  \$58.8 \pm 8.5 \times 10^{-6} \text{ PgC km}^{-2}\$  \(Hugelius et al., 2014\), the differences between WET and DRY near-surface PE values would roughly amount for 250–340 PgC. This is around 40 % of the carbon currently stored in the atmosphere, a large fraction that indicates the importance of uncertainties in frozen-soil hydrology for permafrost carbon storage, with high-impact implications for the permafrost–carbon feedback and, in turn, climate change.](#)

Regarding the LSM depth, the 5L and 11L with respect to the 18L discretization show consistent PE differences in the three hydrological configurations, of  $1.9\text{--}3.1 \cdot 10^6 \text{ km}^2$ . This emphasizes the relevance of LSM depth in reducing deep permafrost degradation. In fact, CMIP6 models ~~having~~with LSMs deeper than 40 m show the smallest ~~ZAA–TTOP PE differences~~differences in ZAA – TTOP PE, indicating that ESM simulations with deep LSM vertical schemes are ~~more reliable to project~~better at projecting the future degradation of deep permafrost and its sensitivity to surface warming. Furthermore, because ice-rich, carbon-bearing deposits (e.g., Yedoma) can extend well below 10 m, using deeper LSMs reduces bottom-boundary warming and leads to more realistic estimates of the vulnerability of these deep soil organic carbon reservoirs.

Finally, the comparison of PE values from the MPIESM-PePE with observation-constrained ESAPCCIv3 estimates and other CMIP6 simulations illustrates how the range of permafrost response that can stem in one model from different assumptions in permafrost hydrology contributes to model uncertainties. Furthermore, observation-constrained products based on data assimilation, i.e., models constrained by observational evidence, are not free from model biases resulting from the decisions on parametrizations or physical processes included. The latter emphasizes the need for more observations and assessment of uncertainties in observation-based products.

*Data availability.* Yearly-resolution temperature, ALT, and PE data from the MPIESM-PePE are available upon request and will be included in a Zenodo repository following the manuscript revision process. Monthly-resolution outputs and other variables of interest can also be provided upon request. The CMIP6 PE data used in this study are publicly available and can be accessed at the following Zenodo repository: <https://doi.org/10.5281/zenodo.10103000> (last access: 22 April 2025). ESAPCCI data are also publicly available and can be downloaded from <https://catalogue.ceda.ac.uk/uuid/6e2091cb0c8b4106921b63cd5357c97c/> (last access: 22 April 2025).

*Author contributions.* FGP conceptualized the study. JFGR, FGP, and NJS prepared the MPIESM-PePE experiments, with the guidance of PdV. FGP processed the output data and performed the calculations. FGP and JFGR prepared the original draft of the manuscript. All the co-authors contributed to the analysis and discussion of the results and to the revision of the draft.

*Competing interests.* PdV is a member of the Editorial board of *The Cryosphere*.

*Acknowledgements.* This work has been developed within the frame of GreatModelS (project no. RTI2018-102305-B-C21) and SMILEME (project no. PID2021-126696OB-C21) from the Spanish Ministry of Science and Innovation (MICINN). We acknowledge the support from the Scientific Network PolarCSIC, funded by the Spanish *Consejo Superior de Investigaciones Científicas* (CSIC). We would also like to thank the Deutsches Klimarechenzentrum (DKRZ) for the resources granted by its ScientiDic Steering Committee (WLA) to run the MPI-ESM1.2-LR P2k+d simulation under project ID bm1026. We also appreciate the efforts of the World Climate Research Programme's Working Group on Coupled Modeling, responsible for CMIP, and thank the contribution by Steinert et al. (2023) for deriving a valuable permafrost

extent database for CMIP6 models. Additionally, thanks to Christian Hauck for editing this paper and Stefan Hagemann and two anonymous  
635 reviewers for their proof-reading and suggestions, which substantially contributed to improve the manuscript. Finally, we would like to have  
a special mention for our colleagues at the MPI-M Veronika Gayler, Tobias Stacke, and ~~Hemlmuth~~Helmuth Haak for their valuable technical  
support.

*Financial support.* This research work has been financially supported by SMILEME (project no. PID2021-126696OBC21), funded by call  
no. MCIN/AEI/10.13039/501100011033 from the Spanish Ministry of Science and Innovation (MICINN). FGP was funded by contract no.  
640 PRE2019-090694 of the MICINN and by the Ministry for the Ecological Transition and the Demographic Challenge (MITECO) and the  
European Commission NextGenerationEU (Regulation EU 2020/2094), through CSIC's Interdisciplinary Thematic Platform Clima (PTI-  
Clima). PdV was funded by the European Research Council under the European Union's Horizon 2020 research and innovation program as  
part of the Q-Arctic project (grant agreement No 951288).

## References

- 645 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, L09 502, <https://doi.org/10.1029/2007GL029536>, 2007.
- Andresen, C. G., Lawrence, D. M., Wilson, C. J., McGuire, A. D., Koven, C., Schaefer, K., Jafarov, E., Peng, S., Chen, X., Gouttevin, I., Burke, E., Chadburn, S., Ji, D., Chen, G., Hayes, D., and Zhang, W.: Soil moisture and hydrology projections of the permafrost region – a model intercomparison, *The Cryosphere*, 14, 445–459, <https://doi.org/10.5194/tc-14-445-2020>, 2020.
- 650 Biskaborn, B., Smith, S., Noetzli, J., Matthes, H., Vieira, G., Streletskiy, D., Schoeneich, P., Romanovsky, V., Lewkowicz, A., Abramov, A., Allard, M., Boike, J., Cable, W., Christiansen, H., Delaloye, R., Diekmann, B., Drozdov, D., Etzelmüller, B., Grosse, G., and Lantuit, H.: Permafrost is warming at a global scale, *Nature Communications*, 10, 264, <https://doi.org/10.1038/s41467-018-08240-4>, 2019.
- Biskaborn, B. K., Lanckman, J.-P., Lantuit, H., Elger, K., Streletskiy, D. A., Cable, W. L., and Romanovsky, V. E.: The new database of the Global Terrestrial Network for Permafrost (GTN-P), *Earth System Science Data*, 7, 245–259, <https://doi.org/10.5194/essd-7-245-2015>, 2015.
- 655 Brown, J., Hinkel, K., and Nelson, F.: The Circumpolar Active Layer Monitoring (CALM) program, *Polar Geography*, 24, 166–258, <https://doi.org/10.1080/10889370009377698>, 2000.
- Brown, J., Ferrians, O., Heginbottom, J. A., and Melnikov., E.: Circum-Arctic Map of Permafrost and Ground-Ice Conditions, Version 2, <https://doi.org/10.7265/skbg-kf16>, 2002.
- 660 Burke, E. J., Jones, C. D., and Koven, C. D.: Estimating the Permafrost-Carbon Climate Response in the CMIP5 Climate Models Using a Simplified Approach, *Journal of Climate*, 26, 4897 – 4909, <https://doi.org/10.1175/JCLI-D-12-00550.1>, 2013.
- 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.
- Cermak, V., Bodri, L., Šafanda, J., Kresl, M., and Dedecek, P.: Ground-air temperature tracking and multi-year cycles in the subsurface temperature time series at geothermal climate-change observatory, *Studia Geophysica et Geodaetica*, 58, 403–424, <https://doi.org/10.1007/s11200-013-0356-2>, 2014.
- 665 Chadburn, S. E., Burke, E. J., Essery, R. L. H., Boike, J., Langer, M., Heikenfeld, M., Cox, P. M., and Friedlingstein, P.: Impact of model developments on present and future simulations of permafrost in a global land-surface model, *The Cryosphere*, 9, 1505–1521, <https://doi.org/10.5194/tc-9-1505-2015>, 2015.
- 670 Chen, D., Rojas, M., Samset, B., Cobb, K., Niang, A. D., Edwards, P., Emori, S., Faria, S., E. Hawkins, P. H., Huybrechts, P., Meinshausen, M., Mustafa, S., Plattner, G.-K., and Tréguier, A.-M.: Introduction, in: *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)], edited by van Dijk, W. R., pp. 147–286, Cambridge University Press, Cambridge, <https://doi.org/10.1017/9781009157896.003>, 2021.
- Cuesta-Valero, F. J., García-García, A., Beltrami, H., and Smerdon, J. E.: First assessment of continental energy storage in CMIP5 simulations, *Geophysical Research Letters*, 43, 5326–5335, <https://doi.org/10.1002/2016GL068496>, 2016.
- de Vrese, P. and Brovkin, V.: Timescales of the permafrost carbon cycle and legacy effects of temperature overshoot scenarios, *Nature Communications*, 12, 2688, <https://doi.org/10.1038/s41467-021-23010-5>, 2021.

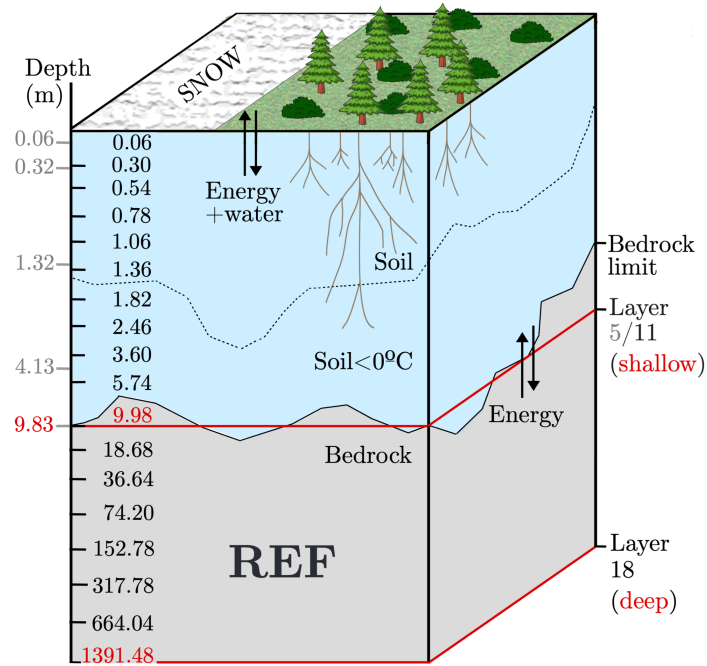
- de Vrese, P., Georgievski, G., González Rouco, J. F., Notz, D., Stacke, T., Steinert, N. J., Wilkenskjeld, S., and Brovkin, V.: Representation of soil hydrology in permafrost regions may explain large part of inter-model spread in simulated Arctic and subarctic climate, *The Cryosphere*, 17, 2095–2118, <https://doi.org/10.5194/tc-17-2095-2023>, 2023.
- Ekici, A., Beer, C., Hagemann, S., Boike, J., Langer, M., and Hauck, C.: Simulating high-latitude permafrost regions by the JSBACH terrestrial ecosystem model, *Geoscientific Model Development*, 7, 631–647, <https://doi.org/10.5194/gmd-7-631-2014>, 2014.
- 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.
- Farquharson, L. M., Romanovsky, V. E., Cable, W. L., Walker, D. A., Kokelj, S. V., and Nicolsky, D.: Climate Change Drives Widespread and Rapid Thermokarst Development in Very Cold Permafrost in the Canadian High Arctic, *Geophysical Research Letters*, 46, 6681–6689, <https://doi.org/10.1029/2019GL082187>, 2019.
- García-Pereira, F., González-Rouco, J. F., Melo-Aguilar, C., Steinert, N. J., García-Bustamante, E., de Vrese, P., Jungclaus, J., Lorenz, S., Hagemann, S., Cuesta-Valero, F. J., García-García, A., and Beltrami, H.: First comprehensive assessment of industrial-era land heat uptake from multiple sources, *Earth System Dynamics*, 15, 547–564, <https://doi.org/10.5194/esd-15-547-2024>, 2024a.
- García-Pereira, F., González-Rouco, J. F., Schmid, T., Melo-Aguilar, C., Vegas-Cañas, C., Steinert, N. J., Roldán-Gómez, P. J., Cuesta-Valero, F. J., García-García, A., Beltrami, H., and de Vrese, P.: Thermodynamic and hydrological drivers of the soil and bedrock thermal regimes in central Spain, *SOIL*, 10, 1–21, <https://doi.org/10.5194/soil-10-1-2024>, 2024b.
- Giorgetta, M. A., Roeckner, E., Mauritsen, T., Bader, J., Crueger, T., Esch, M., Rast, S., Kornblueh, L., Schmidt, H., Kinne, S., Hohenegger, C., Möbis, B., Krismer, T., Wieners, K.-H., and Stevens, B.: The atmospheric general circulation model ECHAM6 - Model description, vol. 135 of *Berichte zur Erdsystemforschung / Max-Planck-Institut für Meteorologie*, <https://doi.org/10.17617/2.1810480>, 2013.
- González-Rouco, J. F., Beltrami, H., Zorita, E., and Stevens, M. B.: Borehole climatology: a discussion based on contributions from climate modeling, *Climate of the Past*, 5, 97–127, <https://doi.org/10.5194/cp-5-97-2009>, 2009.
- González-Rouco, J. F., Steinert, N. J., García-Bustamante, E., Hagemann, S., de Vrese, P., Jungclaus, J. H., Lorenz, S. J., Melo-Aguilar, C., García-Pereira, F., and Navarro, J.: Increasing the depth of a Land Surface Model. Part I: Impacts on the soil thermal regime and energy storage, *Journal of Hydrometeorology*, 22, 3211 – 3230, <https://doi.org/10.1175/JHM-D-21-0024.1>, 2021.
- Gulev, S., Thorne, P., Ahn, J., Dentener, F., Domingues, C., Gerland, S., Gong, D., Kaufman, D., Nnamchi, H., Quaas, J., Rivera, J., Sathiyendranath, S., Smith, S., Trewin, B., von Schuckmann, K., and Vose, R.: Changing State of the Climate System, in: *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)], edited by van Dijk, W. R., pp. 287–422, Cambridge University Press, Cambridge, <https://doi.org/10.1017/9781009157896.004>, 2021.
- Hermoso de Mendoza, I., Beltrami, H., MacDougall, A. H., and Mareschal, J.-C.: Lower boundary conditions in land surface models – effects on the permafrost and the carbon pools: a case study with CLM4.5, *Geoscientific Model Development*, 13, 1663–1683, <https://doi.org/10.5194/gmd-13-1663-2020>, 2020.
- Heuvelink, G. B. M.: Uncertainty and Uncertainty Propagation in Soil Mapping and Modelling, pp. 439–461, Springer International Publishing, Cham, [https://doi.org/10.1007/978-3-319-63439-5\\_14](https://doi.org/10.1007/978-3-319-63439-5_14), 2018.
- Hjort, J., Streletskiy, D., Doré, G., Wu, Q., Bjella, K., and Luoto, M.: Impacts of permafrost degradation on infrastructure, *Nature Reviews Earth & Environment*, 3, 24–38, <https://doi.org/10.1038/s43017-021-00247-8>, 2022.

- Hugelius, G., Bockheim, J. G., Camill, P., Elberling, B., Grosse, G., Harden, J. W., Johnson, K., Jorgenson, T., Koven, C. D., Kuhry, P., Michaelson, G., Mishra, U., Palmtag, J., Ping, C.-L., O'Donnell, J., Schirrmeister, L., Schuur, E. A. G., Sheng, Y., Smith, L. C., Strauss, J., and Yu, Z.: A new data set for estimating organic carbon storage to 3 m depth in soils of the northern circumpolar permafrost region, *Earth System Science Data*, 5, 393–402, <https://doi.org/10.5194/essd-5-393-2013>, 2013.
- 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, <https://doi.org/10.5194/bg-11-6573-2014>, 2014.
- Jungclauss, J. H., Fischer, N., Haak, H., Lohmann, K., Marotzke, J., Matei, D., Mikolajewicz, U., Notz, D., and von Storch, J. S.: Characteristics of the ocean simulations in the Max Planck Institute Ocean Model (MPIOM) the ocean component of the MPI-Earth system model, *Journal of Advances in Modeling Earth Systems*, 5, 422–446, <https://doi.org/10.1002/jame.20023>, 2013.
- Koven, C. D., Riley, W. J., and Stern, A.: Analysis of Permafrost Thermal Dynamics and Response to Climate Change in the CMIP5 Earth System Models, *Journal of Climate*, 26, 1877 – 1900, <https://doi.org/10.1175/JCLI-D-12-00228.1>, 2013.
- Lacroix, F., Zaehle, S., Caldararu, S., Schaller, J., Stimmler, P., Holl, D., Kutzbach, L., and Göckede, M.: Mismatch of N release from the permafrost and vegetative uptake opens pathways of increasing nitrous oxide emissions in the high Arctic, *Global Change Biology*, 28, 5973–5990, <https://doi.org/10.1111/gcb.16345>, 2022.
- Langer, M., von Deimling, T. S., Westermann, S., Rolph, R., Rutte, R., Antonova, S., Rachold, V., Schultz, M., Oehme, A., and Grosse, G.: Thawing permafrost poses environmental threat to thousands of sites with legacy industrial contamination, *Nature Communications*, 14, 1721, <https://doi.org/10.1038/s41467-023-37276-4>, 2023.
- Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., Brovkin, V., Claussen, M., Crueger, T., Esch, M., Fast, I., Fiedler, S., Fläschner, D., Gayler, V., Giorgetta, M., Goll, D. S., Haak, H., Hagemann, S., Hedemann, C., Hohenegger, C., Ilyina, T., Jahns, T., Jiménez-de-la Cuesta, D., Jungclauss, J., Kleinen, T., Kloster, S., Kracher, D., Kinne, S., Kleberg, D., Lasslop, G., Kornblueh, L., Marotzke, J., Matei, D., Meraner, K., Mikolajewicz, U., Modali, K., Möbis, B., Müller, W. A., Nabel, J. E. M. S., Nam, C. C. W., Notz, D., Nyawira, S.-S., Paulsen, H., Peters, K., Pincus, R., Pohlmann, H., Pongratz, J., Popp, M., Raddatz, T. J., Rast, S., Redler, R., Reick, C. H., Rohrschneider, T., Schemann, V., Schmidt, H., Schnur, R., Schulzweida, U., Six, K. D., Stein, L., Stemmler, I., Stevens, B., von Storch, J.-S., Tian, F., Voigt, A., Vrese, P., Wieners, K.-H., Wilkenskjaeld, S., Winkler, A., and Roeckner, E.: Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and Its Response to Increasing CO<sub>2</sub>, *Journal of Advances in Modeling Earth Systems*, 11, 998–1038, <https://doi.org/https://doi.org/10.1029/2018MS001400>, 2019.
- 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.
- Mottaghy, D. and Rath, V.: Latent heat effects in subsurface heat transport modelling and their impact on palaeotemperature reconstructions, *Geophysical Journal International*, 164, 236–245, <https://doi.org/10.1111/j.1365-246X.2005.02843.x>, 2006.
- Mudryk, L., Santolaria-Otín, M., Krinner, G., Ménégoz, M., Derksen, C., Brutel-Vuilmet, C., Brady, M., and Essery, R.: Historical Northern Hemisphere snow cover trends and projected changes in the CMIP6 multi-model ensemble, *The Cryosphere*, 14, 2495–2514, <https://doi.org/10.5194/tc-14-2495-2020>, 2020.

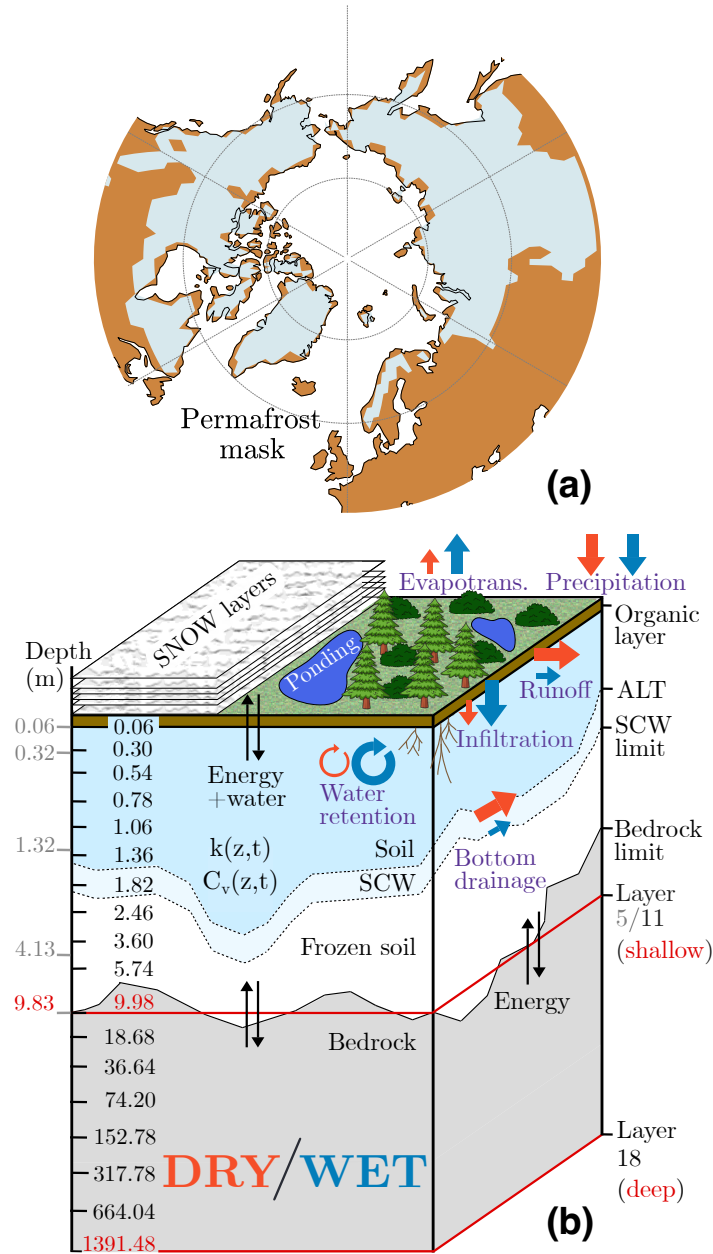
- Nitzbon, J., Schneider von Deimling, T., Aliyeva, M., Chadburn, S., Grosse, G., Laboor, S., Lee, H., Lohmann, G., Steinert, N., Stuenzi, S., Werner, M., Westermann, S., and Langer, M.: No respite from permafrost-thaw impacts in the absence of a global tipping point, *Nature Climate Change*, <https://doi.org/10.1038/s41558-024-02011-4>, 2024.
- Obu, J.: How Much of the Earth's Surface is Underlain by Permafrost?, *Journal of Geophysical Research: Earth Surface*, 126, e2021JF006123, <https://doi.org/10.1029/2021JF006123>, 2021.
- Obu, J., Westermann, S., Bartsch, A., Berdnikov, N., Christiansen, H., Avirmed, D., Delaloye, R., Elberling, B., Etzelmüller, B., Kholodov, A., Khomutov, A., Kääb, A., Leibman, M., Lewkowicz, A., Panda, S., Romanovsky, V., Way, R., Westergaard-Nielsen, A., Wu, T., and Zou, D.: Northern Hemisphere permafrost map based on TTOP modelling for 2000–2016 at 1 km<sup>2</sup> scale, *Earth-Science Reviews*, 193, 299–316, <https://doi.org/10.1016/j.earscirev.2019.04.023>, 2019.
- Obu, J., Westermann, S., Barboux, C., Bartsch, A., Delaloye, R., Grosse, G., Heim, B., Hugelius, G., Irrgang, A., Kääb, A., Kroisleitner, C., Matthes, H., Nitze, I., Pellet, C., Seifert, F., Strozzi, T., Wegmüller, U., Wiczorek, M., and Wiesmann, A.: ESA Permafrost Climate Change Initiative (Permafrost\_cci): Permafrost Ground Temperature for the Northern Hemisphere, v3.0, NERC EDS Centre for Environmental Data Analysis, <https://doi.org/10.5285/b25d4a6174de4ac78000d034f500a268>, 2021.
- Oleson, K., Lawrence, D., Bonan, G., Flanner, M., Kluzek, E., Lawrence, P., Levis, S., Swenson, S., Thornton, P., Dai, A., Decker, M., Dickinson, R., Feddema, J., Heald, C., Hoffman, F., Lamarque, J.-F., Mahowald, N., Niu, G.-Y., Qian, T., and Zeng, X.: Technical Description of version 4.0 of the Community Land Model (CLM), <https://doi.org/10.5065/D6FB50WZ>, 2010.
- Outcalt, S. I., Nelson, F. E., and Hinkel, K. M.: The zero-curtain effect: Heat and mass transfer across an isothermal region in freezing soil, *Water Resources Research*, 26, 1509–1516, <https://doi.org/10.1029/WR026i007p01509>, 1990.
- Parmesan, C., Morecroft, M., Trisurat, Y., Adrian, R., Anshari, G., Arneth, A., Gao, Q., Gonzalez, P., and et al.: Terrestrial and Freshwater Ecosystems and Their Services, p. 197–378, Cambridge University Press, <https://doi.org/10.1017/9781009325844.004>, 2023.
- Rantanen, M., Karpechko, A., Lipponen, A., Nordling, K., Hyvarinen, O., Ruosteenoja, K., Vihma, T., and Laaksonen, A.: The Arctic has warmed nearly four times faster than the globe since 1979, *Communications Earth & Environment*, 3, 168, <https://doi.org/10.1038/s43247-022-00498-3>, 2022.
- Reick, C. H., Gayler, V., Goll, D., Hagemann, S., Heidkamp, M., Nabel, J. E. M. S., Raddatz, T., Roeckner, E., Schnur, R., and Wilkenskjaeld, S.: JSBACH 3 - The land component of the MPI Earth System Model: documentation of version 3.2. Berichte zur Erdsystemforschung, Tech. Rep. 240, MPI für Meteorologie, Hamburg, <https://doi.org/10.17617/2.3279802>, 2021.
- Riahi, K., van Vuuren, D. P., Kriegler, E., Edmonds, J., O'Neill, B. C., Fujimori, S., Bauer, N., Calvin, K., Dellink, R., Fricko, O., Lutz, W., Popp, A., Cuaresma, J. C., KC, S., Leimbach, M., Jiang, L., Kram, T., Rao, S., Emmerling, J., Ebi, K., Hasegawa, T., Havlik, P., Humpenöder, F., Da Silva, L. A., Smith, S., Stehfest, E., Bosetti, V., Eom, J., Gernaat, D., Masui, T., Rogelj, J., Strefler, J., Drouet, L., Krey, V., Luderer, G., Harmsen, M., Takahashi, K., Baumstark, L., Doelman, J. C., Kainuma, M., Klimont, Z., Marangoni, G., Lotze-Campen, H., Obersteiner, M., Tabeau, A., and Tavoni, M.: The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview, *Global Environmental Change*, 42, 153–168, <https://doi.org/10.1016/j.gloenvcha.2016.05.009>, 2017.
- Schaefer, K., Elshorbany, Y., Jafarov, E., Schuster, P., Striegl, R., Wickland, K., and Sunderland, E.: Potential impacts of mercury released from thawing permafrost, *Nature communications*, 11, 4650, <https://doi.org/10.1038/s41467-020-18398-5>, 2020.
- Schuur, E. A., Abbott, B. W., Commane, R., Ernakovich, J., Euskirchen, E., Hugelius, G., Grosse, G., Jones, M., Koven, C., Leshyk, V., Lawrence, D., Loranty, M. M., Mauritz, M., Olefeldt, D., Natali, S., Rodenhizer, H., Salmon, V., Schädel, C., Strauss, J., Treat, C., and

- Turetsky, M.: Permafrost and Climate Change: Carbon Cycle Feedbacks From the Warming Arctic, *Annual Review of Environment and Resources*, 47, 343–371, <https://doi.org/10.1146/annurev-environ-012220-011847>, 2022.
- 795 Slater, A. G. and Lawrence, D. M.: Diagnosing Present and Future Permafrost from Climate Models, *Journal of Climate*, 26, 5608 – 5623, <https://doi.org/10.1175/JCLI-D-12-00341.1>, 2013.
- Steinert, N. J., González-Rouco, J. F., Melo Aguilar, C. A., García Pereira, F., García-Bustamante, E., de Vrese, P., Alexeev, V., Jungclaus, J. H., Lorenz, S. J., and Hagemann, S.: Agreement of Analytical and Simulation-Based Estimates of the Required Land Depth in Climate Models, *Geophysical Research Letters*, 48, e2021GL094 273, <https://doi.org/10.1029/2021GL094273>, 2021a.
- 800 Steinert, N. J., González-Rouco, J. F., de Vrese, P., García-Bustamante, E., Hagemann, S., Melo-Aguilar, C., Jungclaus, J. H., and Lorenz, S. J.: Increasing the depth of a Land Surface Model. Part II: Sensitivity to improved coupling between soil hydrology and thermodynamics and associated permafrost response, *Journal of Hydrometeorology*, 22, 3231 – 3254, <https://doi.org/10.1175/JHM-D-21-0023.1>, 2021b.
- Steinert, N. J., Debolskiy, M. V., Burke, E. J., García-Pereira, F., and Lee, H.: Evaluating permafrost definitions for global permafrost area estimates in CMIP6 climate models, *Environmental Research Letters*, 19, 014 033, <https://doi.org/10.1088/1748-9326/ad10d7>, 2023.
- 805 Steinert, N. J., Cuesta-Valero, F. J., García-Pereira, F., de Vrese, P., Melo Aguilar, C. A., García-Bustamante, E., Jungclaus, J., and González-Rouco, J. F.: Underestimated Land Heat Uptake Alters the Global Energy Distribution in CMIP6 Climate Models, *Geophysical Research Letters*, 51, e2023GL107 613, <https://doi.org/10.1029/2023GL107613>, 2024.
- Stevens, B., Giorgetta, M., Esch, M., Mauritsen, T., Crueger, T., Rast, S., Salzmann, M., Schmidt, H., Bader, J., Block, K., Brokopf, R., Fast, I., Kinne, S., Kornblueh, L., Lohmann, U., Pincus, R., Reichler, T., and Roeckner, E.: Atmospheric component of the MPI-M Earth
- 810 System Model: ECHAM6, *Journal of Advances in Modeling Earth Systems*, 5, 146–172, <https://doi.org/10.1002/jame.20015>, 2013.
- Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An Overview of CMIP5 and the Experiment Design, *Bulletin of the American Meteorological Society*, 93, 485 – 498, <https://doi.org/10.1175/BAMS-D-11-00094.1>, 2012.
- Turetsky, M., Abbott, B., Jones, M., Walter Anthony, K., Olefeldt, D., Schuur, E., Koven, C., McGuire, A., Grosse, G., Kuhry, P., Hugelius, G., Lawrence, D., Gibson, C., and Sannel, A.: Permafrost collapse is accelerating carbon release, *Nature*, 569, 32–34, <https://doi.org/10.1038/d41586-019-01313-4>, 2019.
- 815 Vonk, J. E., Tank, S. E., Bowden, W. B., Laurion, I., Vincent, W. F., Alekseychik, P., Amyot, M., Billet, M. F., Canário, J., Cory, R. M., Deshpande, B. N., Helbig, M., Jammet, M., Karlsson, J., Larouche, J., MacMillan, G., Rautio, M., Walter Anthony, K. M., and Wickland, K. P.: Reviews and syntheses: Effects of permafrost thaw on Arctic aquatic ecosystems, *Biogeosciences*, 12, 7129–7167, <https://doi.org/10.5194/bg-12-7129-2015>, 2015.
- 820 Walvoord, M. A. and Kurylyk, B. L.: Hydrologic Impacts of Thawing Permafrost—A Review, *Vadose Zone Journal*, 15, vzj2016.01.0010, <https://doi.org/10.2136/vzj2016.01.0010>, 2016.
- Westermann, S., Ingeman-Nielsen, T., Scheer, J., Aalstad, K., Aga, J., Chaudhary, N., Etzelmüller, B., Filhol, S., Kääb, A., Renette, C., Schmidt, L. S., Schuler, T. V., Zweigel, R. B., Martin, L., Morard, S., Ben-Asher, M., Angelopoulos, M., Boike, J., Groenke, B., Miesner, F., Nitzbon, J., Overduin, P., Stuenzi, S. M., and Langer, M.: The CryoGrid community model (version 1.0) – a multi-physics toolbox for
- 825 climate-driven simulations in the terrestrial cryosphere, *Geoscientific Model Development*, 16, 2607–2647, <https://doi.org/10.5194/gmd-16-2607-2023>, 2023.
- Yokohata, T., Saito, K., Takata, K., Nitta, T., Satoh, Y., Hajima, T., Sueyoshi, T., and Iwahana, G.: Model improvement and future projection of permafrost processes in a global land surface model, *Progress in Earth and Planetary Science*, 7, 69, <https://doi.org/10.1186/s40645-020-00380-w>, 2020.

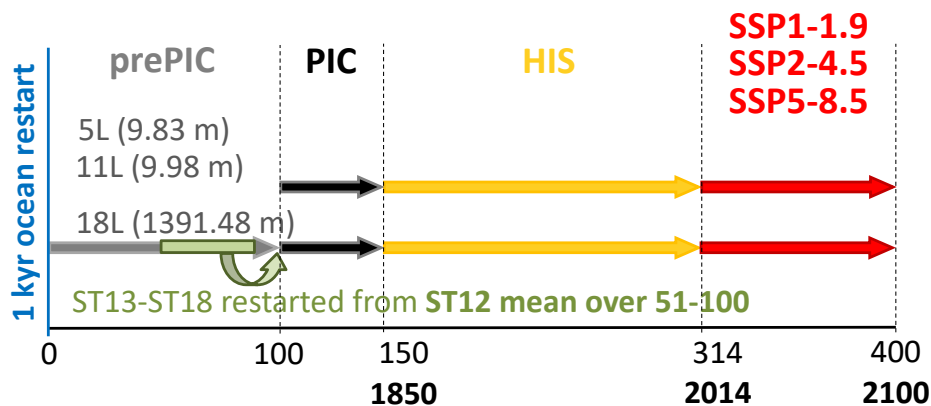
830 Zhu, X., Lee, S.-Y., Wen, X., Wei, Z., Ji, Z., Zheng, Z., and Dong, W.: Historical evolution and future trend of Northern Hemisphere snow cover in CMIP5 and CMIP6 models, *Environmental Research Letters*, 16, 065 013, <https://doi.org/10.1088/1748-9326/ac0662>, 2021.



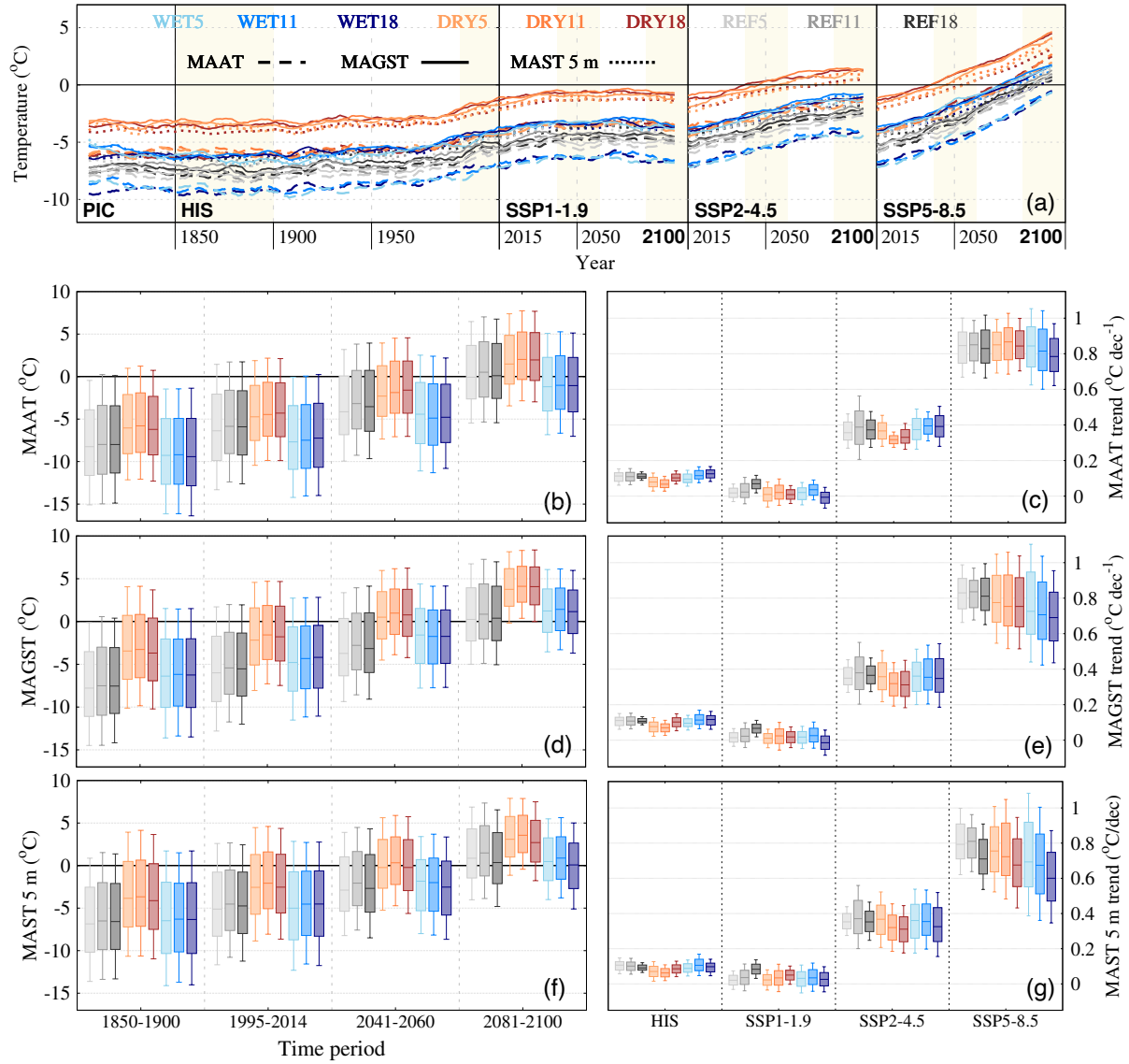
**Figure 1.** Conceptual sketch illustrating the main hydrological and thermodynamic features of the JSBACH standard version (REF). The mid-layer depth values for the default 5-layer (5L) vertical discretization, with a LSM depth at 9.83 m, are given on the left part of the y-axis. The corresponding values for the two enhanced discretizations of 11 (11L) and 18 (18L) layers introduced by this work are also shown, with LSM depths at 9.98 and 1391.48 m, respectively. Note that bedrock is drawn deeper than typical values in permafrost regions to improve visual clarity.



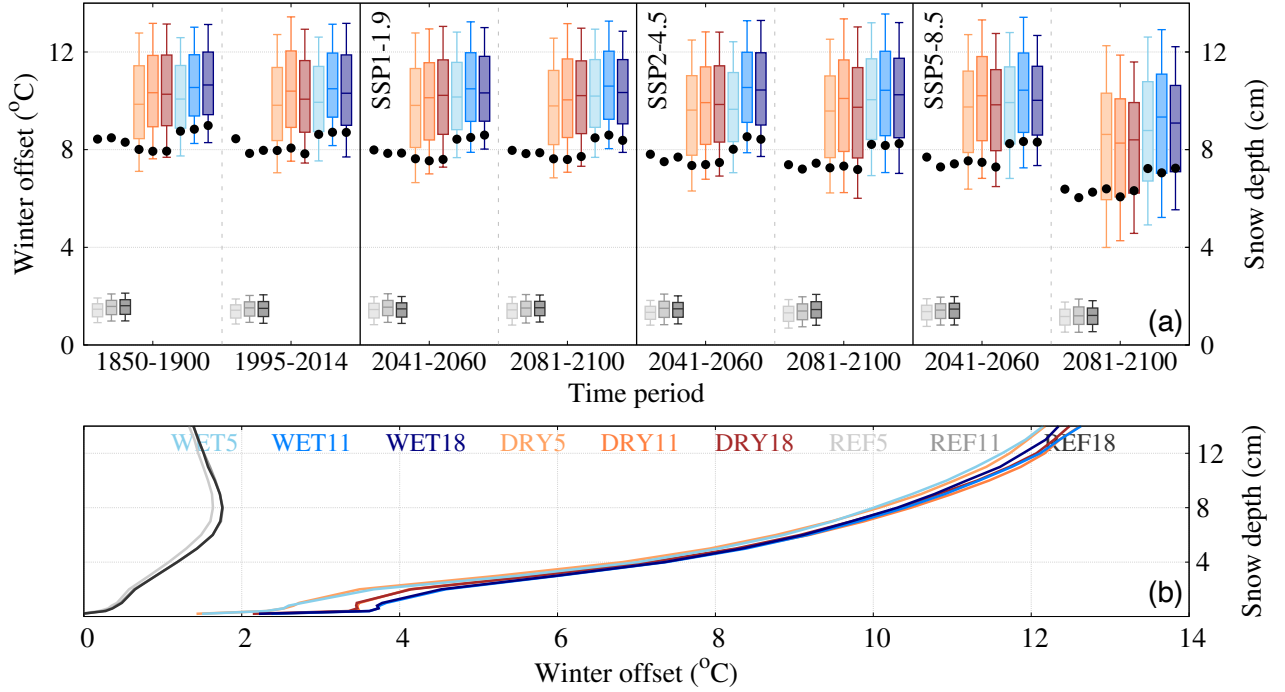
**Figure 2.** HTCp changes applied to JSBACH standard version. (a) Northern Hemisphere (45–90°N) land permafrost affected areas (shaded light blue) in the late historical period (Hugelius et al., 2013, 2014), which are taken as the mask where the JSBACH-HTCp physics are implemented. Outside the area portrayed by the permafrost mask in (a), the REF (Fig. 1) configuration of JSBACH is used. (b) Conceptual sketch showing the features and the intensity (arrow size) of different thermodynamic and hydrological processes for JSBACH-HTCp DRY (red) and WET (blue) configurations. As in Fig. 1, bedrock is drawn deeper than typical values in permafrost regions to improve visual clarity.



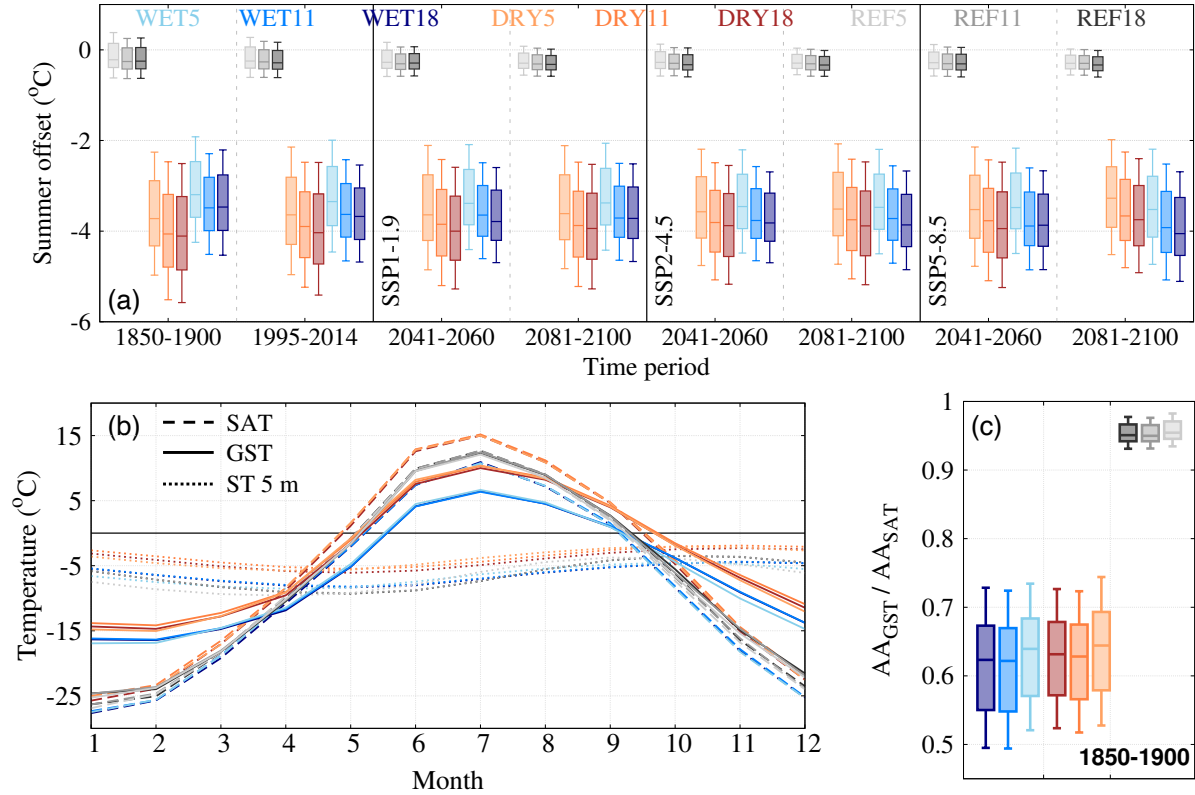
**Figure 3.** Experimental setup of the MPIESM-PePE. Scheme of preindustrial control (PIC, black), historical (HIS, yellow), and 21st century climate change Shared Socioeconomic Pathway forcing scenario (SSP1-1.9, SSP2-4.5, and SSP5-8.5, red) simulations for each of the REF, WET, and DRY configurations of the MPIESM-PePE. This sequence was also run with different vertical discretizations of 5, 11, and 18 layers (5L, 11L, and 18L, respectively); see Table 1 and Fig. 1. PIC is conducted with conditions of 1850 for 50 years. The historical run is started at year 50 of PIC for the three vertical discretizations. To speed up subsurface temperature equilibrium in the vertical column for the 18L simulations, a prior preindustrial control (prePIC) phase of 100 years is run. The subsurface temperature mean for the last 50 years at the 12th layer for the prePIC is used to restart temperatures at layers 13 to 18 in the 18L configuration (see Table 1 and text for the details).



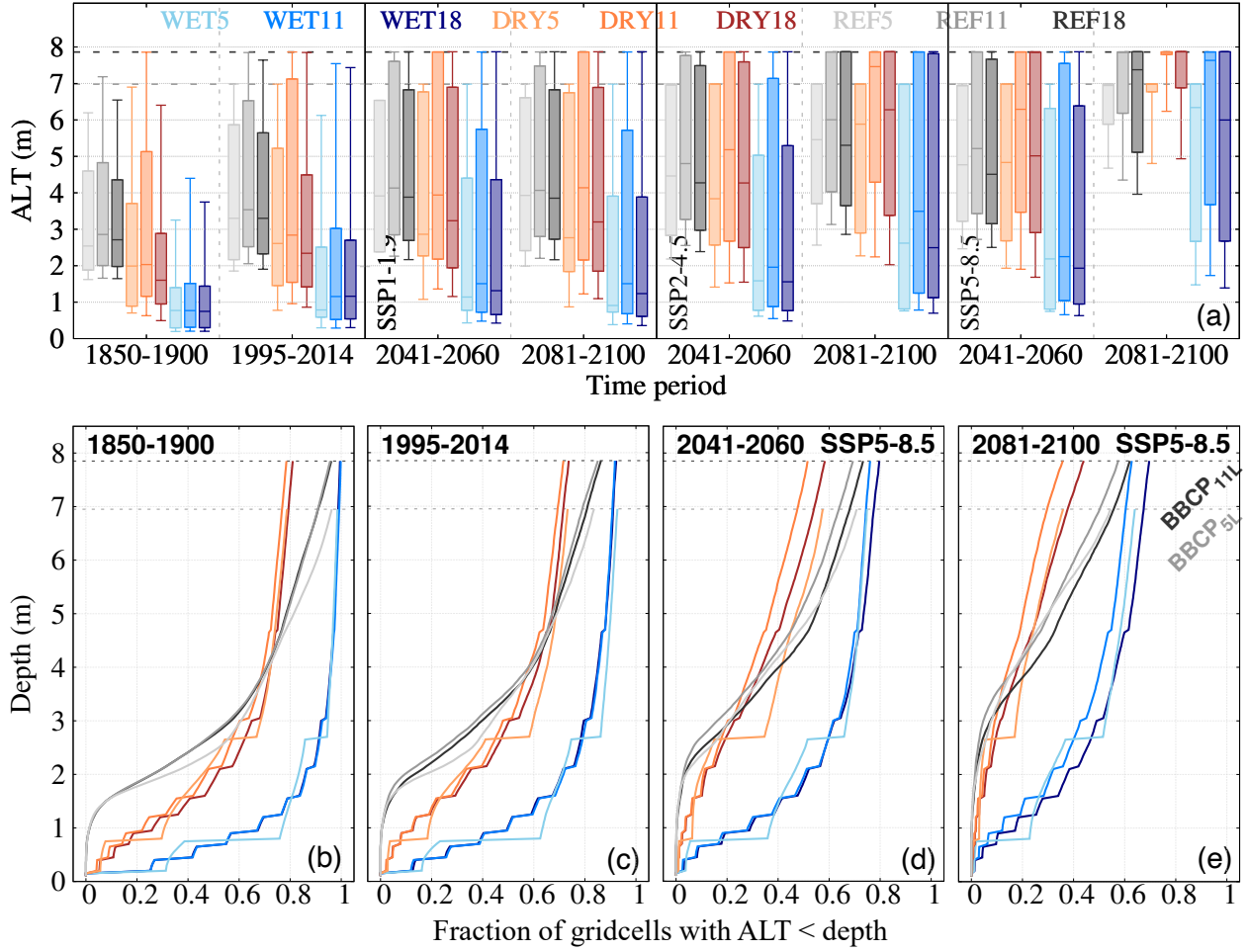
**Figure 4.** Temperature variability in the nine simulations of the MPIESM-PePE (see legend for colors) within the permafrost domain (Fig. 2a). (a) Time evolution (11-year moving average filtered outputs) of the median MAAT (dashed), MAGST (solid), and MAST 5 m (dotted line, °C) in the PIC, HIS, and different SSP scenarios (see Section 2.2). **Spatial distribution of MAAT (b), MAGST (d), and MAST 5 m (f) given by** permafrost domain boxplots of time averages for **the** selected periods highlighted in (a) with light yellow bands: early-historical (1850–1900), late historical (1995–2014), mid-21st century (2041–2060), and late 21st century (2081–2100) in the SSP5-8.5 scenario. The inner tick in the boxplot marks the median value within the permafrost mask area (Fig. 2a), the bottom and top lines the quartiles 1 and 3, and the bottom and top whiskers the 10th and 90th percentiles, respectively. MAAT (c), MAGST (e), and MAST 5 m (g) boxplots indicating the range of trends (°C dec<sup>-1</sup>) within the permafrost domain, for the historical (HIS, 1850–2014) and SSP1-1.9, SSP2-4.5, and SSP5-8.5 **scenarios** **scenario** simulations (2015–2100) are also portrayed.



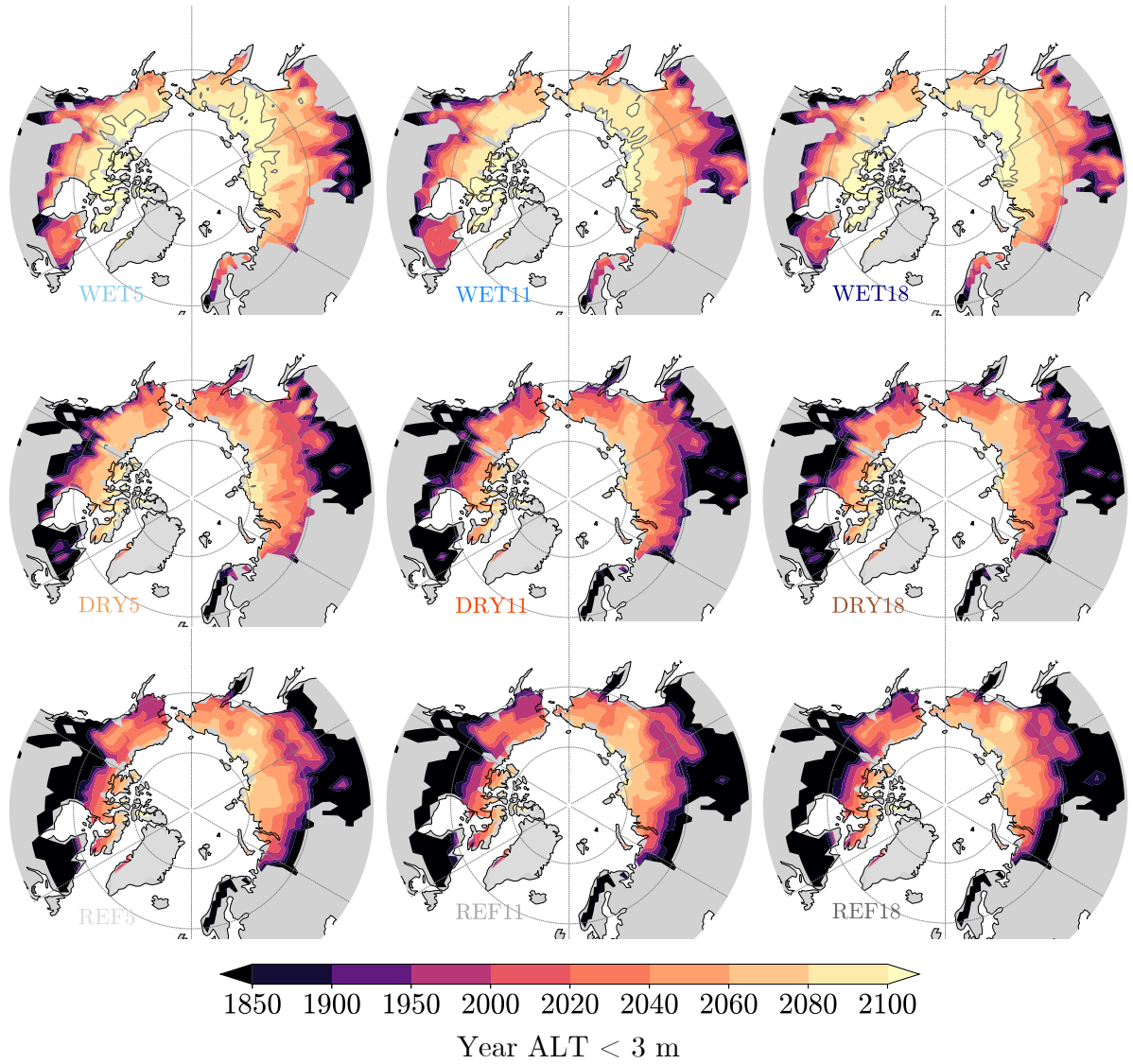
**Figure 5.** Winter offset, defined as GST-SAT in DJF (see Section 2.3.1), and snow depth. (a) Winter offset (°C) permafrost domain boxplots for the nine simulations of the MPIESM-PePE (see legend for colors) in the different periods shown in Fig. 4b,c,e. The inner tick in the boxplot marks the median value within the permafrost domain (Fig. 2a), the bottom and top lines the quartiles 1 and 3, and the bottom and top whiskers the 10th and 90th percentiles, respectively. Snow depth median values (cm) for the same periods and simulations is shown with black dots. (b) Snow depth as a function of the winter offset within the permafrost domain for every simulation of the MPIESM-PePE.



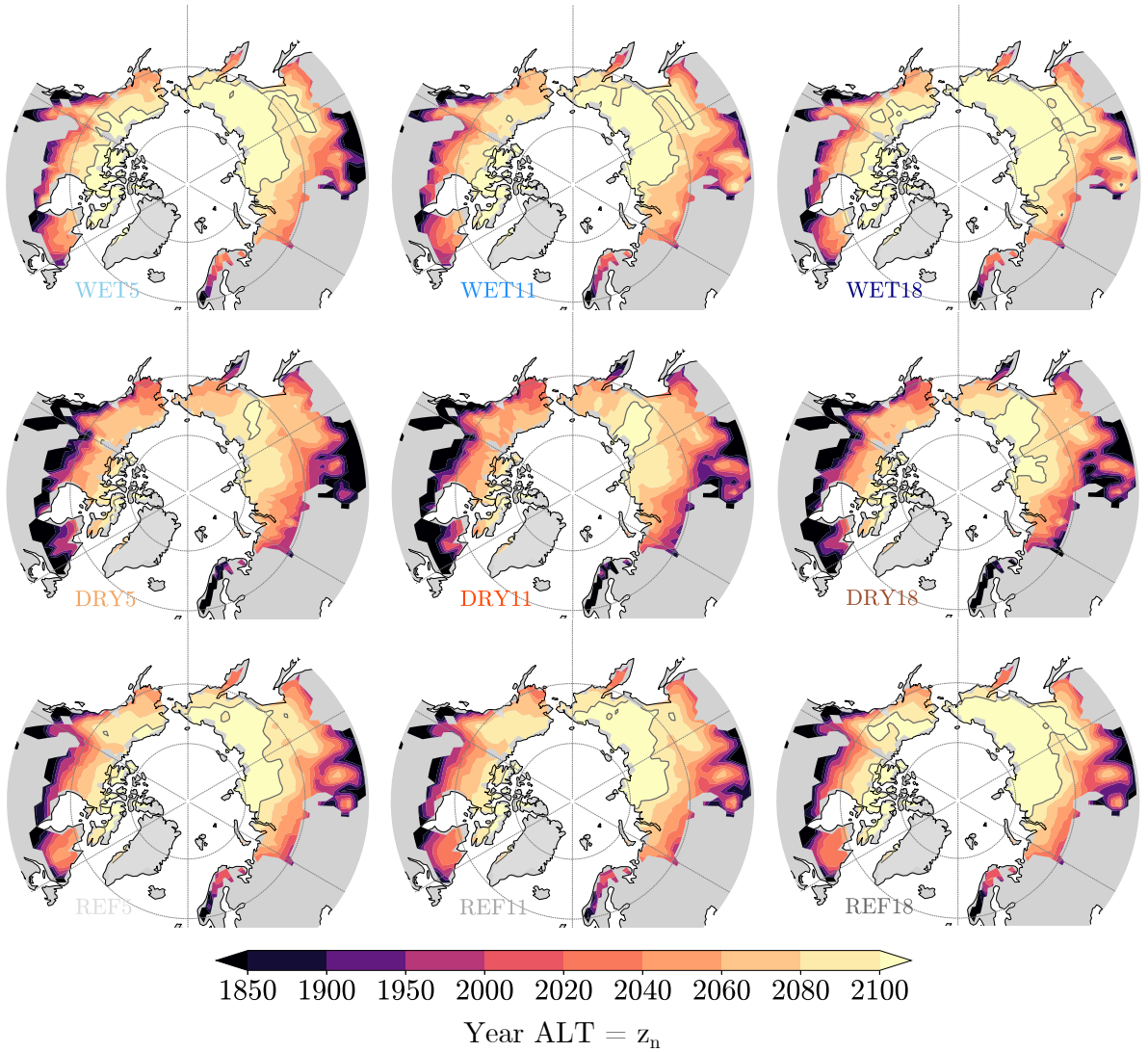
**Figure 6.** Summer offset and annual temperature cycle amplitude. (a) Summer offset (°C) permafrost domain boxplots (see Section 2.3.1) for the nine simulations of the MPIESM-PePE (see legend for colors) in the different periods shown in Fig. 4b,d,f and Fig. 5. The inner tick in the boxplot marks the median value within the permafrost mask area (Fig. 2a), the bottom and top lines the quartiles 1 and 3, and the bottom and top whiskers the 10th and 90th percentiles, respectively. (b) Amplitude (°C) of the permafrost domain median GST (solid), SAT (dashed), and ST 5 m (dotted line) annual cycles during 1850–1900 for the different MPIESM-PePE simulations in (a). (c) Ratio of GST vs. SAT annual cycle amplitude (AA) during 1850–1900 given by permafrost domain boxplots as in (a).



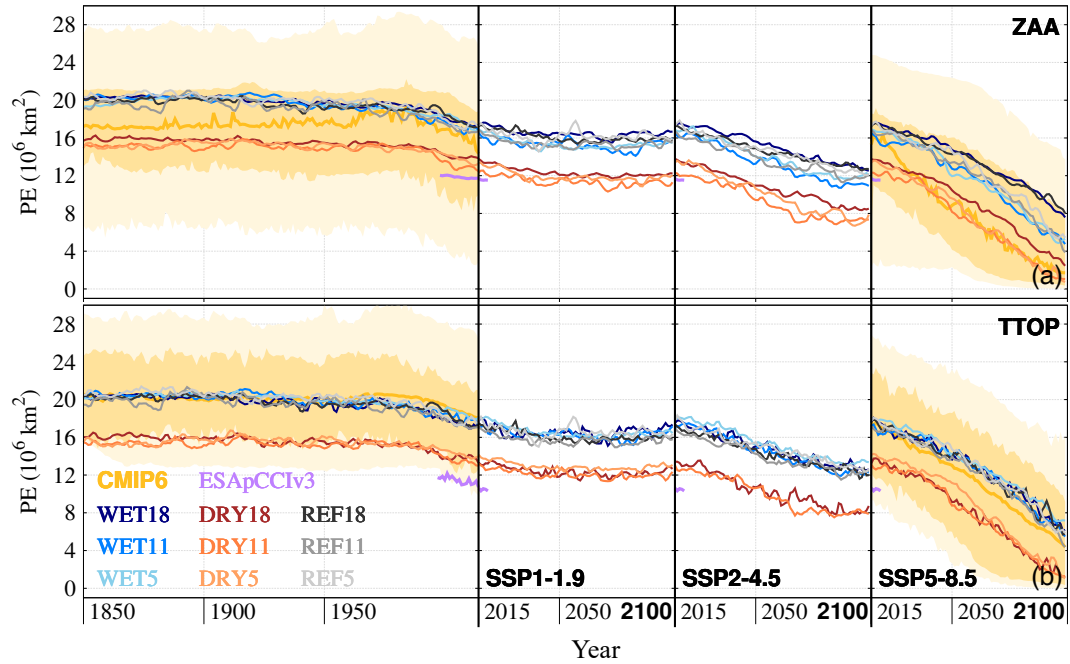
**Figure 7.** (a) Active layer thickness (ALT, m) for the ensemble of MPIESM-PePE simulations in the different periods shown in Fig. 4b,d,f and 5a. Fraction of gridcells within the MPIESM-PePE permafrost mask with a summer ALT greater than the indicated depth in the y-axis for 1850–1900 (b), 1995–2014 (c), 2041–2060 (d), and 2081–2100 (e). For every plot, dark (light) gray dashed line indicates the mid-layer depth of the bottom layer for the 11L (5L) configuration. 18L value is out of the bounds of the panels and, therefore, not included. The mid-layer depth for the bottom layer of the 18L configuration lies outside the range shown in panels a–e and is therefore not included.



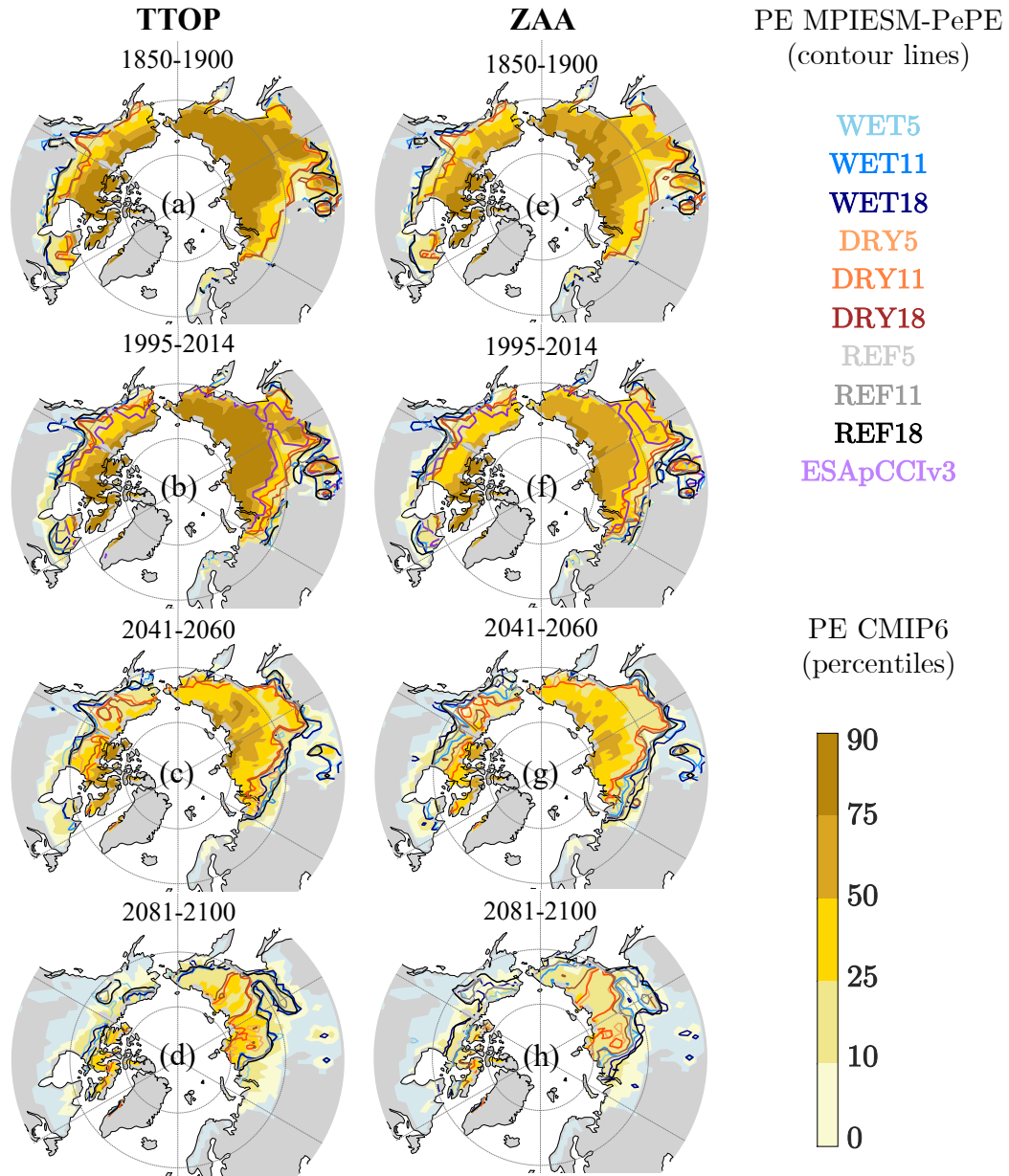
**Figure 8.** Near-surface permafrost presence with time for the combination of the three hydrological configurations (WET, DRY, REF) and vertical discretizations (5L, 11L, 18L). **Spatial**This is represented by the spatial distribution of the last simulation year **at**in which ALT is smaller than 3 m for **every simulation of the each** MPIESM-PePE **simulation**. The time evolution for **the** historical and SSP5-8.5 experiments is shown in **every each** case. Contour lines delimit areas where ALT does not reach 3 m within the temporal extent of the simulation. The colored distribution stands over the permafrost domain displayed in Fig. 2a.



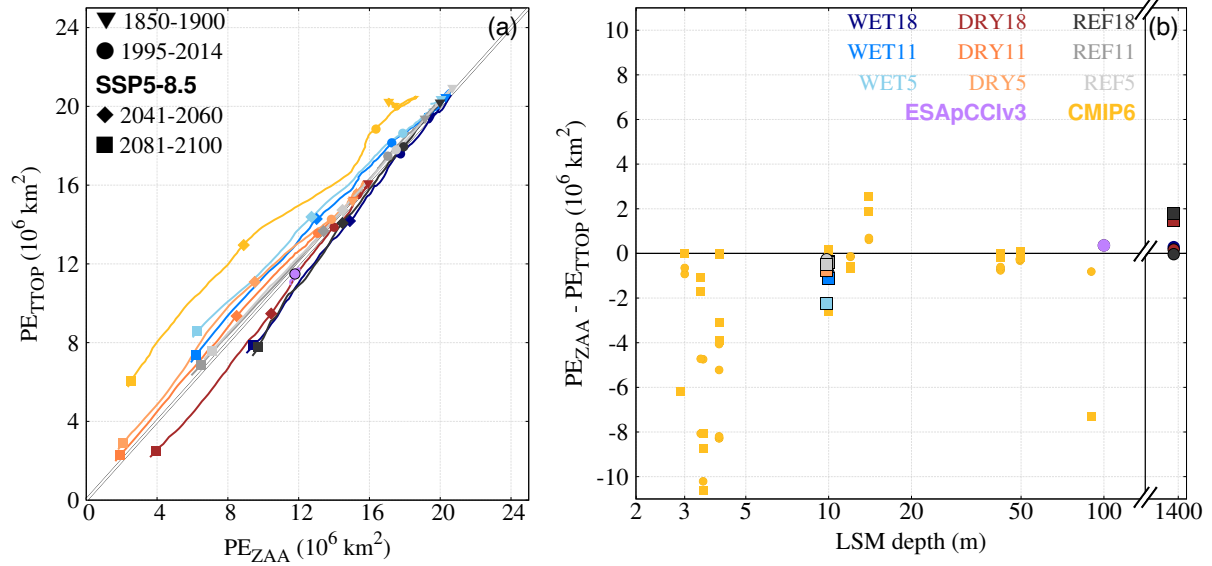
**Figure 9.** Deep permafrost presence with time for the combination of the three hydrological configurations (WET, DRY, REF) and vertical discretizations (5L, 11L, 18L). **Spatial** This is represented by the spatial distribution of the last simulation year (HIS + SSP5-8.5) **at** in which ALT is equal to the mid-layer depth of the bottom layer ( $ALT = z_n$ ) for every simulation of the each MPIESM-PePE simulation. For the 18L simulations, 11L JSBACH  $z_n$  was taken in order to allow for comparison. Contour lines delimit areas where ALT do not reach  $z_n$  within the temporal extent of the simulation. The colored distribution stands over the permafrost domain displayed in Fig. 2a.



**Figure 10.** Permafrost extent evolution according to ZAA and TTOP (see Section 2.3.3). (a) PE ( $10^6 \text{ km}^2$ ) using the ZAA definition for the different MPIESM-PePE simulations (solid lines, see legend for colors), the CMIP6 ensemble (gold solid line for the median, and dark (light) golden shading for the interquartile (interdecile) range), and the observational product ESAPCCiv3 (purple solid line). (b) Same as (a), but for TTOP PE metric. CMIP6 PE data is only available for the historical period and SSP5-8.5 scenario (Steinert et al., 2023).



**Figure 11.** Permafrost extent (PE) for different periods in the MPIESM-PePE and CMIP6 ensembles in the historical and SSP5-8.5 scenario, and observed historical extension. PE CMIP6 10th, 25th, 50th, 75th, and 90th percentiles spatial distribution (shading) for TTOP (a–d) and ZAA (e–h) for 1850–1900, 1995–2014, 2041–2060, and 2081–2100. The extent and distribution of PE for the different MPIESM-PePE experiments in the same periods are depicted with contour lines of different colors (see legend), whilst the observational ESAPCCIV3 historical permafrost distribution in 1997–2019 is shown with a thicker purple contour line.



**Figure 12.** (a) Comparison of ZAA (x-axis) vs. TTOP (y-axis) PE estimates ( $10^6 \text{ km}^2$ ) for the CMIP6 median and the different members of MPIESM-PePE ensemble (see Fig. 10). ~~In this study, ZAA and TTOP definitions are used as indicators of deep and near-surface PE evolution, respectively (see Section 2.3.3).~~ Solid lines depict the year-by-year evolution of ZAA vs. TTOP values. The symbols (see legend) illustrate the mean ZAA vs. TTOP values for different periods. ZAA and TTOP PE estimates are filtered with a 15-year moving average. (b) ZAA - TTOP PE differences ( $10^6 \text{ km}^2$ ) in 1995–2014 (circles), and 2081–2100 (squares) for the 32 members of CMIP6 ensemble stemming from Steinert et al. (2023), the nine MPIESM-PePE simulations presented in this work, and the ESAPCCiv3 observational PE data (only in 1997–2019). Data sources are sorted by their LSM depth (m, logarithmic x-axis). Note the break in x-axis between 100 and 1400 m.