Improving Permafrost Dynamics in Land Surface Models: Insights from Dual Sensitivity Experiments

Dissertation

Submitted to the
Faculty of Mathematics and Natural Sciences
Institute of Physics and Astronomy
of the University of Potsdam
and
the Alfred Wegener Institute
Helmholtz Center for Polar and Marine Research
Potsdam Research Unit
to obtain the academic degree of

Doctor Rerum Naturalium

(Dr. rer. nat.)

in the scientific discipline of "Climatology Physics"

by

Adrien DAMSEAUX

Date of submission: October 31th, 2023, at Potsdam
Date of disputation: April 15th, 2024, at Potsdam

Main supervisor: Prof. Markus Rex
Supervisors: Dr. Annette Rinke, Dr. Stefan Hagemann, Dr. Heidrun Matthes
Reviewers: Prof. Markus Rex, Prof. Guido Grosse, Prof. Priscilla Mooney

Published online on the
Publication Server of the University of Potsdam:
https://doi.org/10.25932/publishup-63945
https://nbn-resolving.org/urn:nbn:de:kobv:517-opus4-639450

ABSTRACT

Arctic permafrost regions are critical for anthropogenic climate change. Permafrost thaw and its subsequent greenhouse gas emissions profoundly influence various societal, ecological, hydrological, and biogeochemical processes in the Arctic and stand as one of the most significant and uncertain positive feedback loops in the context of climate change. To address these critical questions, climate scientists have developed Land Surface Models (LSMs) that encompass a multitude of physical soil processes. This thesis is committed to advancing our understanding and refining precise representations of permafrost dynamics within LSMs, with a specific focus on the accurate modeling of heat fluxes, an essential component for simulating permafrost physics.

The first research question overviews fundamental model prerequisites for the representation of permafrost soils within land surface modeling. It includes a first-of-its-kind comparison between LSMs that participated in CMIP6 to reveal their differences and shortcomings in key permafrost physics parameters. Overall, each of these LSMs represents a unique approach to simulating soil processes and their interactions with the climate system. Choosing the most appropriate model for a particular application depends on factors such as the spatial and temporal scale of the simulation, the specific research question, and available computational resources.

The second research question evaluates the performance of the state-of-the-art Community Land Model (CLM5) in simulating Arctic permafrost regions. Our approach overcomes traditional evaluation limitations by individually addressing depth, seasonality, and regional variations, providing a comprehensive assessment of permafrost and soil temperature dynamics. I compare CLM5's results with three extensive datasets: (1) soil temperatures from 295 borehole stations, (2) active layer thickness (ALT) from the Circumpolar Active Layer Monitoring Network (CALM), and (3) soil temperatures, ALT, and permafrost extent from the ESA Climate Change Initiative (ESA-CCI). The results show that CLM5 aligns well with ESA-CCI and CALM for permafrost extent and ALT but reveals a significant regional cold temperature bias, notably over Siberia. These results echo a persistent challenge identified in numerous studies: the existence of a systematic "cold bias" in soil temperature over permafrost regions. To address this challenge, the following research questions propose dual sensitivity experiments.

The third research question represents the first study to apply a Plant Functional

Type (PFT)-based approach to derive soil texture and soil organic matter (SOM), departing from the conventional use of coarse-resolution global data in LSMs. This novel method results in a more uniform distribution of soil organic matter density (OMD) across the domain, characterized by reduced OMD values in most regions. However, changes in soil texture exhibit a more intricate spatial pattern. Comparing the results to observations reveals a significant reduction in the cold bias observed in the control run. This method shows noticeable improvements in permafrost extent, but at the cost of an overestimation in ALT. These findings emphasize the model's high sensitivity to variations in soil texture and SOM content, highlighting the crucial role of soil composition in governing heat transfer processes and shaping the seasonal variation of soil temperatures in permafrost regions.

Expanding upon a site experiment conducted in Trail Valley Creek by Dutch et al. (2022), the fourth research question extends the application of the snow scheme proposed by Sturm et al. (1997) to cover the entire Arctic domain. By employing a snow thermal conductivity scheme better suited to the snow density profile observed over permafrost regions, this thesis seeks to assess its influence on simulated soil temperatures. Comparing this method to observational datasets reveals a significant reduction in the cold bias that was present in the control run. In most regions, the Sturm run exhibits a substantial decrease in the cold bias. However, there is a distinctive overshoot with a warm bias observed in mountainous areas. The Sturm experiment effectively addressed the overestimation of permafrost extent in the control run, albeit resulting in a substantial reduction in permafrost extent over mountainous areas. ALT results remain relatively consistent compared to the control run. These outcomes align with our initial hypothesis, which anticipated that the reduced snow insulation in the Sturm run would lead to higher winter soil temperatures and a more accurate representation of permafrost physics.

In summary, this thesis demonstrates significant advancements in understanding permafrost dynamics and its integration into LSMs. It has meticulously unraveled the intricacies involved in the interplay between heat transfer, soil properties, and snow dynamics in permafrost regions. These insights offer novel perspectives on model representation and performance.

ACKNOWLEDGEMENT

With immense gratitude, I wish to express my sincere thanks to my mentor, Dr. Heidrun Matthes. Her expertise, motivation, and continuous encouragement were instrumental in the successful completion of this research.

I am also grateful to my supervisors and my thesis advisory committee, Prof. Markus Rex, Dr. Stefan Hagemann, and Dr. Annette Rinke, as well as, Prof. Nick Rutter and Dr. Victoria Dutch for their invaluable guidance and support throughout my research.

I extend my thanks to the University of Potsdam and the Alfred-Wegener-Institut for providing the necessary facilities and resources. Special thanks also go to the staff at the Deutsches Klimarechenzentrum, particularly Dr. Irina Fast, for their assistance in providing computational resources essential for my research.

I acknowledge the support of my colleagues and the entire staff at the Alfred-Wegener-Institut, whose contributions were key during my research.

I wish to extend my profound gratitude to my parents and friends for their unwavering moral support and encouragement whenever needed.

Last but not least, I would like to deeply thank Evie Morin for her meticulous review and editing of my entire thesis, along with her constant encouragement, moral support, patience, and understanding.

TABLE OF CONTENTS

Appendices

List of abbreviations

AD Average Deviation

ALT Active Layer Thickness

BC Brooks and Corey (1966)

BDS Broadleaf Deciduous Shrubs

CALM Circumpolar Active Layer Monitoring Network

CCI Climate Change Initiative

CMIP Coupled Model Intercomparison Project

CLM Community Land Model

CTSM Community Terrestrial Systems Model

DJF December, January, February

ECV Essential Climate Variable

ESA European Space Agency

ESM Earth System Model

ESM-SnowMIP Earth System Model–Snow Model Intercomparison Project

GSOCmap Global Soil Organic Carbon Map

IPCC Intergovernmental Panel on Climate Change

JJA June, July, August

JULES Joint UK Land Environment Simulator

LAI Leaf Area Index

LS3MIP Land Surface, Snow, and Soil Moisture Model Intercomparison Project

LSM Land Surface Model

MAD Mean Absolute Deviation

MAGT Mean Annual Ground Temperature

MAUP Modifiable Areal Unit Problem

MaxAGT Maximum Annual Ground Temperature

MAM March, April, May

MinAGT Minimum Annual Ground Temperature

NCSCD Northern Circumpolar Soil Carbon Database

OMC Organic Matter Content

OMD Organic Matter Density

PCN Permafrost Carbon Network

PFR Permafrost Fraction

RCM Regional Climate Model

RCP Representative Concentration Pathway

RMSE Root Mean Square Error

SES Social-Ecological Systems

SCS Sand, Clay, Silt

SON September, October, November

SWE Snow Water Equivalent

VG van Genuchten (1980)

List of variables

F	Amount of heat conducted across a unit cross-sectional area in unit time
λ	Soil thermal conductivity
∇T	Spatial gradient of temperature
z	Vertical height
F_z	Amount of heat conducted across a unit cross-sectional area in unit time in the vertical direction
H	Heat content per unit volume
t	Time
r_H	Heat sink term
c	Heat storage capacity
$\frac{dT}{dt}$	Rate of change of temperature with respect to time
$\frac{dT}{dz}$	Rate of change of temperature with respect to depth
K_t	Soil thermal diffusivity
θ	Volumetric soil water content
θ_s	Saturated volumetric soil water content
q	Soil water flux
e	Soil moisture sink term
k	Hydraulic conductivity
ψ	Hydraulic potential
L_f	Latent heat of fusion of water
ρ_{ice}	Density of ice

θ_{ice} Amount of ice that freezes or melts within time

ρ_b Bulk density

c_i Heat capacity of the material in question

$c_{s,i}$ Heat capacity of soil solids

ϕ_i Porosity

Δz_i Depth of the soil layer i

$w_{liq,i}$ Liquid water content

$w_{ice,i}$ Ice content

C_{liq} Specific heat capacity of liquid water

C_{ice} Specific heat capacity of ice

n Number of data points

y_i Observed values

$\hat{y}_i$ Predicted (estimated) values

K_{eff} Effective thermal conductivity

λ_{air} Thermal conductivity of air

λ_{ice} Thermal conductivity of ice

ρ_{sno} Snow density

P_a Air pressure

P_0 Reference pressure

c Volumetric soil heat capacity

λ_s Soil thermal conductivity

K_h Hydraulic conductivity

K_{hs} Hydraulic conductivity for saturated soil

K_e Kersten number

S_r Degree of saturation

$\lambda_{bedrock}$ Thermal conductivity of deep ground layers

λ_{sat} Saturated thermal conductivity

λ_{soil} Thermal conductivity of soil solids

$\%sand$ Percentage of sand

$\%clay$ Percentage of clay

λ_{om} Thermal conductivity of SOM solid

f_{om} Soil organic matter fraction

omd Organic matter density

omd_{max} Maximum soil carbon density

c_{soil} Heat capacity of soil solids

c_{st} Heat capacity of soil texture

c_{liq} Heat capacity of liquid water

c_{ice} Heat capacity of ice

k_{sat} Saturated hydraulic conductivity

k_{st} Soil texture hydraulic conductivity

k_{om} Hydraulic conductivity of SOM

z_i Soil depth at layer i

$zsapric$ Depth at which SOM takes on the characteristics of sapric peat

ρ_{sno} Snow density

ρ_T Thermal conductivity of snow

T Atmospheric near-surface temperature

T_{frz} Freezing temperature of water (273.158K)

ρ_w Wind density factor

U 10 m wind speed

ρ_{fs} Snow density including temperature and wind effects

CHAPTER 1

Motivations

1.1 Introduction

Permafrost is a widespread feature of the Earth's northern high-latitude regions, underlying about 9 to 14% of the exposed Earth land surface (Gruber, 2012). It is commonly defined as ground that remains at or below 0°C for at least two consecutive years and contains soil, sediment, and/or rock with varying amounts of ice (Osterkamp and Romanovsky, 1999). Permafrost is a thermal phenomenon that occurs in regions where air temperatures are consistently low enough to keep the ground frozen below a near-surface layer that thaws seasonally, known as the active layer.

In recent years, there has been increasing attention on estimates of carbon stored in soils, particularly in permafrost regions. Studies estimate that permafrost contains between 677 and 949 Pg of soil organic carbon (SOC) in the upper few meters of soil, which is greater than the total amount found in the atmosphere (Palmtag et al., 2022). The accumulation of such high carbon content in permafrost soils can be attributed to plant remains and other organic material that have been buried and frozen since the last ice age, and have been inactive for centuries or even millennia (Schaefer et al., 2014).

The release of carbon from this substantial pool, triggered by warming-induced thawing and the acceleration of microbial decomposition, represents a pivotal positive feedback loop contributing to climate change (Schuur et al., 2015). Arctic Amplification, which represents warming rates in Arctic regions that are nearly four times higher than the global average (Rantanen et al., 2022), exacerbates This process. Climate models indicate that this trend will persist, projecting warming rates between 2.4 to 10.0°C by 2100, relative to the 1995-2014 period and depending on the specific Shared Socio-Economic Pathways considered (IPCC, 2021). Despite this, the Arctic remains one of Earth's least investigated regions (Miner et al., 2022), making predictions challenging.

These significant temperature increases would lead to substantial portions of permafrost areas surpassing the thawing threshold. Model projections indicate a significant 20% reduction in the near-surface permafrost area of the Northern Hemisphere, declining from approximately 15 million square kilometers at the present to 12 million square kilometers by 2040, with this trend remaining relatively consistent across dif-

ferent Representative Concentration Pathway (RCP) scenarios (Overland et al., 2019). Specifically, under RCP 4.5, there is a substantial 50% reduction in permafrost area relative to the present, expected by 2080. For this reason, permafrost areas over high northern latitudes are critical for anthropogenic climate change and constitute a central focus within climate science.

1.1.1 History and classification of permafrost

The term ”permafrost” was first coined by Siemon W. Muller in 1945 to describe the permanently frozen ground in Alaska (Muller, 1945). However, observations of frozen ground date back to the late 19th century, when the Russian scientist P. E. Kropotkin first documented the presence of frozen soils in Siberia (Romanovsky et al., 2010). Since then, permafrost has been extensively studied by scientists from a range of disciplines including geology, geography, hydrology, and climatology.

Permafrost covers extensive areas in the Arctic and subarctic regions, where the mean annual air temperature is below freezing. The depth of permafrost varies from a few meters to several hundred meters, depending on the local climate and geological conditions (Bowden, 2010). The permafrost spatial extent, the actual surface area underlain by permafrost, in the Arctic is estimated to be 13-18$\times 10^6 km^2$ or 9–14 % of the exposed land surface (Gruber, 2012).

Permafrost is commonly categorized into types based on factors such as its extent, distribution, and soil ice content. However, there is currently no universally agreed-upon definition of these classifications within the scientific community. Nonetheless, there is broad consensus on the differentiation between two main categories: ”continuous” and ”discontinuous” permafrost zones. The term ”continuous” refers to regions where permafrost covers more than 90% of the land surface, while ”discontinuous” zones indicate areas where permafrost covers less than 90% of the land surface. Other commonly used classifications include ”sporadic” permafrost, regions where permafrost covers between 10% and 50% of the land surface, and ”isolated” permafrost, areas where permafrost covers less than 10% of the land surface (Gruber, 2012).

1.1.2 Active Layer Thickness

Active Layer Thickness (ALT) refers to the depth of the soil layer that undergoes seasonal freeze-thaw cycles above the permafrost. Fundamentally, ALT is determined by the intersection between the Maximum Annual Ground Temperature (MaxAGT) and the 0°C isotherm, as illustrated in Figure A.2 in Appendix A. It is an important indicator to predict permafrost thaw as it responds rapidly to changes in temperature and precipitation. Changes in ALT have been observed in many Arctic permafrost regions due to climate warming, with increasing ALT linked to permafrost thaw (Lawrence and

Slater, 2005). In addition, an increase in ALT can lead to changes in ecological (Loranty et al., 2018; Yang et al., 2010) and hydrological processes in permafrost regions, with a thicker active layer resulting in increased water infiltration and runoff (Scherler et al., 2010).

1.2 Importance of permafrost for northern social-ecological systems (SES)

Permafrost is a complex system that exerts a profound influence on various ecological, hydrological, biogeochemical processes, ultimately impacting societies in Arctic and Subarctic regions. This interconnected relationship is commonly referred to as social-ecological systems (SES) (Schuur and Mack, 2018).

According to Ramage et al. (2021), approximately 3.3 million people residing in Arctic regions will experience the effects of thawing permafrost, leading to a wide array of SES changes affecting their surrounding communities. For instance, subsidence resulting from thawing permafrost can damage critical infrastructure, including roads, buildings, and pipelines (Jorgenson et al., 2001). Such impacts are expected to significantly escalate from the current decade to the middle of the century (Overland et al., 2019). Furthermore, permafrost thaw can disrupt the water balance within regions, leading to soil erosion and alterations in surface water and groundwater availability (Hinzman et al., 2013). For example, in lowland ecosystems, the loss of ice-rich permafrost has caused a transition from terrestrial to aquatic ecosystems or wetlands, whereas in upland ecosystems, permafrost thaw replaced water-dependent plant species with shrubs (Yang et al., 2010). Permafrost degradation also has the potential to significantly alter soil moisture content, which, in turn, affects the availability of soil nutrients and disturbs species composition (Yang et al., 2010).

These diverse permafrost-driven hazards will ultimately impact society, causing damage to equipment, altering food webs, affecting access to traditional harvest areas, and raising safety concerns for land-users (Walker et al., 2006; Gibson et al., 2021).

1.3 Importance of permafrost for the global climate and carbon cycle

While the local and regional consequences of permafrost thaw are undeniably significant, the most pressing concerns from scientists lie in the broader context of global climate change and the carbon cycle (Biskaborn et al., 2019; Overland et al., 2019).

Prior research, including Humphrey et al. (2021) on soil moisture-atmosphere feedback and Thackeray and Fletcher (2016) on snow albedo temperature feedback, has highlighted the existence of feedback mechanisms between the land surface and the atmosphere. These feedback mechanisms emphasize the complex interplay between the land surface and the atmosphere, playing a vital role in shaping global variations in

modeled land carbon uptake and energy balance.

As the active layer thickens, previously frozen organic matter is exposed to thawing and begins to decompose, releasing greenhouse gases such as carbon dioxide and methane, potent greenhouse gases (Schuur et al., 2008). Increases in ALT and permafrost thaw accelerate the process of decomposition. Aerobic decomposition is the dominant process in the upper part of the active layer where the soil is well-drained and oxygen is available. Anaerobic decomposition occurs in waterlogged and oxygen-depleted soils in the lower part of the active layer, and can result in the release of methane (Paquin and Sushama, 2015). Changes in ALT can also alter the distribution and activity of soil microorganisms, leading to changes in nutrient cycling and biogeochemical processes (Mondav et al., 2017).

As temperatures continue to rise, more permafrost will thaw, leading to increased decomposition and greenhouse gases released. These emissions amplify the greenhouse effect, trapping more heat in the atmosphere and further intensifying temperature increases and permafrost degradation. This cycle creates a positive feedback loop, called "permafrost carbon feedback," that amplifies the rate of climate change and hastens global heating (Schaefer et al., 2014). Crossing this threshold can occur in either direction, meaning permafrost can recover if temperatures decrease. However, the substantial carbon reserves in permafrost have accumulated over millennia. Consequently, emissions from permafrost into the atmosphere over short periods, like decades, would be irreversible within the same time frame (Schuur et al., 2015; Overland et al., 2019). Hence, there is a pressing need to accurately determine the extent of permafrost thaw and the consequent release of sequestered soil carbon into the atmosphere.

1.4 History of permafrost representation in climate models

Climate scientists have developed computational tools, known as models, which integrate a wide array of geophysical processes. The representation of permafrost in climate models has evolved significantly over time. In the early stages of climate modeling, permafrost was often not explicitly included. Models during this era primarily focused on broader climate patterns and lacked detailed representations of land surface processes. Initially, Land Surface Models (LSMs), also called land surface scheme, served as lower boundary conditions for atmospheric models. As these schemes were further developed, more intricate models emerged, encompassing a multitude of processes (Fisher and Koven, 2020). One of the earliest attempts to explicitly depict permafrost physics in a large-scale model dates back to the 1970s when Goodrich (1976) employed a basic energy balance approach with a simple soil scheme.

Towards the end of the 20th century, there was a growing concern regarding the impact of human-induced climate change, which led to an increased focus on studying

permafrost soils and their relationship with the climate system. IPCC reports recognized the significance of permafrost and its potential responses to anthropogenic warming. The first IPCC (1990) report briefly mentioned that higher temperatures "could" increase the emissions of methane in permafrost. The second IPCC (1996) report featured a section on permafrost, discussing its observed behaviour, potential reactions to human-induced warming, and the concern of a permafrost carbon feedback, thereby emphasizing the potential implications of permafrost thawing on carbon release into the atmosphere. This lead to a growing recognition of the need to include permafrost in LSMs as part of Earth System Models (ESMs). LSMs simulate the interactions between the land surface and the atmosphere, making them essential for understanding permafrost dynamics (Fisher and Koven, 2020).

In the mid-1990's, the Community Land Model (CLM) made a significant advancement in including permafrost into ESMs (Bonan, 1998). The CLM is a LSM that simulates the exchanges of energy, water, and carbon between the land surface and the atmosphere. The model includes a sophisticated representation of thermal ground regime that simulates heat diffusion and soil water transfer in a six-layer soil column. The inclusion of net land-atmosphere CO_2 exchange is an important component of the land model, enabling its use in studying the global carbon cycle. Other LSMs that include ground thermal processes have also been developed in parallel, such as the Interaction Soil-Biosphere-Atmosphere model (ISBA, Noilhan and Planton 1989), the Canadian Land Surface Scheme (CLASS, Verseghy 1991), and the Tiled ECMWF Scheme for Surface Exchanges over Land model (TESSEL, Viterbo and Beljaars 1995). These models have been used to study the response of permafrost to climate change and the impacts of permafrost thaw on the carbon cycle.

In the late 90s and early 2000s, as the computational capacity and resolution of numerical modelling systems increased, LSMs started to include more complex processes relevant to Arctic ecosystems such as such as river routing (Lohmann et al., 1996), more advanced land cover classifications (Oleson and Bonan, 2000), the nitrogen cycle (Kucharik et al., 2000), multi-layer snowpack (Bonan et al., 2002b), vertical heterogeneity in soil texture (Bonan et al., 2002b), and soil organic matter (Lawrence et al., 2008; Rinke et al., 2008). These advancements marked a significant step forward in our ability to simulate and understand the intricate feedbacks between the Arctic environment and the global climate system, facilitating more accurate assessments of future climate change scenarios in permafrost regions.

1.5 Recent advances in permafrost representation in climate models

In recent years, significant progress has been made in incorporating permafrost processes into LSMs. Chapter 2 delves into the advances made in modeling techniques,

with a particular focus on the representation of processes in Arctic regions.

However, numerous models initially designed for unfrozen soil applications have been extended for use in frozen soils. As highlighted by He et al. (2021), several of these models were not adequately validated with Arctic-conditions measurements. Their simulation accuracy remains insufficiently demonstrated (Hu et al., 2017).

Furthermore, the level of sophistication in representing these processes varies considerably among models. To assess the current state, Burke et al. (2020) conducted an intercomparison project that evaluated the representation of permafrost dynamics in both the Coupled Model Intercomparison Project 6 (CMIP6) and CMIP5 models. Their findings indicate an overall improvement in the representation of permafrost since CMIP5, although further enhancements are still necessary. Notably, certain models within the CMIP6 ensemble demonstrated improved snow insulation, resulting in reduced variability in permafrost extent compared to the CMIP5 ensemble. However, the ability of the models to accurately simulate summer thaw depths shows only limited improvement between the ensembles.

1.5.1 Systematic cold bias in LSMs

Accurate representation of heat fluxes in LSMs is paramount for understanding and simulating permafrost dynamics. A persistent challenge identified in numerous studies is the existence of a systematic 'cold bias' in soil temperature over Arctic permafrost regions. This bias has been consistently observed across various LSMs, as documented in a substantial body of research, including JULES (Dankers et al., 2011; Burke et al., 2013), JSBACH (Ekici et al., 2014), CLASS (Paquin and Sushama, 2015; Oogathoo et al., 2022), ISBA (Decharme et al., 2016; Barrere et al., 2017), Noah-MP (Li et al., 2020), and CLM5 (Dutch et al., 2022).

It is worth noting that some studies evaluating LSMs have reported the absence of a cold bias (Chadburn et al., 2015, 2017). However, these studies rely on sparse in-situ measurements that may not fully represent the entire Arctic domain. An exception is the LSM ORCHIDEE, which suggests a distinctive warm bias (Guimberteau et al., 2018).

Furthermore, evaluation studies of LSMs in the Arctic region typically concentrate on annual averages, neglecting the seasonality represented by these models. A limited number of studies have explored this aspect and have generally revealed that the discrepancy between simulated and observed soil temperatures becomes more pronounced as the freezing period progresses (Dutch et al., 2022). Similarly, Yang et al. (2022) also observed a cold bias over the permafrost regions on the Qinghai–Tibet Plateau and found this bias tends to amplify with increasing depth.

The cold bias has received minimal attention in the literature, despite indicating a significant underestimation of soil temperatures in LSM simulations. This under-

estimation is primarily because summer temperatures and MaxAGT exert a dominant influence on ALT and, thus, permafrost thaw. However, it is essential to recognize that winter temperatures leave a "memory effect" that carries into the subsequent summer. Neglecting any extra-cold transfer of temperature from winter to summer could lead to warmer-than-predicted summer temperatures, an increase in MaxAGT, and, consequently, an underestimation of the maximum permafrost thaw. Furthermore, the winter cold bias is likely to trigger interconnected effects within soil, snow, vegetation, and atmosphere physics. For instance, Oogathoo et al. (2022) observed a similar cold bias with the CLASS model and pointed out that it would likely impact various processes, including transpiration. This has a direct feedback on the climate system, particularly affecting precipitation and air temperature. The inaccurate projection of precipitation and air temperature by climate models will, in turn, affect the timing of snowmelt onset, soil temperature, available soil liquid water, and eventually the length of the growing season. Similarly, recent developments in the next version of CLM5 have resulted in an unexpected occurrence of dead vegetation in the Arctic due to extreme cold temperatures represented by the model (`https://github.com/NCAR/LMWG_dev/discussions/3`). The complex interplay among these variables highlights the need for a comprehensive approach to soil temperature modeling that provides a more holistic view of permafrost behavior within LSMs.

1.6 Research questions

The primary objective of this thesis is to offer insights into enhancing the accuracy and representation of seasonal, depth and regional variations within LSMs. It also seeks to address and improve our comprehension of the systematic cold bias observed in the majority of LSMs across permafrost regions. This endeavor is driven by the goal of achieving a more accurate understanding of permafrost dynamics and interconnected processes.

To address the overarching theme of permafrost representation and the mitigation of the cold bias in Arctic soils, this research follows a methodical progression. First, I conduct a thorough analysis of the fundamental requirements to represent permafrost dynamics. Then, I review the existing methodologies used by LSMs for characterizing soil heat and water transfer to shed light on the performance and limitations of current approaches. Subsequently, I evaluate one of these models, Community Land Model (CLM5), to understand the factors underpinning its capacity to accurately represent permafrost. Following this, I undertake two sensitivity experiments to rectify the identified shortcomings within CLM5. Finally, I extend the findings to the broader domain of LSMs for widespread applicability. This systematic approach is summarized in the following four research questions:

- **Research Question 1:** Which physical processes and variables determine a model's capability to represent permafrost soils? Which methods do land surface models use to simulate heat and water transfer in permafrost regions?
- **Research Question 2:** How accurately does the Community Land Model (CLM5) simulate permafrost and soil temperature dynamics in the Arctic region, in comparison to observed data on permafrost depth and temperature profiles? In particular, to what extent can CLM5 effectively capture the spatial and seasonal variabilities inherent in permafrost physics across the Arctic region?
- **Research Question 3:** What is the impact of using a plant functional type (PFT)-based approach to derive soil texture and soil organic matter in CLM5 on permafrost dynamics, and how does it compare to traditional approaches that use fixed soil properties?
- **Research Question 4:** To what extent can modifications to the parameterization of snow thermal conductivity improve the performance of CLM5 in reproducing snow cover dynamics over Arctic permafrost regions?

In summary, this thesis aims to improve our understanding of permafrost dynamics and provide more accurate predictions of future permafrost dynamics within LSMs, with a particular emphasis on accurate modeling of heat and water fluxes. Addressing these challenges and refining permafrost representation in LSMs is crucial for better understanding the complex interactions between permafrost soils and the Earth system.

1.7 Outline of thesis

Following this introduction, Chapter 2 explores the essential requirements for modeling permafrost dynamics and an overview of essential land model components. Chapter 3 describes the Community Land Model version 5 (CLM5) and assess its ability to reproduce permafrost dynamics against observations obtained through in-situ and remote sensing data. In Chapter 4, I conduct a first sensitivity experiment by replacing the default fixed soil properties used to derive soil texture and soil organic matter with a plant functional type (PFT)-based approach. I then examine the impact of this change on the model's ability to simulate permafrost dynamics. Chapter 5 presents a second sensitivity experiment where I replace the current snow scheme with one designed specifically for Arctic permafrost regions. I evaluate the impact of this modification on the model's ability to accurately simulate permafrost dynamics. Finally, in Chapter 6, I summarize the main findings of this thesis and provide concluding remarks on the current state of permafrost modeling, as well as the challenges that lie ahead.

CHAPTER 2

Model requirements for representation of permafrost soils

2.1 Introduction

Permafrost is a key component of the Earth's cryosphere and carbon cycle and plays an important role in the global climate system. A good representation of permafrost in LSM is crucial for accurate climate projections, as well as for understanding the feedback between permafrost and the climate system (Miner et al., 2022).

Model assessments of projecting changes in permafrost extent and carbon stocks vary widely for both historical trends and future scenarios (Burke et al., 2020). These discrepancies can, in part, be attributed to variations in the formulations of soil thermodynamics and associated variables such as the thermal properties of soil organic matter (Zhu et al., 2019) and snow pack (Langer et al., 2013). In a correlation analysis involving 27 Earth System Models (ESMs) contributing to the fifth phase of the Climate Model Intercomparison Project (CMIP5, Taylor et al. 2012), Zhu et al. (2019) underscored the critical importance of improving the parameterization of soil thermal properties. Finally, scholars urge for the development of appropriate soil schemes for permafrost regions to enhance LSMs predictability and sensitivity (Zhu et al., 2019; Burke et al., 2020; Yang et al., 2022; Hu et al., 2023).

This chapter describes fundamental concepts and core theories in soil physics of LSMs, along with their relevance to permafrost physics. After this theoretical analysis, I examine key variables that form the basis of accurately representing permafrost dynamics within LSMs. I conclude this chapter with a thorough comparison of nine LSMs from the Coupled Model Intercomparison Project Phase 6 (CMIP6, Eyring et al. 2016), aiming to provide a comprehensive evaluation of their capabilities and functionalities.

2.2 Core theories in soil physics of Land Surface Models

The core objective of land surface modeling is to capture the energy, matter, and momentum exchanges between the land surface and the atmosphere. One key component of this system is the heat and water transfers through the soil. In permafrost regions, as the freeze-thaw process unfolds, soils undergo diurnal and seasonal transformations,

regulated by the interactions among solid soil particles, ice formations, liquid water, and gas water.

2.2.1 Heat transfer

Soil temperatures are primarily controlled by heat energy transfer. Heat energy can be transported through soil via various mechanisms, including conduction, convection of heat by flowing liquid water, convection of heat by moving air, and convection of latent heat. Among these mechanisms, conduction is the most significant process (Jury et al., 1991).

Fluids have a natural tendency to balance concentration gradients in order to reach a state of equilibrium. This process leads to the movement of particles from regions of higher concentration to regions of lower concentration, driven by molecular motion. The transfer of heat through this molecular collision is referred to as conduction. The first law of heat conduction based on the first law of thermodynamics (which describes conservation of energy), also known as the differential form of Fourier's law, governs this phenomenon:

$$F = -\lambda \nabla T \tag{2.1}$$

where F is the amount of heat conducted across a unit cross-sectional area in unit time (Wm^{-2}), λ is soil thermal conductivity ($Wm^{-1}K^{-1}$), and ∇T is the spatial gradient of temperature ($Km - 1$). In one dimension, Fourier's law can be expressed as follows:

$$F_z = -\lambda \frac{\partial T}{\partial z} \tag{2.2}$$

where z is in the vertical direction, z and F_z are positive upward.

The steady-state heat flow condition describes only a limited subset of the possible heat transport processes in soil. In reality, heat transport is typically a function of both time and space, leading to transient or time-dependent flows that require a more comprehensive mathematical description than steady-state flows. The initial phase of developing a complete transient heat flow description is to define the heat conservation equation, also known as the mass balance or continuity equation:

$$\frac{\partial F_z}{\partial z} + \frac{\partial H}{\partial t} + r_H = 0 \tag{2.3}$$

where H is the heat content per unit volume, t is time in (s) and r_H is the heat sink term.

In simple terms, F_z represents the heat energy moving out of the soil volume, minus the heat energy moving into it, H is the increase in heat energy stored in soil volume,

and r_H is the amount of heat energy that disappeared from soil volume by reactions, during a time interval of $\Delta t \to 0$. r_H should be included in the heat balance whenever a source or sink of heat generates or consumes quantities of energy (e.g., radioactive soil, chemical reactions). Most LSMs assume that $r_H = 0$.

Following Jury et al. (1991), the heat content per unit volume H may be written as:

$$H = c(T - T_{ref}) \tag{2.4}$$

where c is the heat storage capacity ($Jm^{-3}K^{-1}$) and T_{ref} is an arbitrary reference temperature at which $H = 0$.

If we account for non-steady or transient conditions, we can insert the heat flux equation 2.2 and the heat content 2.4 into the heat conservation equation 2.3, resulting in the heat transfer equation:

$$c\frac{dT}{dt} = \frac{d}{dz}[\lambda_s \frac{dT}{dz}] \tag{2.5}$$

- c represents the heat storage capacity ($Jm^{-3}K^{-1}$) of the material in question. It is a measure of the amount of energy that is required to warm a unit mass for one Kelvin. It represents how much heat energy can be stored in a given amount of material, and it depends on the density (Appendix B) and properties of the material. In the context of permafrost, it represents the amount of heat energy that can be stored in the soil and permafrost at a given depth.

- $\frac{dT}{dt}$ represents the rate of change of temperature with respect to time. It describes how quickly the temperature at a given depth changes over time due to heat flow.

- λ represents the heat transfer capacity or soil thermal conductivity ($Wm^{-1}K^{-1}$) of the material in question. It is a measure of how fast heat energy can flow through the material, and it depends on the material's ability to pass over kinetic momentum on the molecular/atomic level. The conductivity of a substance increases as the molecules become closer together. Higher conductivity means heat can better propagate. Gases typically have low soil thermal conductivity, while liquids exhibit higher values. Solid materials with an organized, grid-like structure, such as crystalline substances like quartz, are excellent conductors of heat. Similarly, ice possesses a high soil thermal conductivity due to its solid and structured nature. In the context of permafrost, it represents how easily heat can be transferred through the soil and permafrost at a given depth. The estimation of soil thermal conductivity is recognized to be a challenging endeavor, given its dependence on a multitude of influencing variables (Hu et al., 2017) that are discussed throughout this chapter. This complexity is further amplified within permafrost

regions due to soil freeze-thaw processes and the presence of unfrozen water and ice.

- $\frac{dT}{dz}$ represents the rate of temperature change with respect to depth. It describes how quickly the temperature changes as you move deeper into the ground.

If the z dependence of λ_s is neglected, 2.5 reduces to:

$$\frac{dT}{dt} = K_t \frac{d^2T}{dz^2} \tag{2.6}$$

where $K_t = \frac{\lambda_s}{c}$ is the soil thermal diffusivity.

To summarize, the heat flow equation 2.6 describes the flow of heat in one dimension through a material controlled by the strength of the heat flux (the gradient), and by the material properties, soil thermal conductivity λ_s and heat capacity c, in terms of the spatial variable z and time t. This equation links the rate of temperature change at a particular depth to the second derivative of temperature with respect to that depth, revealing how temperature evolves over time within a material. Essentially, it shows that the temperature at a given location within the material is influenced by the amount of heat that is transferred into or out of that point over time, which depends on the material's thermal diffusivity.

Numerical models solve this equation by subdividing the entire soil column into smaller discrete layers. This vertical discretization is crucial due to the substantial variations in soil properties with depth. Moreover, the model establishes specific boundaries and initial conditions. For a given soil column, the boundary conditions typically involve constant temperatures at the surface (attributed to the atmosphere) and at the base (associated with the geothermal heat flux). The initial condition refers to the temperature distribution at time zero. Solving the heat transfer equation with these specified boundaries and initial conditions enables us to calculate the temperature profile in the subsurface as a function of time.

Apart from vertical heat transfer, lateral heat flow can also significantly influence the temperature evolution of permafrost and has been demonstrated to notably improve the freeze-thaw depth simulation compared to the original model (Hu et al., 2023). However, due to its high computational cost, lateral heat flow is not commonly included in LSMs (Ou and Zhang, 2022). Despite this limitation, understanding the potential impact of lateral heat flow on permafrost dynamics should be a priority for future research and model development to improve comprehension of how permafrost responds to climate change.

2.2.2 Water transfer

This thesis primarily emphasizes heat flows, but it is important to note that water flows are equally critical for accurately representing permafrost soils. Most state-of-the-art LSMs have been designed to simulate heat and soil water transfer processes separately, although recent studies have proposed considering coupled water and heat transfer developments (Hu et al., 2023). However, these coupled models have yet to gain widespread acceptance within the land surface modeling community.

Soil water input from precipitation is shared between changes in water storage in snow, ice, and soil moisture, and is balanced against losses from evapotranspiration, sublimation and river discharge. In addition to the heat transfer equation, a water balance equation is needed to describe the movement of water for modeling permafrost. Starting with the conservation of mass, we can formulate a one-dimensional vertical water flow in soils equation:

$$\frac{\partial \theta}{\partial t} = -\frac{\partial q}{\partial z} - e \tag{2.7}$$

where θ is the volumetric soil water content (mm^3 of water / mm^3 of soil), z is height (mm), q is soil water flux (kgm^-2s^-1), and e is a soil moisture sink term (mm of water mm^-1 of soil s^-1). The soil water flux q can be described by Darcy's law:

$$q = -k \left[\frac{\partial \left(\psi + z \right)}{\partial z} \right] \tag{2.8}$$

where k is the hydraulic conductivity (mms^{-1}), and ψ is the hydraulic potential (mm), which describes how a soil retains water based on its texture. If we manipulate and substitute equation 2.8 into equation 2.7, assuming e = 0, we can derive the commonly called Richards equation (Richards, 1931):

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z} \left[k \left(\frac{\partial \psi}{\partial z} + 1 \right) \right] \tag{2.9}$$

Similar to the heat flow equation 2.5, this equation is generally solved numerically in LSMs by dividing the soil column into multiple layers with upper and lower boundaries.

2.2.3 Latent heat energy

Freezing or melting is the process by which a substance undergoes a phase change from a liquid to a solid state or vice versa. The latent heat is the amount of energy required to break up the inter-particle bonds in order to melt a substance. The same amount of energy is released when the bonds of the solid phase build up. During a phase change, the addition or removal of energy to or from the substance does not cause a change in temperature. The amount of energy consumed or released during the phase change

without a change in temperature is referred to as latent heat.

Within LSMs, the incorporation of latent heat into equation 2.5 entails introducing an extra energy term that functions as either a source or a sink, which is expressed as follows:

$$c\frac{dT}{dt} = \frac{d}{dz}[\lambda_s \frac{dT}{dz}] + \rho_{ice} L \frac{d\theta_{ice}}{dt} \tag{2.10}$$

where L_f is the fusion of water latent heat (kJ/kg), ρ_{ice} is the density of ice ($kg/m3$), and θ_{ice} (m3/m3) is the amount of ice that freezes or melts within time t (s). This equation operates under several assumptions: the absence of thermal advection, no phase change involving the gas phase, and the exclusion of mechanical effects related to soil freezing, such as the expansion of the total soil volume (Gouttevin et al., 2012).

During the phase change of permafrost near 0°C, the heat transfer caused by the latent heat is greater than that caused by heat conduction (Hu et al., 2023). The large amount of energy needed to melt ice acts as an additional heat capacity, which leads to the commonly called zero-curtain effect during every autumn and spring in permafrost soils: soil temperatures stay at 0°C until the full amount of soil water (ice) is frozen (melted), although air temperatures have already decreased (increased) substantially lower (higher). The same process also introduces high thermal inertia on longer timescales wherever ice (water) contents are high. This explains why permafrost areas on the border of permafrost zones with soil temperatures close to 0°C only warm slowly with observed climate warming in comparison to cold Arctic permafrost regions, where warming rates are higher (Romanovsky et al., 2010; Zhu et al., 2019).

2.3 Key variables in the representation of permafrost soils

This section discusses key variables within permafrost physics, including soil texture, soil organic matter (SOM), snow, soil moisture and ground ice, Arctic vegetation, atmospheric forcings, and lower boundary fluxes. Additionally, I explore their integration and implementation within LSMs. These factors wield significant influence over soil heat transfer and storage capacities, directly impacting permafrost soils temperature. A comprehensive understanding of their interplay and effects on heat and water transfers within permafrost soils is paramount for accurately modeling permafrost in LSMs (Nicolsky and Shakhova, 2010; Yang et al., 2022).

2.3.1 Soil texture

Soil texture is a fundamental variable to the composition and properties of soils (Warrick, 2002; He et al., 2021). A soil can be described as a ”soil matrix” composed of

soil material, water and/or ice, solutes and air. Due to this intricate composition, soils exhibit a more complex response to changes in energy than pure substances.

In the context of LSMs, soil texture typically represents the relative percentages of sand (particle size of 0.05 to 0.2 mm), silt (particle size of 0.002 to 0.05 mm), and clay (particle size below 0.002 mm) (Warrick, 2002).

Mineral soil composition plays a critical role in determining soil properties and thermal dynamics in permafrost regions. The impact of minerals on soil thermal conductivity depends on two main factors:

- Firstly, soil texture effects porosity (the fraction between the volume of pore spaces in the soil and the volume of soil), which subsequently influences latent heat exchange during freeze and thaw cycles (Jury et al., 1991; Ochsner et al., 2001; Balland and Arp, 2005; Zhu et al., 2019). Sandy soils have higher porosity, allowing for an increase in thermal conductivity with more moisture content, while cohesive soils with silty or clayey texture show a smaller variation of conductivity with moisture (Luo et al., 2009).

- Secondly, different minerals found in soils possess varying thermal conductivities, affecting the overall heat transfer through the soil (Jury et al., 1991). For instance, sandy soils are characterized by a high quartz content, which is one of the minerals with the highest soil thermal conductivity. In contrast, clay soils have higher percentages of silicate minerals, resulting in weaker heat conduction (Luo et al., 2009).

In addition to regulating heat transfer, soil texture also plays a crucial role in various non-thermal soil processes, including water infiltration, nutrient retention, carbon sequestration and storage, protection of soil organic carbon against microbial decomposition (Kravchenko and Guber, 2017), and the exchange of water and energy between the land surface and the atmosphere (Shao et al., 2022). The unique hydraulic properties of soils with different textures influence water movement and its retention capacity, ultimately determining water availability for plant growth and the exchange of heat and moisture between the soil and the atmosphere. For example, while sandy soils are extremely porous, clay soils have a higher water holding capacity due to their smaller pore sizes, facilitating rapid water flow. The greater water-holding capacity also induces a greater latent heat release during freezing (Zhu et al., 2019).

In theory, sandy soils tend to experience warmer winter soil temperatures due to the higher amount of latent heat released during freezing. In contrast, during the summer months, sandy soils demonstrate colder temperatures due to enhanced water drainage and lower water retention capacity compared to soils with higher clay or silt content. However, using a network of 184 in-situ stations, Zhu et al. (2019) only found a weak positive correlation between sand content and thermal conductivity.

Further study is required to better understanding the complex interplay between soil texture, its properties, and ecosystem processes (Dai et al., 2019; Zhu et al., 2019; Hu et al., 2023).

Chapter 4 provides a detailed discussion of how soil texture is included in LSMs.

2.3.2 Soil organic matter

Soil organic matter (SOM) plays a crucial role in providing a broad range of ecosystem services including production of food and fiber, nutrient supply, and flood and erosion controls. Some scholars also refer to it as organic matter content (OMC).

Arctic soils, particularly in Northern Eurasia, are rich in organic carbon because the low soil temperatures in these regions inhibit the decomposition of dead plant material that accumulates over time, thereby forming commonly called peat deposits (Minayeva et al., 2016). A peat deposit refers to an accumulation of partially decomposed organic matter, primarily composed of plant remains such as mosses, sedges, and other vegetation, which have gathered in waterlogged environments over extended periods. The waterlogged conditions cause organic material to decompose slowly, forming a distinctive layer of peat.

Figure A.1 in Appendix A represents the soil organic carbon content of two soil horizons (0–30 and 30–70 cm) aggregated at a 0.5° by 0.5° horizontal resolution and estimated from Global Soil Organic Carbon Map (ITPS, 2020). The data represent a unique global soil inventory providing information on the morphological, chemical and physical properties of soils at very high resolution.

Soil organic carbon exhibits distinct hydraulic and thermal properties when compared to mineral soil. The presence of organic carbon augments available water capacity (Hudson, 1994), soil porosity, and saturated hydraulic conductivity (Lawrence and Slater, 2008; Decharme et al., 2016; Zhu et al., 2019). These enhancements lead to soils with elevated organic content serving as insulators, thereby retarding the rate of heat transfer into the soil and resulting in significantly reduced soil thermal conductivity, while concurrently elevating soil heat capacity in comparison to mineral soils, ultimately manifesting as lower temperatures during the summer months. This phenomenon has been particularly studied over Arctic permafrost regions (Rinke et al., 2008; Lawrence and Slater, 2008; Dankers et al., 2011; Paquin and Sushama, 2015; Decharme et al., 2016; Domine et al., 2016).

The insulating influence of SOM during summer exhibits greater significance in contrast to its relatively modest warming effect during winter, primarily due to the insulating nature of snow cover (Zhu et al., 2019).

How SOM is incorporated in LSMs is explored in Section 2.4.2.3 and Chapter 4.

2.3.3 Snow

Snow cover, also called snowpack, plays a significant role in permafrost regions as it covers a large portion of the land surface, exerting essential influences on permafrost physics (Sturm et al., 1997). Royer et al. (2021) show that current state-of-the-art snow models fail to represent the main traits of Arctic snowpacks.

Similar to Blome (2014), we can identify four key physical properties or processes associated with snow that are crucial in modulating permafrost physics:

- Insulating ”barrier”: Snow acts as an insulating ”barrier” between the soil and the overlying air from fall to early spring.
- Optical ”reflector”: With its unique optical and radiative properties, snow impacts the surface energy balance.
- Latent heat ”sink”: The considerable amount of latent heat consumed during snow melting acts as an energy sink in the surface energy balance.
- Snow/soil hydrology ”regulator”: The accumulation and release of water due to freezing and melting significantly impact soil hydrology.

These properties have subsequent effects on soil temperature and water content, which in turn have feedback effects on snow thickness. Each of these points are further explored below.

2.3.3.1 Insulating ”barrier”

The low soil thermal conductivity of snow is primarily due to air pockets present between ice crystals. These air pockets act as barriers to heat transfer by conduction, as air is a poor conductor of heat. When heat energy attempts to pass through the snow, it must traverse these air-filled gaps, significantly slowing the heat transfer process. Consequently, a snow cover effectively decouples air and soil temperatures, resulting in a warming of the soil surface.

Numerous scholars have demonstrated that snow cover significantly impacts the thermal regime of permafrost by providing insulation to the ground during Arctic winters (Nicolsky and Shakhova, 2010; Park et al., 2015; Li et al., 2021). This influence is closely tied to both the extent and timing of variations in snowfall.

2.3.3.2 Optical ”reflector”

Snow cover can substantially modify the albedo of a land surface, thereby affecting the amount of solar radiation absorbed by the surface. The extent of this effect depends on factors such as the shape and size of the snow grains, the solar zenith angle, and cloud

conditions. In most cases, the effect leads to a reduction in absorbed energy, as snow has one of the highest albedo values among natural land surface materials, ranging from 0.6 for old, wet snow to 0.85 for fresh snow. This modification of albedo affects the emission of thermal radiation, thereby influencing surface temperatures and resulting in a cooling effect on the land surface when compared to bare surfaces.

The net effect of albedo variation depends on the geographical area and season. During autumn, the albedo of fresh snow is high, and, with a reduced sun elevation angle, the developing snow cover further decreases energy input to the land surface. In spring, the albedo of old snow is reduced, while the sun's elevation angle is almost at its yearly maximum. Consequently, the cooling effect of snow cover is generally larger in spring than in autumn.

Under cloudy conditions, snow cover reacts differently, leading to the development of near-surface temperature inversions due to strong radiative cooling. On the other hand, moist, cloudy conditions with a high proportion of back-scattered longwave radiation result in higher snow surface temperatures.

2.3.3.3 Latent heat "sink"

Snow can influence the partitioning of energy between sensible and latent heat fluxes. The latent heat of the fusion of water is 334 J/g, whereas the volumetric heat capacity of ice is only 2.1 J/gK. The melting of snow cover during spring acts as a significant energy sink, warming of the surface, albeit for a short duration of two weeks or less. The effects on the surface energy balance to date are not yet fully understood. However, due to its relatively short duration, the mean annual ground temperature (MAGT) is not strongly influenced by this effect (Blome, 2014).

2.3.3.4 Snow/soil hydrology "regulator"

Snow accumulation and the freezing of soil water result in a net storage of water from October to April. When the snow melts and soil thaw begins in spring, water becomes available for plant uptake and growth. However, this process is confined to a short period of less than two weeks. As the land surface remains frozen, the infiltration capacity is low, leading to significant runoff into rivers with peak discharge rates occurring in May-June (Hu et al., 2023) and large flatland area flooding from May to September.

During the initial stages of soil thawing, high surface moisture values correspond to the melting of ice in near-surface layers, while deeper layers remain frozen and less permeable (Blome, 2014). Consequently, the drainage capacity is limited, leading to the accumulation of high soil moisture within the shallow thawed upper layers. Furthermore, the refreezing of infiltrated snowmelt water further exacerbates this phenomenon (Swenson and Lawrence, 2014). As the warm season progresses and the ALT deepens,

the storage capacity for water increases. The permafrost table acts as a barrier within the deeper soil, facilitating the development of lateral flow, which contributes to the formation of slower subsurface runoff. This enlarged water storage capacity, coupled with improved drainage, can lead to drier conditions within the upper soil layers during this period of the year.

2.3.3.5 Snow in LSMs

Most LSMs include a snow module in their code, which encapsulates snow accumulation, distribution, and melt processes. Operating at the interface between the atmosphere and the land surface, the snow module plays a fundamental role in simulating the interactions between meteorological conditions, energy exchange, and the evolving snowpack. By encompassing a range of physical and thermal properties unique to snow, such as thermal conductivity, albedo, density, and water content, the snow module enables LSMs to accurately capture the complex feedback mechanisms between the snowpack and the surrounding environment. In Chapter 5, I provide an example of a snow module from CLM5.

2.3.4 Soil moisture and ground ice

Soil moisture refers to the amount of water present within the pores of the soil, usually expressed as a percentage of the total soil volume. It represents the liquid and frozen water content within the soil matrix, including both the water adsorbed on soil particles and the water filling the soil pores.

The distinct cold Arctic temperatures give rise to a soil moisture pattern specific to permafrost regions, governed by freezing and thawing processes. Consequently, soil moisture content plays a pivotal role in permafrost modeling, influencing soil's thermal and hydrological processes (He et al., 2021).

Despite often being characterized by a dry climate, permafrost regions frequently exhibit high soil frozen water contents. This can be attributed to the low temperatures, which cause soil moisture to freeze during the cold months. As a result, the soil moisture becomes "trapped" in a frozen state and is not readily available for evapotranspiration or water discharge at the beginning of the warm season until it melts. This extended resting period of soil moisture, combined with freezing-induced cracks in the soil, allows water from the surface to infiltrate and accumulate, eventually leading to the formation of large ice wedges and frozen soil underneath, commonly called Yedoma permafrost. High thermal conductivity of ice generally leads to increased soil thermal conductivities as compared to the same soil without ice (Farouki, 1981). Furthermore, soil moisture significantly contributes to the latent heat energy, as explained earlier.

Within LSMs, soil moisture depends on the water fluxes represented by the model,

typically formulated using the aforementioned one-dimensional form of the Richards equation. Yet, physical processes associated with ground ice, such as soil profile cracking, are generally omitted.

According to the literature, precise representation of soil moisture content is key for permafrost modeling, given its substantial influence on permafrost dynamics' thermal and hydrological processes. For instance, Li et al. (2019) underscore that soil moisture content regulates soil thermal conductivity. Meanwhile, Lee et al. (2014)'s model sensitively study examining large-scale repercussions of ground ice on projected permafrost temperatures revealed a significant impact on soil temperature's vertical structure and potential influence on permafrost thaw timing and rate. Importantly, Burke et al. (2020) found that current CMIP6 ESMs overlook the critical ground ice process in permafrost soils, highlighting a gap in these models' predictions of permafrost dynamics.

2.3.5 Arctic Vegetation

Several studies have stressed the importance of Arctic shrubs' impact on permafrost regions (Lawrence and Swenson, 2011; Loranty et al., 2011; Bonfils et al., 2012; Myers-Smith et al., 2011; Pearson et al., 2013; Chen et al., 2021; Heijmans et al., 2022). For example, Royer et al. (2021) significantly increased the performance of their model by implementing a more complex representation of the interaction between Arctic plants and snow. Incorporating better vegetation physics in LSMs should be a priority in future research.

Arctic plants possess unique water holding capacity characteristics. The extreme conditions of permafrost, including cold soil temperatures, limited nutrient availability, and the narrow root zone confined to the shallow active layer thickness (ALT), make it challenging for larger plants to thrive. In areas with discontinuous permafrost and deeper ALT, taller tree species with deeper roots like aspen can be found. Conversely, smaller species like black spruce have better adaptations and can survive in more northern regions.

Mosses, with their sponge-like structure, can hold nearly their own volume of water, which results in higher conductive properties, particularly when frozen. They act as excellent insulators when dry, which is often the case at the end of summer (Blome, 2014).

Moreover, non-vascular vegetation types, such as mosses and lichens, have a tendency to "protect" permafrost through their specific seasonal thermal impacts on the ground, in contrast to vascular vegetation types like grasses and shrubs, which contribute to summer warming and deepen the ALT (Rinke et al., 2008).

Vegetation also influences heat transport into the ground through its interaction with snow. The most abundant vegetation types, mosses and lichens in the Arctic typically

reach a maximum height of a few centimeters, limiting their ability to trap snow to the same extent as taller vegetation types like tussock grasses. Consequently, winter cooling is more efficient in regions dominated by low vegetation (Blome, 2014). However, the reverse effect is also observed: a thicker snow cover that persists until late spring delays soil warming at the beginning of the warm season, resulting in a relatively cooler ground compared to adjacent locations with a thinner snow cover. Trees and shrubs, on the other hand, trap a substantial amount of snow in their canopies, which reduces snow depth at the ground and leads to lower soil temperatures (Myers-Smith et al., 2011). Furthermore, forested areas accumulate thick organic matter layers, which decompose poorly due to the cold climatic conditions. Combined with the shading effect of the canopy and reduced solar energy input, forests generally contribute to permafrost stabilization.

Most LSMs represent plant dynamics using a homogeneous single-layer canopy (known as the "big-leaf" approach) assuming that all vegetation within a grid cell behaves as a single, uniform "big leaf" when interacting with the atmosphere. In reality, there can be various types of vegetation within a grid cell with different characteristics such as leaf area index (LAI), stomatal conductance, and photosynthetic capacity. By simplifying these complexities, the big-leaf approach is computationally less demanding than a representation that considers individual plant types or even leaves within the grid cell.

LSMs commonly group plant species into plant functional types (PFTs). These categories represent the diversity of vegetation types and their specific properties, such as photosynthesis pathways, LAI, rooting depth, or specific leaf area. They are parameterized based on the averaged characteristics of the vegetation within the grid cell. Different vegetation types have distinct effects on energy and water exchanges with the atmosphere, impacting local climate, evapotranspiration rates, and soil heat transfer.

PFTs influence soil heat transfer in several ways depending on the model:

- Albedo and canopy structure: Albedo is the amount of solar radiation reflected by the land surface. Vegetation with darker leaves and a more open canopy structure tend to absorb more solar radiation, increasing the surface temperature and consequently affecting soil heat transfer.

- Evapotranspiration: PFTs regulate evapotranspiration rates, which influence the cooling effect on the land surface. Transpiration, the process of water movement through plants and its subsequent release into the atmosphere, cools the vegetation and surrounding soil. Higher transpiration rates from certain PFTs can reduce the soil temperature.

- Rooting depth: Deeper-rooted vegetation can extract water from deeper soil layers, affecting soil moisture distribution and heat transfer at different depths.

- Litter and organic matter: The type and amount of litter and organic matter added to the soil by vegetation influence soil properties, such as soil thermal conductivity and heat capacity. These properties impact the transfer of heat within the soil profile.

- Land cover change: Changes in land cover and vegetation distribution, driven by human activities or natural processes can significantly alter local and regional climate conditions, including soil heat transfer patterns.

There are two common PFTs used by LSMs to describe Arctic vegetation:

- C_3 Arctic grass - C_3 arctic grass is a type of vegetation that falls under the C_3 photosynthetic pathway, one of the three major pathways through which plants fix carbon during photosynthesis. C_3 plants generally have higher photosynthetic rates at lower temperatures, making them well-suited for colder regions. These grasses contribute to the vegetation cover in Arctic ecosystems and play a vital role in the carbon and energy exchange between the land surface and the atmosphere.

- Broadleaf deciduous shrub - boreal - Broadleaf deciduous shrubs (BDS) are characterized by broad-shaped leaves that shed seasonally (deciduous) in contrast to needle-shaped leaves of coniferous trees. Boreal shrubs are adapted to the cold and harsh conditions of the boreal forest, which is a subarctic biome characterized by cold winters and short growing seasons. These shrubs are an essential component of the understory vegetation in boreal forests and play significant roles in ecosystem processes, including energy balance, nutrient cycling, and carbon storage.

However, some experts argue that relying on just two PFTs may be insufficient and recommend the inclusion of a greater variety (Sulman et al., 2021).

2.3.6 Atmospheric forcings

Permafrost is a thermal phenomenon, and thus depends on the temperature of the overlying air masses. Additionally, precipitation, wind, solar radiation, specific humidity also play a crucial role in permafrost dynamics by determining the energy balance at the surface.

In LSMs, these atmospheric conditions, or forcings, are often represented using reanalysis data, which provide a consistent and comprehensive record of atmospheric conditions. Reanalysis data are generated by assimilating observations from multiple sources, including satellites, radiosondes, and surface stations, into a global atmospheric model. This produces a gridded dataset of atmospheric variables, such as air

temperature, precipitation, and wind speed, with varying spatial and temporal resolutions.

However, the use of reanalysis data in LSMs also has limitations, as these data may contain biases and errors that can affect the simulation of permafrost dynamics. For example, reanalysis data may not capture local-scale variations in atmospheric conditions, such as the effect of topography on wind patterns. In addition, reanalysis data may be subject to errors in the assimilation process, such as observation data or model physics error. Therefore, it is important to evaluate the performance of LSMs using reanalysis data against independent observations, such as in-situ measurements or remote sensing data.

2.3.7 Lower boundary fluxes

Lower boundary fluxes refer to the exchanges of energy, moisture, and other relevant properties between the land surface and the lowermost layer of the model, which typically represents the underlying bedrock.

According to Hu et al. (2023), the bottom layer depth must be below minimum 30 meters to run a realistic simulation of the soil temperature profile over time, especially for permafrost. However, in the case of deeper soil configurations, the presence of extended system memories makes it challenging to attain equilibrium, thereby adding an additional layer of complexity to the process of model initialization.

A standard practice in LSMs is to default the heat flux at the base of the bedrock column to zero. Nevertheless, certain experiments have deviated from this practice, replacing the zero-flux lower boundary condition with a geothermal heat flux, resulting in positive simulation outcomes (Xiao et al., 2013).

2.4 Comparison of Land Surface Models used in CMIP6

LSMs are crucial components of Earth System Models (ESMs) used to represent what is happening under and at earth's surface. Different LSMs employ various schemes and parametrizations to simulate the dynamic exchange of heat, water, and energy in the land and in exchange with the atmosphere.

To provide a comprehensive analysis, I selected a range of LSMs commonly utilized in ESMs, detailed in table 2.1, with a focus on models that participated in CMIP6 (Eyring et al., 2016). I exclusively considered models that include studies on permafrost, have good documentation, and have a maximum soil depth exceeding 3 meters. This selection ensures the inclusion of relevant models with a particular emphasis on permafrost-related research. A list of models that were not selected from CMIP6 are in table 2.2.

Table 2.1 Land Surface Models and Earth System Models

Land Surface Models (old versions)	Earth System Model(s)	Reference paper	Arctic paper(s) with model version used in this study
CTSM/CLM5 (CLM4.5, CLM4)	CESM2, FGOALS, NorESM2, TaiESM1, CMCC, SAM0-UNICON	Lawrence et al. (2019)	Burke et al. (2020), Dutch et al. (2022)
ISBA/Surfex 8.0c	CNRM	Decharme et al. (2016)	Burke et al. (2020), Barrere et al. (2017), Decharme et al. (2016)
ELMv0	E3SM	Golaz et al. (2019)	Burke et al. (2020)
GFDL-LM4 (LM3)	GFDL, KIOST-ESM	Zhao et al. (2018)	Burke et al. (2020), Milly et al. (2014)
ORCHIDEE-MICT v8.4.1	IPSL	Guimberteau et al. (2018)	Guimberteau et al. (2018), Chadburn et al. (2017)
JULES	UKESM1, HadGEM3	Best et al. (2011)	Chadburn et al. (2015)
JSBACH	MPI-ESM, AWI-CM, NESM3	Reick et al. (2021)	Chadburn et al. (2017), Ekici et al. (2014)
MATSIRO6.0	MIROC6	Takata et al. (2003)	Burke et al. (2020)
CLASSIC1.0	CanESM5	Melton et al. (2020)	Paquin and Sushama (2015), Melton et al. (2019)

Table 2.2 Land Schemes not selected

Land Scheme	Earth System Model(s)	Why not selected in this study?
CABLE	ACCESS-ESM1-5	Too shallow (2.9m)
BCC-AVIM2	BCC-CSM2-MR	Too shallow (2.9m)
CoLM	CAMS-CSM1-0, CAS-ESM2.0	Too shallow (2.9m)
HTESSEL	EC-Earth3, ERA5-Land	Too shallow (1.4m)
GISS LSM	GISS-E2-1-G	Too shallow (2.7m)
?	KACE	Poor documentation
HAL 1.0	MRI-ESM2-0	Poor documentation

In the following, CTSM/CLM5 are referenced as CLM5, ISBA/Surfex 8.0c as ISBA, and ORCHIDEE-MICT v8.4.1 as ORCHIDEE for simplicity.

Throughout the next two sections, I explore and compare the different soil schemes employed within these selected LSMs. By examining their distinct characteristics, strengths, and limitations, I aim to enhance our understanding of how these models simulate soil processes and contribute to the overall comprehension of land-atmosphere interactions.

I do not delve into detail of the potential biogeochemistry modules integrated within these LSMs. Instead, my primary emphasis are on heat diffusion processes. By narrowing my scope, I can effectively explore the essential aspects of these soil schemes and their implications for the broader understanding of land surface dynamics over permafrost areas.

ESMs employed in CMIP experiments exhibit a broad spectrum of sophistication in their incorporation of land surface processes. Reflecting this diversity, prior CMIP analyses (Koven et al., 2013; Burke et al., 2020) have demonstrated substantial disparities in the simulation outcomes of permafrost-related parameters across distinct models. Consequently, this section is dedicated to an in-depth comparison and understanding of these models.

Concurrent with the CMIP initiatives, the Land Surface, Snow, and Soil Moisture Model Intercomparison Project (LS3MIP, Van Den Hurk et al. 2016) aims to evaluate the representation of land surface processes within LSMs and to identify inherent systematic biases. Nevertheless, as of the present, the findings from this initiative remain unpublished.

Numerous LSMs have been applied in various studies, often with configurations differing from the one under current discussion. The objective here is to systematically compare these models in their default configurations, which is employed in CMIP6. An exception is made for ORCHIDEE, which is examined in an alternative configuration (Guimberteau et al., 2018). This specific configuration was designed with a focus on the Arctic region and offers a more extensive documentation compared to the still-in-development ORCHIDEE v2.0.

2.4.1 Soil discretization

Figure 2.1 provides a visual representation to compare different soil vertical discretizations used by all LSMs. The vertical discretization of soil layers is an essential aspect of LSMs, as it directly influences the representation of soil processes.

Most LSMs employ a geometrical distribution of soil layers, with more layers allocated at shallower depths and fewer layers at greater depths. However, some models deviate from this general pattern. Notably, ISBA, GFDL-LM4, and JULES exhibit less

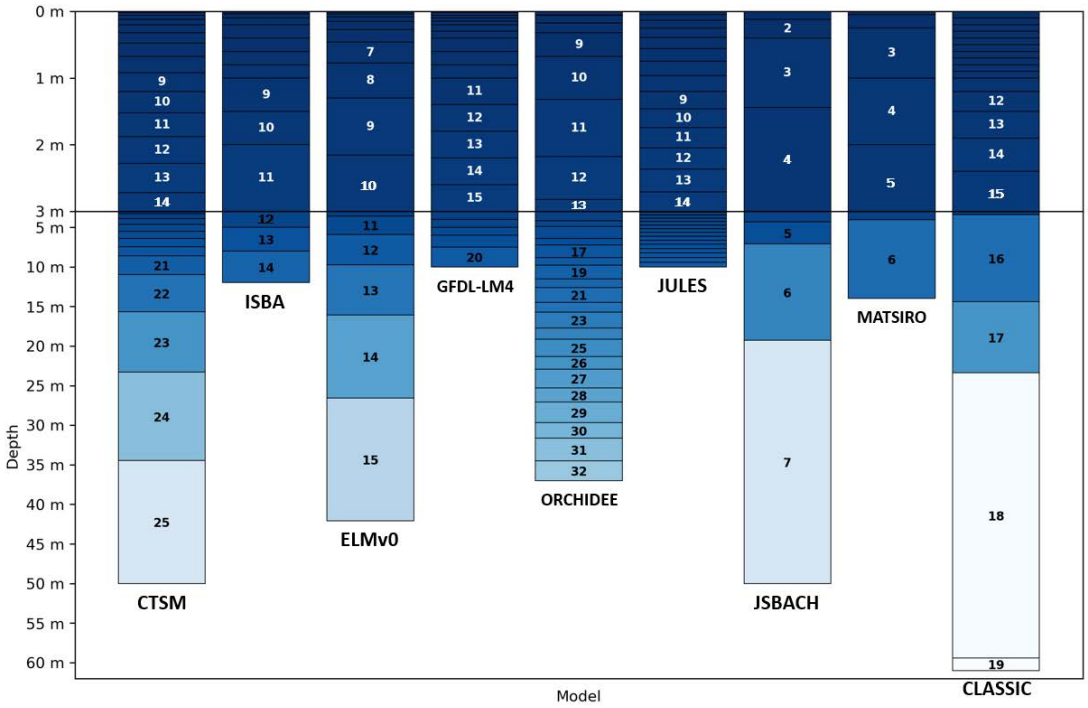


Fig. 2.1 Soil vertical discretization of LSMs. Each rectangle represents a soil layer. Note that the y-axis is not linear. Lighter blue colours indicate deeper layers.

pronounced variations in layer density along the soil column.

The number of soil layers and their distribution, as well as the maximum depths considered by the models, show significant variability. Simple models like MATSIRO6.0 use only six layers, while more complex models like ORCHIDEE have 32 layers. Other models, such as JULES and CLM5, have a comparable number of layers but differ in their maximum depths.

To accurately simulate soils in permafrost regions, two key requirements have been identified from previous studies:

- Lawrence et al. (2019) emphasize the need for high resolution within the top 3 meters of the soil column to improve the representation of the active layer thickness (ALT) within the permafrost zone.
- Alexeev et al. (2007) demonstrate that the soil column should extend to a minimum depth of 30 meters to properly resolve the annual cycle in temperature. In addition, only a geothermal steady heat flow at depths greater than 30 m as the lower boundary condition is suitable to simulate 100-year scale permafrost changes (Hu et al., 2023).

Based on the two key necessities identified, we can categorize the nine models into four distinct groups:

- Group 1: Models with low resolution in the top 3 meters and a shallow maximum depth. This group includes MATSIRO6.0.

- Group 2: Models with low resolution in the top 3 meters but with a relatively deep maximum depth. This group includes JSBACH.
- Group 3: Models with high resolution in the top 3 meters but with a shallow maximum depth. This group includes ISBA, GFDL-LM4, and JULES.
- Group 4: Models with high resolution in the top 3 meters and a deep maximum depth. This group includes CLM5, ELMv0, ORCHIDEE, and CLASSIC.

Some models apply different physics to specific soil layers. Notably, CLM5, ELMv0, and ORCHIDEE designate the last 5, 5, and 21 layers, respectively, as "bedrock." These layers are hydrologically and biogeochemically inactive in CLM5 and ELMv0, while in ORCHIDEE, they remain biogeochemically active. This distinction is a valuable asset, as hydrological physics become less important in these deep layers (Guimberteau et al., 2018), allowing for substantial computational savings.

2.4.2 Soil physics

Table 2.3 summarizes the differences in soil physics among those models. Most models represent the process of heat transport in soil with the 1-D heat flow equation, including latent heat as a sink or source (see Equation 2.10). They employ a similar discrete finite-difference form to solve this equation. However, JSBACH, MATSIRO6.0, and CLASSIC1.0 neglect latent heat in their calculations.

As expected, none of the models under study incorporate lateral heat flow, primarily due to its computational cost.

The key differences lie in how they estimate soil heat capacity and soil thermal conductivity. Studies have consistently demonstrated the significance of these two parameters in land surface modeling (Lawrence and Slater, 2008; Hu et al., 2017; Zhao et al., 2018; Dai et al., 2019). They play a crucial role in governing a wide range of physical, biological, and chemical processes by influencing energy partitioning at the ground surface and energy distribution within subsurface soil layers.

2.4.2.1 Soil heat capacity scheme

Most models use a similar soil heat capacity scheme derived from de Vries (1963), which depends on the heat capacities of the soil solid, liquid water, and ice constituents:

$$c_i = c_{s,\,i}\,(1-\phi_i) + \frac{w_{ice,\,i}}{\Delta z_i} C_{ice} + \frac{w_{liq,\,i}}{\Delta z_i} C_{liq} \tag{2.11}$$

where $c_{s,i}$ represents the heat capacity of soil solids, ϕ_i is the porosity, Δz_i is the depth of the soil layer i, and $w_{liq,i}$ and $w_{ice,i}$ are the liquid water and ice contents,

Table 2.3 Soil Physics Characteristics of Land Surface Models

LSM	References	Max. no layers (max. depth)	Latent heat included	Soil heat capacity scheme	Soil thermal conductivity scheme	Organic carbon included	Soil hydrology parametrization
CLM5	Lawrence et al. (2019)	20 soil + 5 bedrock layers (50 m)	Yes	de Vries (1963)	Lawrence and Slater (2008)	Yes	Richards equations (BC model)
ISBA	Boone et al. (2016), Barrere et al. (2017), Decharme et al. (2016), Le Moigne (2018)	14 soil layers (12 m)	Yes	de Vries (1963)	Peters-Lidard et al. (1998)	Yes	Mixed form of Richards equation (BC or VG model)
ELMv0	Golaz et al. (2019), Oleson et al. (2013)	15 soil layers (42.1 m)	Yes	de Vries (1963)	Lawrence and Slater (2008)	Yes	Richards equations (BC model)
GFDL-LM4	Zhao et al. (2018), Milly et al. (2014)	20 soil layers (10 m)	Yes	Fixed depending on soil type	Fixed depending on soil type	No	”Bucket” scheme
ORCHIDEE	Guimberteau et al. (2018)	32 soil layers (37 m) - soil hydrology only in 11 layers (2m)	Yes	de Vries (1963)	Wang et al. (2016a)	Yes	Richards equations (VG model)
JULES	Best et al. (2011), Chadburn et al. (2015)	28 soil layers (10 m)	Yes	Cox et al. (1999)	Peters-Lidard et al. (1998) or Cox et al. (1999) or Dharssi et al. (2009)	Yes	Richards equations (BC or VG model)
JSBACH	Reick et al. (2021)	7 soil layers (30.1 m)	No	de Vries (1963)	Johansen (1975)	Yes	Richards equations (VG model)
MATSIRO6.0	Guo et al. (2021), Saito (2008)	6 soil layers (14 m)	No	Saito (2008)	Sass et al. (1971)	Yes	Richards equations (BC model)
CLASSIC1.0	Melton et al. (2020),Verseghy (2012)	20 soil layers (61 m)	No	Verseghy (1991)	Côté and Konrad (2005)	Yes	Richards equations (BC model)

respectively. C_{liq} and C_{ice} are the specific heat capacities (J kg^{-1} K^{-1}) of liquid water and ice.

Most other schemes are derived from de Vries (1963) and include additional processes. For example, Cox et al. (1999) incorporates the apparent heat capacity associated with phase change, while Saito (2008) and Verseghy (1991) do not consider depth variation or porosity in their schemes. Only GFDL-LM4 adopts a fixed soil heat capacity dependent on soil type, which may adversely impact soil temperature representation.

2.4.2.2 Soil thermal conductivity scheme

Table 2.4 Soil Thermal Conductivity Schemes used by LSMs and Comparisons. Results from He et al. (2021).

Soil Th. Cond. Scheme	Model(s)	RMSE	MAD
Lawrence and Slater (2008), Peters-Lidard et al. (1998), Johansen (1975)	CLM5, ELMv0, ISBA, JULES, JSBACH, CLASSIC1.0	0.6	0.17
Wang et al. (2016a)	ORCHIDEE	0.49	0.03
Cox et al. (1999)	JULES	0.68	0.31
Dharssi et al. (2009)	JULES	1.6	0.76
Sass et al. (1971)	MATSIRO6.0	1.09	0.98
Côté and Konrad (2005)	CLASSIC1.0	0.58	-0.28

Soil thermal conductivity plays a crucial role in determining the vertical profiles of soil heat flux and soil temperature, and it significantly influences soil freeze-thaw cycles. Due to its strong influence on soil thermodynamics, soil thermal conductivity is considered one of the most important physical parameters in land modeling studies (Li et al., 2019; Dai et al., 2019; He et al., 2021; Hu et al., 2023).

As exemplified in this chapter, soil thermal conductivity is dependent on numerous variables, king its estimation challenging in soil sciences. In response to the controlling variables, several theoretical and empirical models have been developed to predict soil thermal conductivity. Column 6 in Table 2.3 presents eight distinct soil thermal conductivity schemes from the selected list of LSMs. Notably, some models allow for the use of different schemes, such as JULES. Among the selected LSMs, GFDL-LM4 adopts a fixed soil thermal conductivity dependent on soil type, which may adversely affect soil temperature representation.

A detailed description of these soil thermal conductivity schemes can be found in Dai et al. (2019) or He et al. (2021). Chapter 4 elaborates on the scheme used in the model employed for this study (Lawrence and Slater, 2008).

In order to assess the effectiveness of these schemes, He et al. (2021) conducted a comparison involving 39 soil thermal conductivity schemes. This comparison was

carried out using a comprehensive dataset of soil thermal conductivity measurements in 27 frozen soils, which included 331 measurements located in various regions around the world. Table 2.4 reports the results of this study, focusing on the eight schemes used by the models evaluated in this chapter. The variability of the models around the mean is assessed using the Root Mean Square Error (RMSE - Appendix B - Equation B.6) and the mean is evaluated using the Mean Absolute Deviation (MAD - Appendix B - Equation B.7). Both RMSE and MAD should approach 0 to indicate close agreement with the observations.

The results do not differ for Johansen (1975), Peters-Lidard et al. (1998), and Lawrence and Slater (2008) due to the assumption made by He et al. (2021) that the thermal conductivity of quartz is equivalent to that of sand (although this assumption overlooks the impact of pore space in sands). Additionally, their study does not account for the thermal conductivity of soil organic matter (SOM). Extending this study to include such differences would be highly valuable.

Based on He et al. (2021) findings, the Wang et al. (2016a) scheme used by ORCHIDEE yields the best results, followed by Côté and Konrad (2005). Lawrence and Slater (2008), Peters-Lidard et al. (1998), and Johansen (1975) have similar scores, with Cox et al. (1999) closely behind. Dharssi et al. (2009) and Sass et al. (1971) schemes perform relatively worse compared to He et al. (2021) observations. Interestingly, all schemes tend to overestimate soil thermal conductivity, with the exception of Côté and Konrad (2005), indicating a potential for improvement by considering the impact of SOM.

This comparison is limited to a single study, but provides valuable insights into the soil thermal conductivity used by the nine models. Further research and extended studies would enhance our understanding of the soil thermal conductivity representation in LSMs.

2.4.2.3 Soil organic matter parametrization

A substantial body of land surface modeling literature underscores the pivotal role of soil organic matter (SOM) in governing soil thermodynamics within permafrost regions (Rinke et al., 2008; Lawrence et al., 2008; Chadburn et al., 2015; Decharme et al., 2016; Zhu et al., 2019; Zhang et al., 2021). Although all models, with the exception of GFDL-LM4, incorporate a representation of SOM, they diverge in terms of parameterization and levels of sophistication.

Lawrence and Slater (2008) were pioneers in integrating SOM within a large-scale climate model. They introduced a linear weighted combination of organic soil properties with standard mineral soil properties. CLM5 and ELMv0 have adopted this approach, which is discussed in Chapter 4. ISBA, ORCHIDEE, and JULES also base their

SOM parameterization on Lawrence and Slater (2008), albeit with added complexity. For instance, ISBA incorporates a pedotransfer function linking soil water retention at various pressure levels to the fiber content of organic soils (Decharme et al., 2016). ISBA and JULES calculate soil thermal conductivities as geometric averages of organic and mineral soils, in contrast to the linear approach proposed by Lawrence and Slater (2008). Meanwhile, ORCHIDEE introduces a parametrization of the SOM effect on soil water capacity (Guimberteau et al., 2018).

Exclusively, CLASSIC1.0 diverges from Lawrence and Slater (2008) or Guimberteau et al. (2018), deriving the thermal conductivities of organic soils based on the analysis conducted by Côté and Konrad (2005).

Lastly, JSBACH does not explicitly represent SOM. Alternatively, it represents a moss and/or lichen layer at the surface, which possesses dynamic moisture contents and thermal properties, thus serving as a physical representation of the surface organic layer (Chadburn et al., 2017). MATSIRO6.0 adopts a similar approach by incorporating a top organic layer (Saito, 2008).

The diverse array of complexities introduced by these various approaches remains ripe for regional-scale testing to comprehensively assess their additional values.

2.4.2.4 Soil hydrology parametrization

The majority of models solve water mass transfer by using Darcy's law and soil moisture dynamics by using the Richards equations in a multi-layered representation of the soil column (Tabe 2.3, column 7). However, we assumed that GFDL-LM4 employ a "bucket" scheme, which treats the soil column as a single, well-mixed unit. By investigating different setups of soil hydrology schemes with JSBACH, Hagemann and Stacke (2015) conclude that the utilization of hydrological multi-layers model leads to a more realistic representation of soil hydrological processes.

From the models using the multi-layers approach, ISBA deviates by adopting a "mixed-form" of the Richards equations. In this approach, the time tendency is resolved in relation to volumetric water, while the hydraulic gradient is addressed through water pressure head. This alternative methodology contrasts with the moisture-based form utilized by other models, where both time tendency and hydraulic gradient are addressed in terms of volumetric water content. Decharme et al. (2011) argue that the latter approach is constrained to strictly unsaturated conditions and homogeneous media. This limitation arises due to the discontinuity of soil moisture at the interfaces of soil layers, thereby limiting its applicability within hydrological investigations.

In the context of soil hydrology, two commonly used closed-form models predict the relationship between volumetric water content and soil hydraulic conductivity, as well as matric potential (or water pressure head). The model proposed by Brooks and

Corey (1966) (BC - Appendix B - Equation B.2, B.3) offers simple analytical power functions for soil-water retention, suction and hydraulic conductivity. It is a more simplistic representation of soil's hydraulic properties than other schemes, but the required parameters can be determined from the sand, silt and clay fractions of soils, which are available in many global soil datasets (Best et al., 2011). Some models also refer to this method as the Campbell (1974) or Clapp and Hornberger (1978) model. On the other hand, the Van Genuchten (1980) (VG - Appendix B - Equation B.4, B.5) model provides more complex analytical functions. However, the specific parameters required for this formulation have not been traditionally available within soil datasets, making it difficult to use. Still, more recent datasets now include these parameters within their soil information (Best et al., 2011). The BC model is predominantly used within the atmospheric community, while hydrologists generally prefer the VG model (Decharme et al., 2011). In a site-study consisting of a fallow field in southwestern France, Decharme et al. (2011) found that BC gives the best results for simulating the soil moisture profile, while VG displays the best score at the surface. However, their results were not significantly different. Another study indicates that when appropriate parameter values are selected, both models exhibit similarities across a wide range of soil moisture levels (Dharssi et al., 2009). More research is needed to explore this question.

2.4.3 Snow physics

Various studies have aimed to compare the snow physics employed in LSMs. The Earth System Model–Snow Model Intercomparison Project (ESM-SnowMIP, Krinner et al. 2018) is an extension of LS3MIP focusing on assessing the performance of different snow models. Menard et al. (2021) conducted an innovative comparison and discussion involving 27 models from ESM-SnowMIP, excluding ELMv0 and GFDL-LM4. In this section, I emphasize essential snow properties related to soil heat transfer, which we believe have not been explored in previous research.

Table 2.5 provides a summary of the snow layering, water phase transitions, snow densification, and snow thermal conductivity parameterizations used in the models studied. Each model employs its own developed snow model, except for ELMv0, which utilizes an older version of CLM5. ISBA has been used with a more sophisticated model like CROCUS (Barrere et al., 2017), but this particular version is not present in the recent CMIP6 dataset.

2.4.3.1 Snow layering

Concerning snow layering, most of the models incorporate a multi-layer representation of snow, with the exception of CLASSIC1.0 and GFDL-LM4, which use a single-layer, simpler representation of snow.

Table 2.5 Snow Modules Characteristics in Land Surface Models

LSM	Snow model reference	Snow layers	Snow densification processes	Fresh snow density scheme depends on	Snow thermal conductivity scheme	Water phases
CLM5	Oleson et al. (2013), van Kampenhout et al. (2017)	Dynamic (max 12)	compaction, destructive and melt metamorphism	wind speed and temperature	Jordan (1991)	Ice, liquid
ISBA	Decharme et al. (2016) - CROCUS coupling exists	Static (12)	compaction, destructive and melt metamorphism	wind speed and temperature	Yen (1981)	Ice, liquid, vapour
ELMv0	Oleson et al. (2013)	Dynamic (max 5)	compactation, temperature	Jordan (1991)	Jordan (1991)	Ice, liquid
GFDL-LM4	Milly et al. (2014)	Static (1)	Fixed	fixed	Fixed	Ice
ORCHIDEE	Wang et al. (2013)	Static (3)	compaction, destructive metamorphism	temperature	Anderson (1976)	Ice, liquid, vapour
JULES	Best et al. (2011)	Dynamic (max 8?)	compaction	fixed	Yen (1981)	Ice, liquid, vapour
JSBACH	Ekici et al. (2014)	Dynamic (max 5)	Fixed	fixed	Fixed	Ice
MATSIRO6.0	Takata et al. (2003)	Dynamic (max 3)	Fixed	fixed	Fixed (0.3)	Ice, liquid
CLASSIC1.0	Verseghy (1991)	Static (1)	Fixed	fixed	Mellor (1977)	Ice, liquid

Within the multi-layer approach, some models use a dynamically evolving set of snow layers, where these layers actively interact with each other and can be added, removed, merged, or split over time, triggered by specific snow depth thresholds. Conversely, other models adopt a static number of layers, similar to how it is done in soil and atmospheric applications, with fixed positions in space or pressure levels. Research indicates that a non-boundary dynamic system works better for snowpack physics (Winstral et al., 2019).

The models' minimum number of layers is three, as seen in MATSIRO6.0 and ORCHIDEE, which has been shown to adequately resolve the thermal gradients within the snow cover, as recognized by Sun et al. (1999). On the other end of the spectrum, the most sophisticated models, such as CLM5 and ISBA, use up to twelve layers to capture snow processes in detail.

2.4.3.2 Snow densification

As expected, the most basic models, namely GFDL-LM4, JSBACH, MATSIRO6.0, and CLASSIC1.0, have fixed snow density. However, for the other models, snow density can change due to three distinct processes:

- Compaction: Snow layers compact due to overburden pressure stress, leading to sintering and mechanical creep.
- Destructive metamorphism: Under thermal equilibrium conditions, water molecules undergo sublimation and condensation on snow crystals to minimize surface free energy. As a result, newly formed snow crystals with a high surface area to mass ratio experience transformative metamorphism, where sharp-edged crystals change into cohesive aggregates with smoother, rounded, or irregular grains (Yen, 1981). This process causes the snowpack to settle and its density to increase.
- Melt metamorphism: The presence of liquid water and melt-freeze cycles cause changes in the snow crystal structure.

A fourth process called constructive metamorphism (Yen, 1981), related to temperature gradients and the formation of depth hoar, is not included in any of the snow models studied here or in most state-of-the-art snowpack models (Dutch et al., 2022).

All models without fixed snow density incorporate compaction due to overburden pressure. Notably, CLM5 takes an additional step in representing compaction by including a simple representation of compaction by drifting snow (van Kampenhout et al., 2017), which has been recognized as a potential key factor to represent specific snow density over Tundra regions (Guimberteau et al., 2018). Destructive metamorphism is included in ORCHIDEE, ISBA, and CLM5, while melt metamorphism is considered only by ISBA and CLM5.

Additionally, fresh snow density is parameterized differently in column 5 of Table 2.5 for each model. ELMv0 and ORCHIDEE use Anderson (1976) parametrization, which depends linearly on temperature. On the other hand, ISBA and CLM5 use a more complex dependence of fresh snow density on temperature and wind speed (Decharme et al., 2016; van Kampenhout et al., 2017). This increase in complexity was introduced to address excessive subsurface melt predicted in Antarctica (Lenaerts et al., 2019). However, the potential unintended consequences of applying this increased complexity to the Arctic snowpack are yet to be explored (Dutch et al., 2022). I further address this issue in Chapter 5.

2.4.3.3 Water phases

With the exception of GFDL-LM4 and JSBACH, most models in this study allow for snow to undergo melting or refreezing processes. These phase transitions between ice and liquid water significantly impact the energy distribution within the soil and exert a crucial influence on the upper snow layer (Wang et al., 2017).

To handle these phase changes, LSMs typically use temperature as the prognostic variable and employ a threshold freezing point to determine the occurrence of water phase change. However, Wang et al. (2017) point out that this approach can lead to instability issues with mass or heat conservation during numerical simulations. The sudden variation in bulk (soil and water) hydraulic and thermal properties, especially a dramatic change in soil hydraulic conductivity, at the freezing point, can cause numerical instability when solving the coupled energy and water balance equations.

In addition, ISBA and ORCHIDEE models take into account water vapor in their snow representation. They include the parametrization proposed by Sun et al. (1999), which represents the soil thermal conductivity due to vapor transfer in the snow described in the next section. Estimates provided by researchers for the heat transported by water vapor vary significantly, ranging from 10% to 40%, depending on the specific characteristics of the snowpack studied (Sturm et al., 1997), while Mellor (1977) concluded that at very low snow density values, vapor diffusion dominates the heat-transfer process.

The consideration of phase changes and the inclusion of water vapor in these models add complexity to better represent the energy distribution and improve the simulation of snow processes. However, the impact of these considerations have not yet been tested over the Arctic.

2.4.3.4 Snow thermal conductivity scheme

In this section, I briefly introduce snow thermal conductivity and explain various schemes utilized by the models listed. A more comprehensive analysis of this topic in LSMs is

presented in Chapter 5.

Snow thermal conductivity plays a pivotal role in determining the rate of heat transfer to the underlying soil, which is arguably the most crucial model parameter within snow physics (Hu et al., 2023) and one of the largest sources of uncertainty in LSM (Langer et al., 2013).

Commonly, LSMs parameterize snow thermal conductivity as a function of simulated snow density or snow temperature, K_{eff} (Yen, 1981). Multiple statistical relationships have been proposed based on experiments conducted in laboratories or on diverse snowpacks around the world. Consequently, it is reasonable to assert that a statistical relationship derived from experiments involving an alpine snowpack would significantly differ from the actual K_{eff} of a snowpack found in the Arctic tundra or on an ice sheet, and vice versa. Moreover, the temperature range observed when those samples were collected can significantly differ from Arctic permafrost regions (Sturm et al., 1997).

Within the list of models studied here, three use fixed values for snow thermal conductivity (GFDL-LM4, JSBACH, and MATISRO6.0). Among the other non-static K_{eff} schemes, four distinct K_{eff} approaches are employed:

- Jordan (1991) for CLM5 and ELMv0, is derived from a single study made of 5 observations from New Hampshire mountain areas snowpacks (Yen, 1962) and is formulated as:

$$K_{\text{eff}} = \lambda_{\text{air}} + \left(7.75 \times 10^{-5} \rho_{\text{sno}} + 1.105 \times 10^{-6} \rho_{\text{sno}}^2\right) \left(\lambda_{\text{ice}} - \lambda_{\text{air}}\right) \tag{2.12}$$

  where λ_{air}, λ_{ice} are the thermal conductivity of air $= 0.023\ Wm^{-1}K^{-1}$, and ice $= 2.29\ Wm^{-1}K^{-1}$, and ρ_{sno} is the snow density in kgm^{-3}.

- Yen (1981) for ISBA and JULES, is determined from snow density using an equation fitted to results obtained by seven studies. Snow samples were collected in St. Petersburg - Russia (13), Uppsala - Sweden (33), the French Alps (56), other locations in Russia (16), Leipzig - Germany (18), and New Hampshire - USA (5), plus an additional study from Russia I could not retrieve, for a total of minimum 141. The resulting equation is:

$$K_{\text{eff}} = \lambda_{\text{ice}} \left(\frac{\rho_{\text{sno}}}{\rho_{\text{water}}}\right)^{1.885} \tag{2.13}$$

- Anderson (1976) for ORCHIDEE, is derived by the same exact group of studies from Yen (1981) with two additional theoretical expressions from Schwerdtfeger (1963) and Woodside (1958). It is formulated as:

$$K_{\text{eff}} = 2.5 \times 10^{-6} \left(\rho_{\text{sno}}\right)^2 + 0.02 \tag{2.14}$$

- Mellor (1977) for CLASSIC1.0, is similarly derived by the same group of studies from Yen (1981), but without samples from New Hampshire, and with five additional studies. The additional snow samples were collected in Sapporo - Japan (16), Massachusetts - USA (8), and New Brunswick - Canada (5), plus three undocumented references for a total of minimum 170. It is presented in Verseghy (1991) as:

$$K_{\text{eff}} = 2.576 \times 10^{-6} \left(\rho_{\text{sno}}\right)^2 + 0.074 \tag{2.15}$$

From equations 2.13 and 2.14, the additional simple parametrization from Sun et al. (1999), considering vapor heat transfer within the snow, must be added:

$$\frac{P_0}{P_a} \times \max(0, k_1 - \frac{k_2}{T_{sno} - k_3}) \tag{2.16}$$

where P_a is the air pressure (hPa), P_0 is a reference pressure equal to 1000 hPa, and the coefficients k_1 = -0.06023 $Wm^{-1}K^{-1}$, k_2 = 2.5425 Wm^{-1}, and k_3 = 289.99 K.

Ignoring vapour heat transfer, Figure A.3 in Appendix A compares the four snow thermal conductivity schemes depending on snow density. Overall, the equations differ more at lower density values. Two groups emerge: (a) Yen (1981) and Anderson (1976) and (b) Jordan (1991) and Mellor (1997) with higher K_{eff} values for the later. Below 200 kgm^{-3} the two groups differ widely, and there is up to a factor of 2 as already discussed in Wang et al. (2013), which are close to the typical snow density values found in the Arctic (Royer et al., 2021). Equations differ greatly below 100 kgm^{-3} snow density but such values are rarely observed in Arctic snowpacks.

Ultimately, Sturm et al. (1997) have conducted a comprehensive evaluation of all the studies employed to establish the four K_{eff} aforementioned equations, shedding light on inherent issues. For instance, they emphasized that the studies underpinning the derivation of the initial three equations relied on weathered boxed snow samples. In addition, the latter three equations were derived without incorporating measurements of crucial variables such as solar radiation and utilized compressed snow samples that were not in thermal equilibrium. Moreover, their outcomes were considerably influenced by data smoothing techniques. The insufficiency of documentation and precision inherent in these methodologies are addressed in Chapter 5.

2.4.4 Vegetation representation

All models adopt a ”big-leaf” approach and have Plant Functional Types (PFT). Table 2.6 presents the variations in PFT utilized by each model for grid points designated as vegetated. One significant difference observed among the models is the implementation of a ”mosaic PFT tile,” or tile distribution. In this approach, a single grid point can represent a distribution of PFTs rather than being characterized by a single PFT. This is

Table 2.6 Vegetation Characteristics in Land Surface Models

LSM	Vegetated PFTs	PFTs linked to Arctic vegetation
CLM5	15 (tile distribution)	C_3 arctic grass, BDS boreal
ISBA	15 (tile distribution)	C_3 arctic grass, BDS boreal
ELMv0	14 (tile distribution)	C_3 arctic grass, BDS boreal
GFDL-LM4	5 (single)	None
ORCHIDEE	11 (tile distribution)	None
JULES	9 (tile distribution)	None
JSBACH	17 (tile distribution)	Tundra, BDS boreal
MATSIRO6.0	10 (tile distribution)	None
CLASSIC1.0	4 (tile distribution)	None

crucial in accurately representing vegetation heterogeneity over a region, as each grid point typically covers an area of at least 10 km^2. All studied models except GFDL-LM4 incorporate this mosaic PFT tile approach, contributing to a more realistic depiction of the diverse vegetation in the region.

Regarding the PFTs associated with arctic vegetation used by the models, only four LSMs employ the same two Arctic-based PFTs: CLM5, ISBA, ELMv0, and JSBACH. However, JSBACH uses a "tundra" PFT, which bears similarities to C_3 grass but with a reduced maximum rate of carboxylation in leaves (Chadburn et al., 2017). Curasi et al. (2023) demonstrates that the inclusion of similar PFTs enhances the ability of LSMs in simulating the carbon cycle over the Arctic region.

2.5 Conclusion and further research directions

This chapter aims to understand and consolidate fundamental model prerequisites for the representation of permafrost soils in LSMs. After examining foundational heat and water transfer principles, I examined soil physics, snow physics, and other key variables in more detail. Then, I compared LSMs in CMIP6 to reveal their differences, capabilities, and shortcomings in soil physics, soil hydrology, snow physics and vegetation representation.

First, I delved into the different soil discretization methods employed in LSMs. I observed significant variability in factors such as the number of soil layers, their distribution, and the maximum depths in the models, highlighting significant variability. Based on two key criteria, the need for high-resolution modeling within the top 3 meters of the soil column (Lawrence et al., 2019) and the recommended minimum depth (Alexeev et al., 2007; Hu et al., 2023), I categorized the nine models studied into four distinct groups.

This study then highlighted the profound importance of soil thermal conductivity

schemes in accurately representing soil temperature and permafrost dynamics by examining various schemes from selected LSMs and observing their implications on soil temperature representation. Through external comparative analysis, I identified the Wang et al. (2016a) scheme used in ORCHIDEE yielded the most promising results. Furthermore, I highlighted the value of considering the influence of soil organic matter. Continued research in this field holds the potential to further enhance our understanding of soil thermal conductivity representation in LSMs.

Next, I observed the diversity in approaches to representing soil hydrology, ranging from multi-layered models to "bucket" schemes. While the majority of models employ the Richards equations to simulate water mass transfer and soil moisture dynamics, the choice between two methods to derive the relationship between volumetric water content and soil hydraulic conductivity lacks general consensus.

I examined various snow physics parameters including snow layering, densification, water phases, and thermal conductivity. The models demonstrated differing approaches to representing these complex processes, and I highlighted the importance of phase transitions and water vapor in refining energy distribution within the soil. Additionally, the discussion on snow thermal conductivity shed light on the challenges associated with deriving accurate equations, given the wide range of snowpack conditions across different regions. A comprehensive evaluation of existing studies by Sturm et al. (1997), further discussed in Chapter 5, highlights limitations in current equations, emphasizing the need for precise documentation and data.

Plant Functional Types (PFTs) is the common way to represent vegetation, where mosaic PFT tiles emerged as a crucial approach to accurately portray vegetation heterogeneity. The significance of aligning PFT choices with Arctic-based vegetation was highlighted, particularly in terms of enhancing the carbon cycle simulations over the Arctic.

It is crucial to note that even for models encompassing the same processes, results can significantly diverge. Such disparities arise due to variations in the finer points of implementation, differing parameterizations, or resolutions-topics not examined here. Overall, each of these LSMs represent a unique approach to simulating soil processes and their interactions with the climate system. Choosing the most appropriate model for a particular application depends on factors such as the spatial and temporal scale of the simulation, the specific research question, and the available computational resources. Low complexity models like GFDL-LM4 are likely to be more computationally efficient. All of these elements are interconnected, further highlighting the importance of synthetic evaluation of different yet related model aspects in improving our understanding of the system.

In conclusion, the interactions between soil, snow, and vegetation within Arctic permafrost regions are highly complex and sensitive to local conditions. The advance-

ments in LSMs discussed here underscore the continuous efforts towards refining our understanding and modeling capabilities. Further research and studies are essential to validate and improve these models, facilitating their applicability in predicting climate-related impacts and informing policy decisions in the face of changing environmental conditions.

In the subsequent chapters, I delve into case studies involving the application of CLM5 over the permafrost region, aiming to validate its performance and enhance our insights into its potential strengths and limitations.

CHAPTER 3

Evaluation of CLM5 against in-situ and grid-based observations

3.1 Introduction

Understanding the complex dynamics of land surface processes is crucial for unraveling the intricate interplay between the Earth's surface and the atmosphere. In permafrost studies, the significance of accurate LSMs cannot be overstated. These models allow us to represent the balance and exchanges of energy, heat, water, and carbon fluxes operating in the soil.

The integration of high-resolution Regional Climate Models (RCMs) further enhances our ability to dissect land surface-atmosphere interactions with precision. As highlighted by Sushama et al. (2007), the fine spatial resolution of RCMs enable a comprehensive examination of localized climate phenomena. Unlike their global counterparts, RCMs are bounded by lateral climatic data, which mitigates the risk of drifting into divergent climatic states. This intrinsic constraint ensures that any deviations from observed conditions are inherently linked to variations in model physics rather than an evolving climate "background". By utilizing reanalysis data for driving these models, we can align the simulated climate as closely as possible with observed conditions, setting the stage for rigorous model evaluation. Burke et al. (2020) evaluated that the restricted resolution of ESMs may contribute to the models' difficulties in accurately simulating summer thaw depths. RCMs offer a useful solution to this challenge.

This chapter evaluates the ability of the state-of-the-art CLM5, to accurately replicate seasonal and annual ground temperature profiles, which are pivotal benchmarks for representing the underlying permafrost dynamics. After a comprehensive description of CLM5, I overview the different validation data and validation procedures used in this study. Subsequently, I present the results and a discussion of this evaluation before drawing conclusions in the final section.

Through this holistic investigation, I aim to unearth the strengths and limitations of CLM5 in capturing the complex land surface processes that underpin Arctic permafrost regions. By subjecting this model to rigorous scrutiny, I hope to enrich our understanding of its performance and contribute to the ongoing refinement of land surface modeling approaches for permafrost studies.

3.2 Community Land Model (CLM5) description

This study uses the Community Earth System Model (CTSM) (`https://github.com/ESCOMP/CTSM`). CTSM is released by the National Center for Atmosphere Research (NCAR) and is the default land component of the community-developed CESM2, part of the CMIP6 experiment. CTSM is a three-dimensional process-based model of the land surface and the terrestrial biosphere that calculates water, energy, and carbon fluxes between the surface and different soil layers. The land component of CTSM, known as the Community Land Model (CLM5), is prominently featured in this study. In the subsequent sections, I refer to CTSM as CLM5 to simplify the discussion.

A comprehensive model description and global evaluation can be found in Lawrence et al. (2019) and in the technical description in Lawrence et al. (2018).

Several updates and parameterizations have been introduced in CLM5 compared to CLM4.5, enhancing its performance in simulating various processes. Among others, these updates include improvements in snow density, hydraulic redistribution, and nitrogen cycling.

The model stratigraphy includes a complex 25 soil layers representation, with a bottom layer at around 50 meters. Compared to the previous iteration of CLM5, the soil layer resolution of the default model is increased, especially in the first 3 meters, to more explicitly represent the active layer thickness (ALT) in permafrost areas, a key element in this study. Soil texture (fraction of sand, clay and silt) and organic matter density (OMD) are derived from Bonan et al. (2002a) and Hugelius et al. (2013), respectively. The fractions vary with depth for the first 10 layers but are constant for the following 15 layers.

The representation of spatial land surface heterogeneity in CLM involves a nested subgrid hierarchy. In this hierarchy, grid cells are composed of multiple land units, columns, and patches, as illustrated in Figure A.4, Appendix A. Notably, each grid cell can have a different number of land units, with each land unit accommodating a distinct number of columns. Furthermore, multiple patches can exist in the vegetated land units, each representing a specific Plant Functional Type (PFT).

In my configuration, vegetation heterogeneity is characterized by using a fraction of 15 + 2 different PFTs in each grid cell, including 1 PFT for crops, and 1 PFT for bare soil (absence of vegetation), sharing the same soil column. Radiation, temperature, and interception are calculated for each PFT separately. All PFTs share common equations but with different parameters, except for leaf phenology. This study uses the satellite phenology mode where stem and leaf area indices are derived from monthly climatologies.

CLM5 calculates radiative fluxes for shortwave radiation as an interaction of the optical properties of the canopy and surface with the descending shortwave radiation.

Shortwave radiation in the visible and near-infrared range is considered separately for diffuse and direct light. The net shortwave radiation is therefore dependent on a number of optical properties of the vegetation and snow, including the stem and leaf area index, ice grain size distribution of the snow pack, and aerosol deposits in the snow.

The heat transfer equation 2.5, presented in Chapter 2, is solved numerically to calculate the soil, snow, and surface water temperatures for the 25-layer soil column and zero heat flux at the bottom of the soil column. The numerical solution is described in https://escomp.github.io/ctsm-docs/versions/master/html/tech_note/Soil_Snow_Temperatures/CLM50_Tech_Note_Soil_Snow_Temperatures.html#numerical-solution. At each time iteration, the snow/soil temperatures are evaluated to determine if phase change takes place. Chapters 2 and 4 describe the equations in detail.

Hydrology calculations, including the Richards equation 2.9 presented in Chapter 2, occur in the upper 20 soil layers, while the 5 bedrock layers are considered to be impermeable for water. The numerical solution is described in https://escomp.github.io/ctsm-docs/versions/master/html/tech_note/Hydrology/CLM50_Tech_Note_Hydrology.html#numerical-solution. Vertical soil moisture transport in the model is driven by infiltration, surface and subsurface runoff, gradient diffusion, gravity, canopy transpiration through root extraction, and interactions with groundwater, respecting the conservation of mass.

van Kampenhout et al. (2017) recently substantially modified the snow module. It includes a calculation of snow compaction, fraction of snow covered, ice and water content driven by mass conservation, and particle deposition (of black, organic carbon and mineral dust). The vertical discretization includes a maximum of twelve snow layers in an Eulerian method, in which each layer has a prescribed maximum thickness. The model assumes a maximum snow depth (or snow cap) of 10 meters snow water equivalent (SWE) to avoid runaway snow. Any mass exceeding the snow cap (whether rain or snow) is channelled to the river component. The snow module is further discussed in Chapter 5.

Figure 3.1 shows a schematic illustration of processes simulated by CLM5. Overall, CLM5 is characterized by comprehensive and well-developed hydrothermal processes and has been extensively used to study permafrost (Cai et al., 2019; Lawrence et al., 2019; Birch et al., 2020; Dutch et al., 2022; Cheng et al., 2023), making it an effective tool for this study.

3.2.1 Model set-up

The model version used throughout this study is ctsm5.1.dev086, forked version https://github.com/AdrienDams/CTSM/.

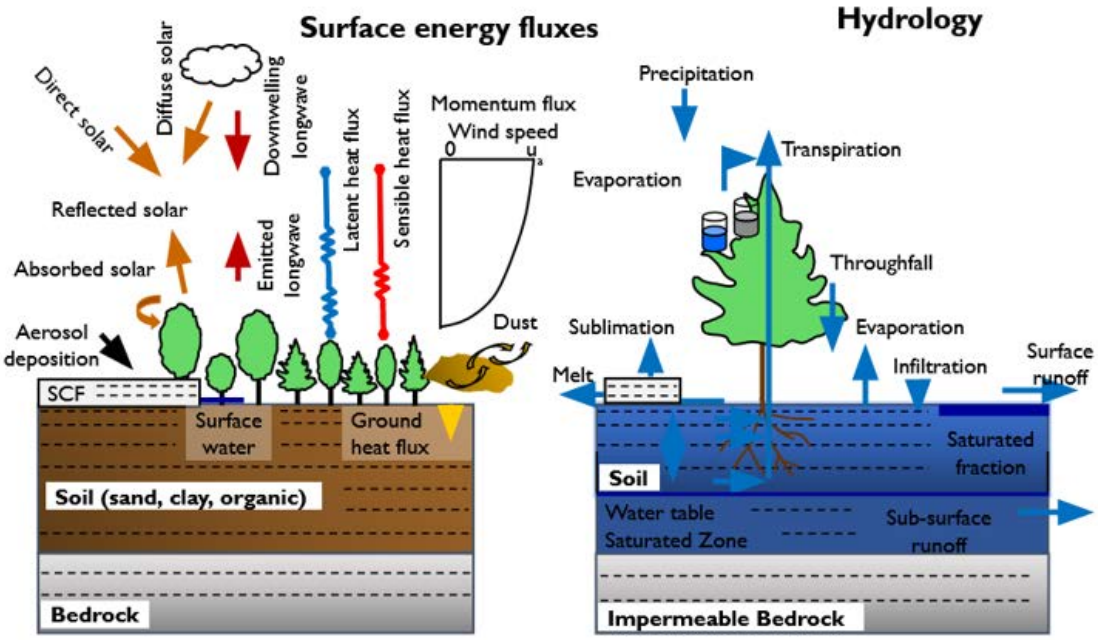


Fig. 3.1 Schematic representation of primary processes in CLM5. SCF = snow cover fraction. Note that not all soil levels are shown. Not all processes are depicted. Optional features that are not active in default configurations are italicized (adapted from Lawrence et al. 2019).

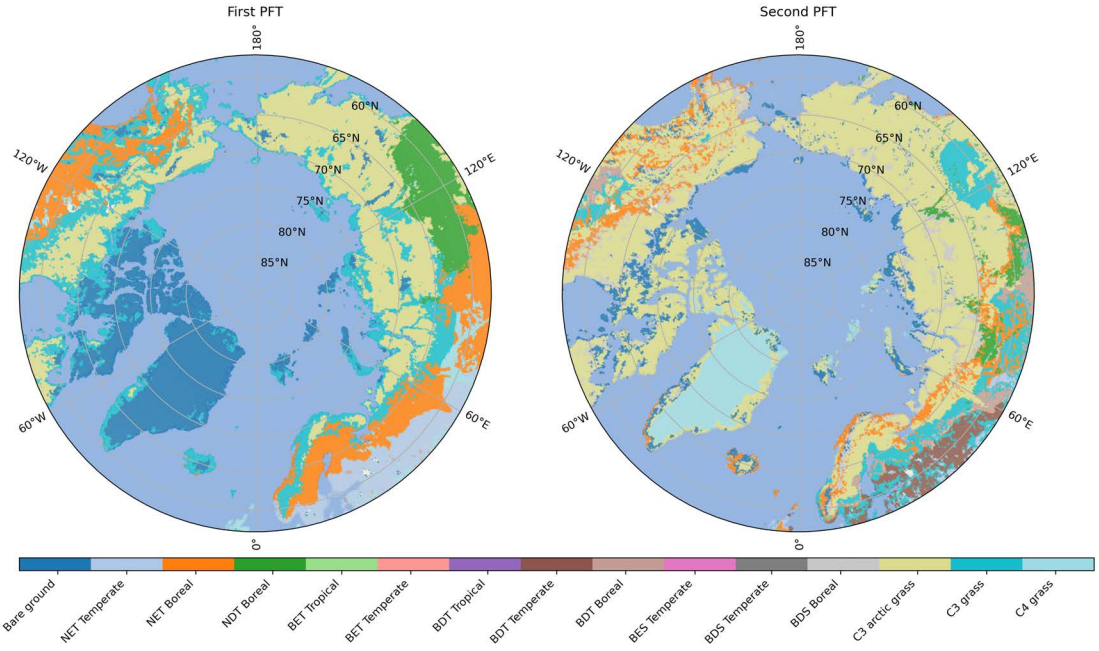


Fig. 3.2 Domain used in this study. The common coordinate projection used is a north polar stereographic with a central longitude of 0°and central latitude of 90°N. The colours represent the most dominant (left map) and second most dominant (right map) PFT for each grid point. NET - Needleleaf evergreen tree, NDT - Needleleaf deciduous tree, BET - Broadleaf evergreen tree, BDT - Broadleaf deciduous tree, BES - Broadleaf evergreen shrub, BDS - Broadleaf deciduous shrub.

The domain is specially developed for this case study. It covers the entire north pole from the latitude 57 N. The surface is represented by a non-regular icosahedral grid developed using the ICON tool (`https://webservice.dwd.de/cgi-bin/spp1167/webservice.cgi`). This type of grid was used to compare the results of this study to future works using an ICON-CLM5 coupling infrastructure. To our knowledge, this is the second time CLM5 has run on such domain, following a prior study by Birch et al. (2020), who used a similar Pan-Arctic domain but at a coarser resolution. The domain consists of 204086 grid points with a triangular resolution which varies approximately between 116.25 and 179.37 km^2, giving a rectangular resolution of around 12 km^2. Figure 3.2 represents the domain used in this study with the two most prominent PFT for each grid point.

The default meteorological forcings data (CRU/GSWP3) are replaced by the finer resolution ERA5 forcings from 1980 to 2021. ERA5 description can be found in Hersbach et al. (2020). The forcing is at an hourly timestep and on a 0.25°(around 14 km^2) latitude-longitude grid. To our knowledge, this is the second time that CLM5 was used with large scale resolution forcings, after Cheng et al. (2023). While this increase in resolution should represent a substantial improvement over previous global reanalysis methods used (Albergel et al., 2018), it also introduces additional uncertainty, as the model was not parameterized with these settings as its default configuration.

Supplementing the default model configuration, several adjustments were effectuated. To save computation time, this study uses the satellite phenology (SP) set-up, meaning it does not include complex carbon cycle interactions, and the land-ice and river routing models (MOSART) were deliberately deactivated. Furthermore, recalibrations to the snow initialization protocol were made, with the snow water equivalent (SWE) restricted to 0.8 meters, deviating from the default 10 meters. This refinement aims to avert the unrealistic distortion that was engendered by the former ceiling of 10 meters, leading to snow heights upwards of 20 meters in non-glaciated regions. Addressing this significant bias requires further study.

To start the run in an equilibrium state, a spin-up of 30 years using ERA5 reanalysis (looping from 1980 to 1989) was used before running the model from 1980 to 2021 (42 years). Results from the spin-up can be found in Appendix A, Figure A.5.

3.3 Validation data

I used three datasets to evaluate the Arctic region, including two different in-situ datasets and one derived from remote sensing products, all of which offer complementary perspectives.

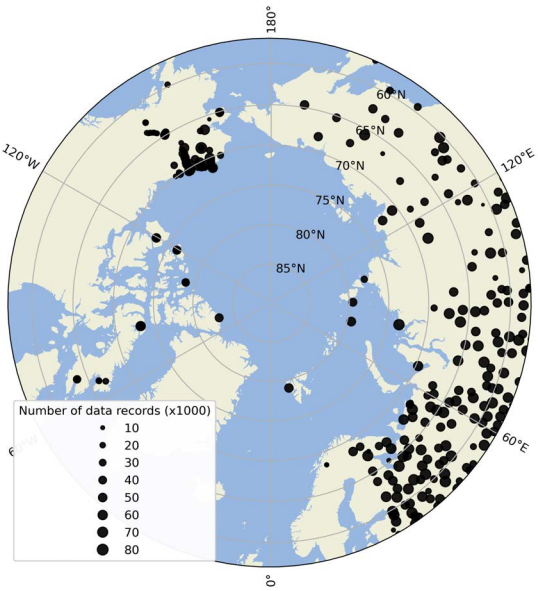


Fig. 3.3 Location of the 295 borehole stations used. The size of each point represents the number of data records per station over the whole period and for all depths.

3.3.1 In-situ ground temperature stations data

In the western Arctic region, the availability of daily long-term soil temperature records (spanning a minimum of 10 years) is notably limited. The inception of continuous measurements only commenced in the late 1990s, and even these measurements rarely extend beyond a depth of 1 meter (Matthes et al., 2017).

I expanded upon the dataset used by Matthes et al. (2017) to comprise a distinctive and diverse network encompassing 295 borehole stations across the entire Arctic pole, represented in Figure 3.3. This comprehensive dataset is denoted as "295GT" and encapsulates a temporal span of 42 years, from 1980 to 2021. I use monthly averages to expedite the analysis, preventing an an investigation into daily variation. Data were obtained from various sources: the Permafrost Laboratory website (`https://permafrost.gi.alaska.edu`), the GTN-P database (`http://gtnpdatabase.org`), the Nordicana D website (`https://nordicana.cen.ulaval.ca/`), and the Roshydromet network (`http://aisori-m.meteo.ru/`).

Ground temperatures have been recorded across 278 distinct depth levels, ranging from 0.01 m to 60 m. To align with the soil ceiling defined by CLM5, I confined my analysis to a depth of 50 m.

3.3.2 Circumpolar Active Layer Monitoring Network (CALM)

In addition to the stations measuring ground temperature, I also used the Circumpolar Active Layer Monitoring Network (CALM), as described in Shiklomanov et al. (2012). CALM is a large-scale effort to construct a station network where the seasonal depth

of thaw is measured over 124 CALM stations operated by researchers from Canada, Denmark/Greenland, Norway/Svalbard, Russia, and the United States/Alaska.

3.3.3 ESA Climate Change Initiative

In parallel to in-situ based data, I also used grid-based products from the European Space Agency (ESA) Climate Change Initiative (CCI). Throughout this study, I reference this dataset as "ESA-CCI." A detailed account of the data and its validation procedures can be found in Heim et al. (2021).

The ESA-CCI initiative constitutes a global monitoring program meticulously designed to deliver Earth Observation Essential Climate Variables (ECVs) products over extended temporal spans.

ESA-CCI products encompass ECVs with a very high pixel resolution of 1 km^2 and include mean annual ground temperature (MAGT) at distinct ground depths of 1, 2, 5, and 10 meters, Active Layer Thickness (ALT), and permafrost fraction (PFR). The geographical extent of these products spans the Northern hemisphere above 30°N within an Arctic stereographic circumpolar projection. The temporal coverage for MAGT, ALT, and PFR time series ranges from 1997 to 2019 at an annual resolution.

Considering the evaluation conducted by Heim et al. (2021), the ESA-CCI dataset exhibits a median bias of -1.12°C for MAGT. Disparities are concentrated along the southern boundary of the permafrost extent in the Eurasian region. With regard to ALT, residuals exceeding 1 meter are prominent within warmer permafrost zones situated in forested areas encompassing Alaska, Canada, and Central Siberia. Lastly, PFR demonstrates a high agreement with observed values.

3.4 Validation procedures and algorithms

This study introduces a new validation procedure. The semi-automated evaluation includes different techniques of comparisons with in-situ based and grid-based data. All the algorithms can be found in `https://github.com/AdrienDams/cegio` and are described below.

My validation methods aim to comprehensively integrate and analyze (1) temporal scale variations, (2) spatial distributions, and (3) depth variations, in the context of permafrost assessments. Generally, previous studies tend to prioritize one of these aspects at the expense of the others. In contrast, my approach aims to emphasize all three dimensions concurrently. The aim is to maintain the depth spread that I find essential to evaluate the capabilities of a land model.

3.4.1 295GT

To compare the model outputs with the 295GT dataset, I compare each station to the nearest grid point and perform a linear interpolation using the two closest CLM5 depth level options.

3.4.2 CALM

ALT is computed from the model outputs by deriving a curve of maximum annual ground temperature (MaxAGT) depending on depth, using a spline to fit the available data points. I find the depth at which this curve crosses from unfrozen to frozen soils and calculate ALT as the depth corresponding to 0°C. Thalik formations were disregarded. Figure A.2 in Appdenix A shows a schematic representation of how I calculated ALT.

3.4.3 ESA-CCI

To compare the results to ESA-CCI products, I aggregated ESA-CCI products to the domain grid using a conservative second-order regridding equation described in Jones (1999).

The equations to determine the PFR at each grid point within CLM5 is determined by the following equation:

$$\text{PFR} = \begin{cases} 1, & \frac{1}{M}\sum_{y=1}^{M} \min_{z=1,N} \max_{t(y)=1,2Y} T_i(z, t(y)) < 273.15\text{K} \\ 0, & \frac{1}{M}\sum_{y=1}^{M} \min_{z=1,N} \max_{t(y)=1,2Y} T_i(z, t(y)) \geq 273.15\text{K} \end{cases} \tag{3.1}$$

where M is the number of years minus one (41 in our setup), z is the index for the depth, N is the number of depths, t is the index for the days in the year y and the next year, Y is the number of days in a year, and $T_i(z, (t(y))$ is the temperature depending on the day, depth and grid cell.

In simple terms, I first calculated the maximum temperature over a two-year period for each grid cell and each layer. Then, I calculated the vertical minimum to see if there is one layer that is continually frozen over these two years. From this, I obtained a temperature data grid for each year, which I then averaged over the period spanning 1997 to 2019 to fit ESA-CCI products period. Subsequently, I classified grid points into two categories: those with temperatures below 0°C were designated as permafrost, while those with temperatures above 0°C were classified as non-permafrost. It is worth noting that this method provides a binary definition of permafrost, in contrast to other datasets like ESA-CCI, which offer a quantitative representation of permafrost ranging from 0 to 100% resulting from to their ensemble-members experiments.

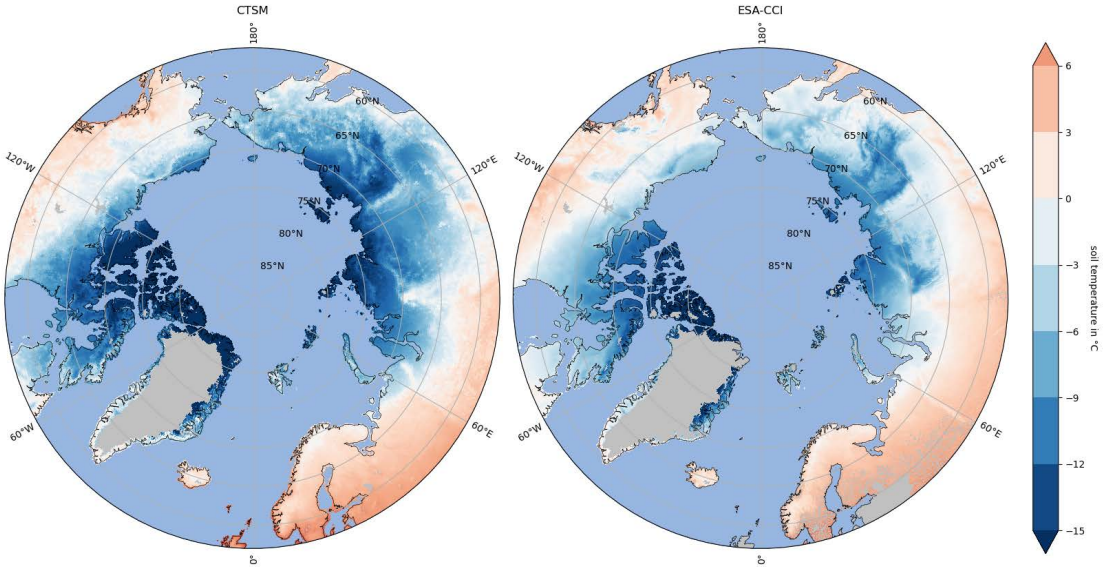


Fig. 3.4 Period (1997 to 2019) MAGT at 1m depth in°C for CLM5 (left) and ESA-CCI (right). Darker blue indicates that colder temperature. ESA-CCI data are aggregated on the CLM5 grid using a conservative projection method.

To reconcile this difference, I adopted three classes for the ESA-CCI data: continuous if greater than 90%, discontinuous if between 50% and 90%, and free if less than 50%. Six distinct masks were then created to facilitate the comparison between the two CLM5 model permafrost classifications and three ESA-CCI product permafrost classifications (see Fig. 3.11).

The determination of ALT at each grid point within CLM5 is calculated in a similar way as described in the section above. The resulting ALT data for both the model and ESA-CCI were subsequently period-averaged from 1997 to 2019.

To obtain the final maps, I made a simple grid difference between the CLM simulation and ESA-CCI for (a) MAGT, (b) PFR and (c) ALT period-averaged products. In addition, I calculated the Mean Absolute Deviation (MAD - Equation B.7 in Appendix B) and Root Mean Square Error (RMSE - Equation B.6 in Appendix B) for MAGT and ALT. In my case, predicted values are results from the model and observed values are ESA-CCI products. A lower MAD indicates that the model's predictions are closer to the actual values, while a lower RMSE indicates that the predictions are closer to the actual values on average.

3.5 Results

3.5.1 Soil temperature

Fig. 3.4 and Fig. 3.5 display the mean annual ground temperature (MAGT) at 1 meter between CLM5 and ESA-CCI. I deliberately chose to show values at 1 meter of depth exclusively because my results show that the spatial variation is much stronger than the

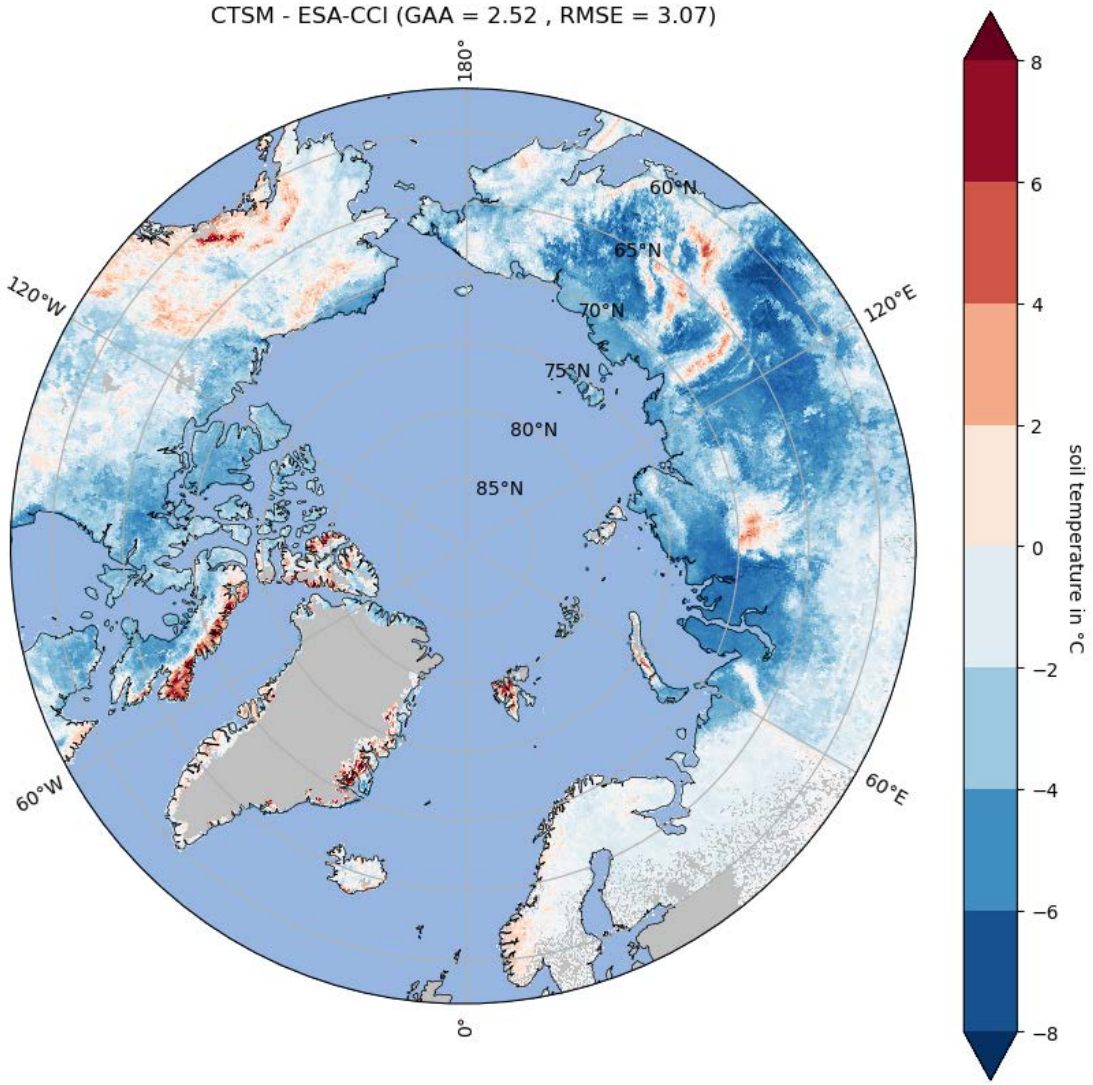


Fig. 3.5 Period (1997 to 2019) MAGT at 1m depth difference between CLM5 and ESA-CCI in°C. Darker blue colour represent CLM5 soil temperature is colder than ESA-CCI. ESA-CCI data are aggregated on the CLM5 grid using a conservative projection method.

depth variation, making it more relevant to this analysis.

The climatology pattern in Fig. 3.4 is similar to ESA-CCI, but their difference in Fig. 3.5 is very significant. In general, the model shows a much colder domain, especially over Siberia where temperature differences are up to -8°C. Over the Alaskan region, the model is much closer to ESA-CCI, with some areas still showing a difference around -2°C.

Some positive bias is observed over mountainous regions, such as the north coast of Baffin Island. However, this is likely attributed to the difference in resolution of the two products. ESA-CCI derives its forcings from daily ERA5 reanalysis data and are downscaled using the 1 km Global Multi-resolution Terrain Elevation Data (Westermann et al., 2020). Using this process, ESA-CCI uses atmospheric forcings on a much finer resolution (1 km^2) than the atmospheric forcings utilized in this study (12 km^2). This difference in resolution allows ESA-CCI to more accurately represent cold high-elevation areas, explaining the observed warm bias.

Fig. 3.6 compares soil temperature at 278 depth levels between CLM5 (simulation on y-axis) and 295 stations (observation in x-axis). I used the 295GT (in-situ observations database) because it provides monthly averages instead of period annual averages, creating a better picture of the model's accuracy in representing soil temperature. Additionally, this offers an effective tool to analyze seasonality of permafrost.

The R^2 coefficients are highly dependent on the region, and there is a linear gradient from high coefficients in the west to low coefficients in the east. Alaska has the highest coefficient of 0.77, while Eastern Russia has a low coefficient of 0.51. This indicates that the model is better able to represent soil temperature in some regions than others.

Regarding the data points on the scatter plot, one can observe that they are spread around both 0-degree lines, forming what I refer to as the "zero-curtain cross anomaly". This phenomenon can be attributed to the mismatch in timing between thawing and freezing during phase transition. For instance, when observations cross the 0°C threshold in a positive direction, two scenarios may unfold: (1) the model is in a thawing phase, and its predicted values closely match the observations, or (2) the model has not yet initiated thawing, and its predicted values are colder than the observed ones. These scenarios can occur in both directions and are a consequence of the model's simplified representation of latent heat energy.

Looking at specific regions does not reveal a clear trend in Alaska in terms of depth differences (Figure 3.7a). However, when we examine the correlations between months and soil temperature, the coefficients are much lower (Figure 3.7b) than the year-averaged coefficient. Specifically, for every month, the correlation loses 0.1 points on average, with a more pronounced decrease of 0.2 points in winter.

We can observe a different trend in Eastern Siberia. Unlike Alaska, there is a clear decrease in soil temperature correlation with increasing depth (Fig. 3.8). Interestingly,

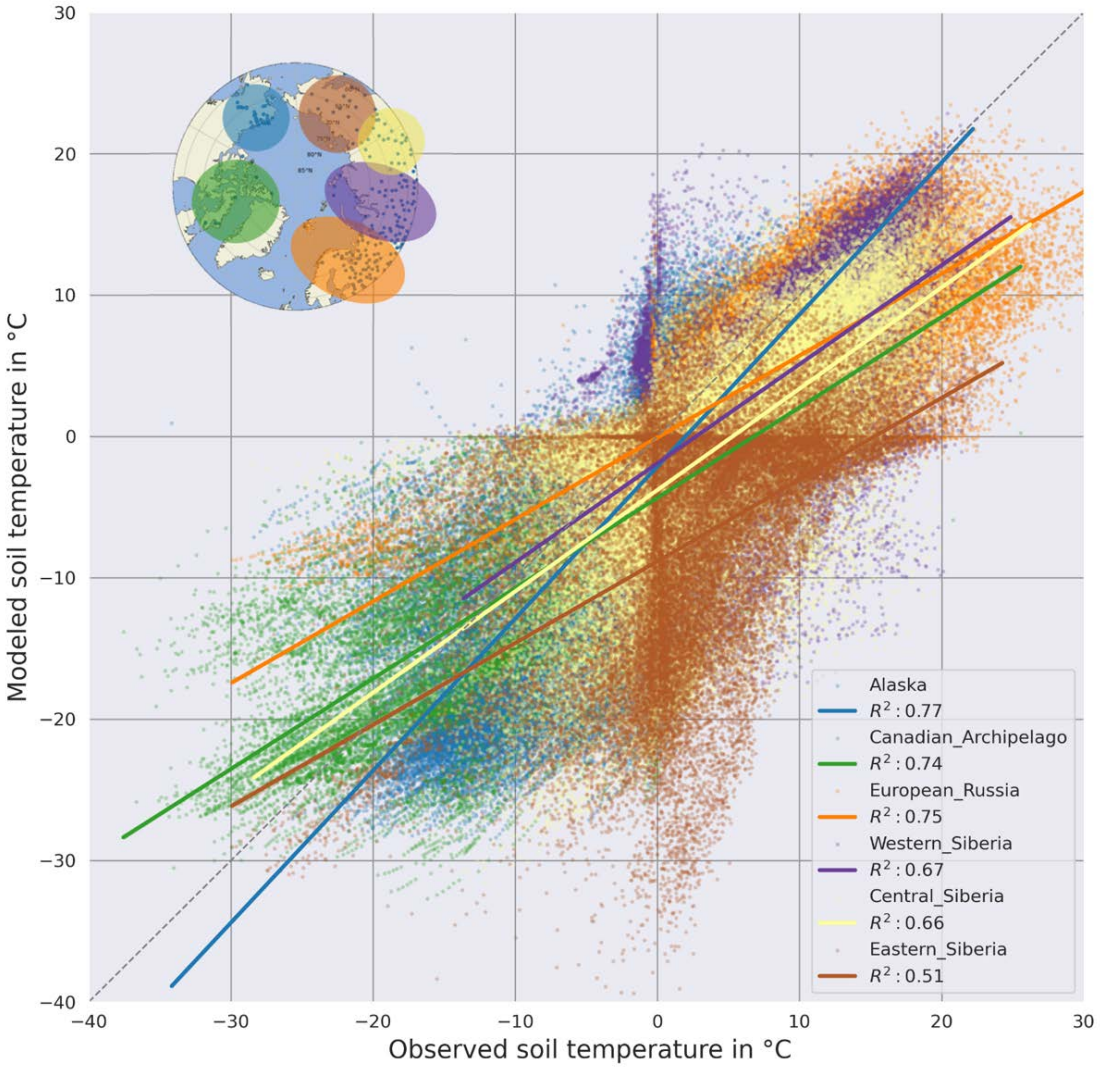
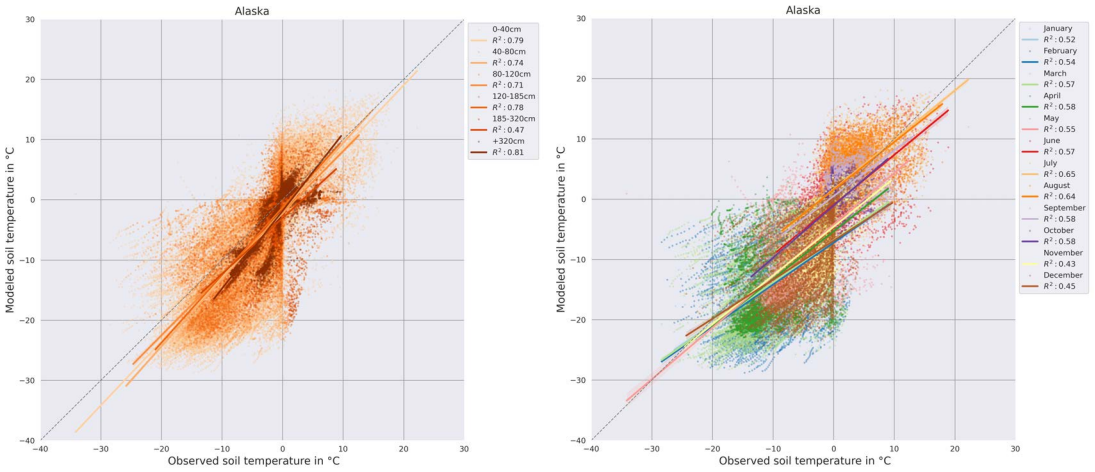
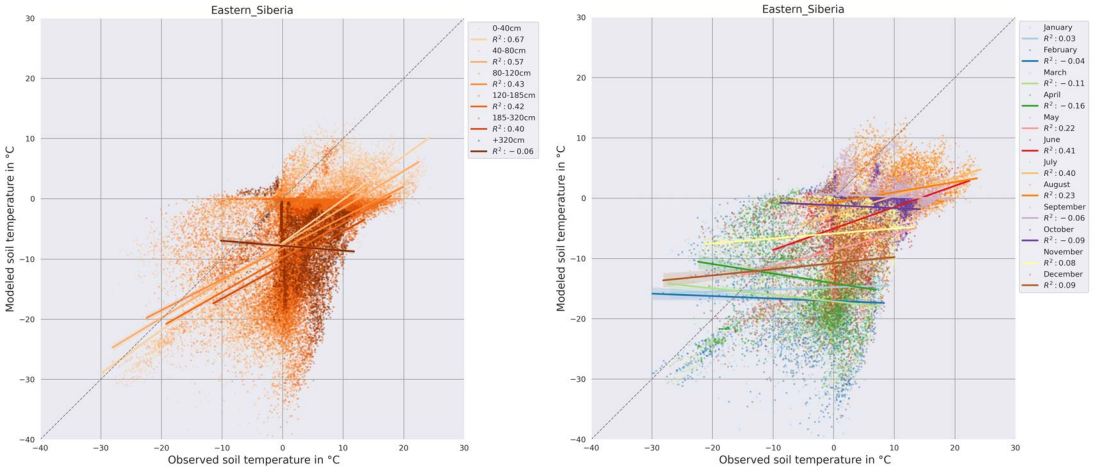


Fig. 3.6 Soil temperature at 278 depth levels comparison between CLM5 (simulation on y-axis) and 295GT (observation in x-axis) in°C. Monthly averages between 1980 to 2020. Each colour represents the region location (see legend). Lines represent an attempt at correlation between the observed and simulated values and R coefficients are shown in the legend for each region. The map shows an approximation of the regionalisation.



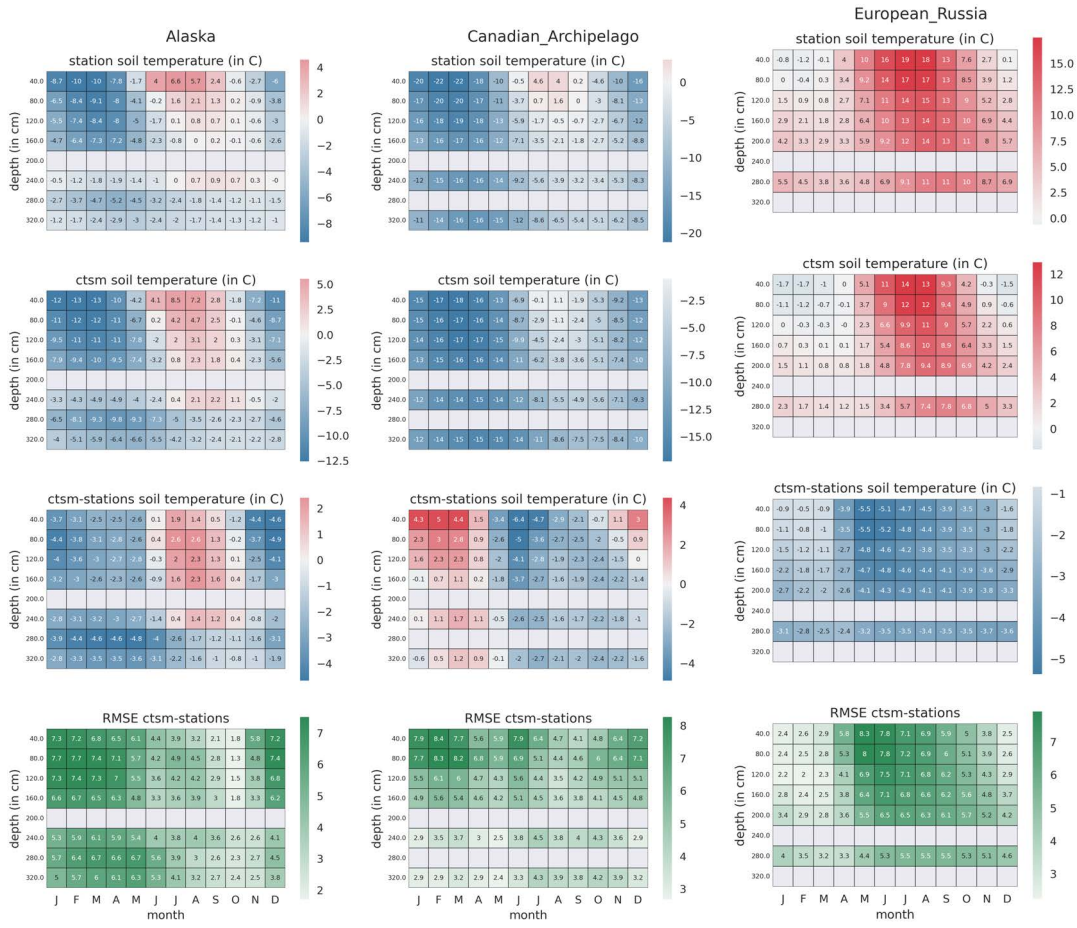
(a) Soil temperature at different depths **(b)** Soil temperature for different months

Fig. 3.7 Comparison of soil temperature between CLM5 and observations in Alaska. Figure 3.7a shows the scatter plot of soil temperature at different depths, while Figure 3.7b shows the scatter plot of soil temperature for different months. The different colours represent the depth bins and the twelve months, respectively. The lines represent the attempted correlation between the CLM5 simulations and observations, and the R coefficients for each panel are shown in the legend.



(a) Soil temperature at different depths **(b)** Soil temperature for different months

Fig. 3.8 Comparison of soil temperature between CLM5 and observations in Eastern Siberia. Figure 3.8a shows the scatter plot of soil temperature at different depths, while Figure 3.8b shows the scatter plot of soil temperature for different months. The different colours represent the depth bins and the twelve months, respectively. The lines represent the attempted correlation between the CLM5 simulations and observations, and the R coefficients for each panel are shown in the legend.



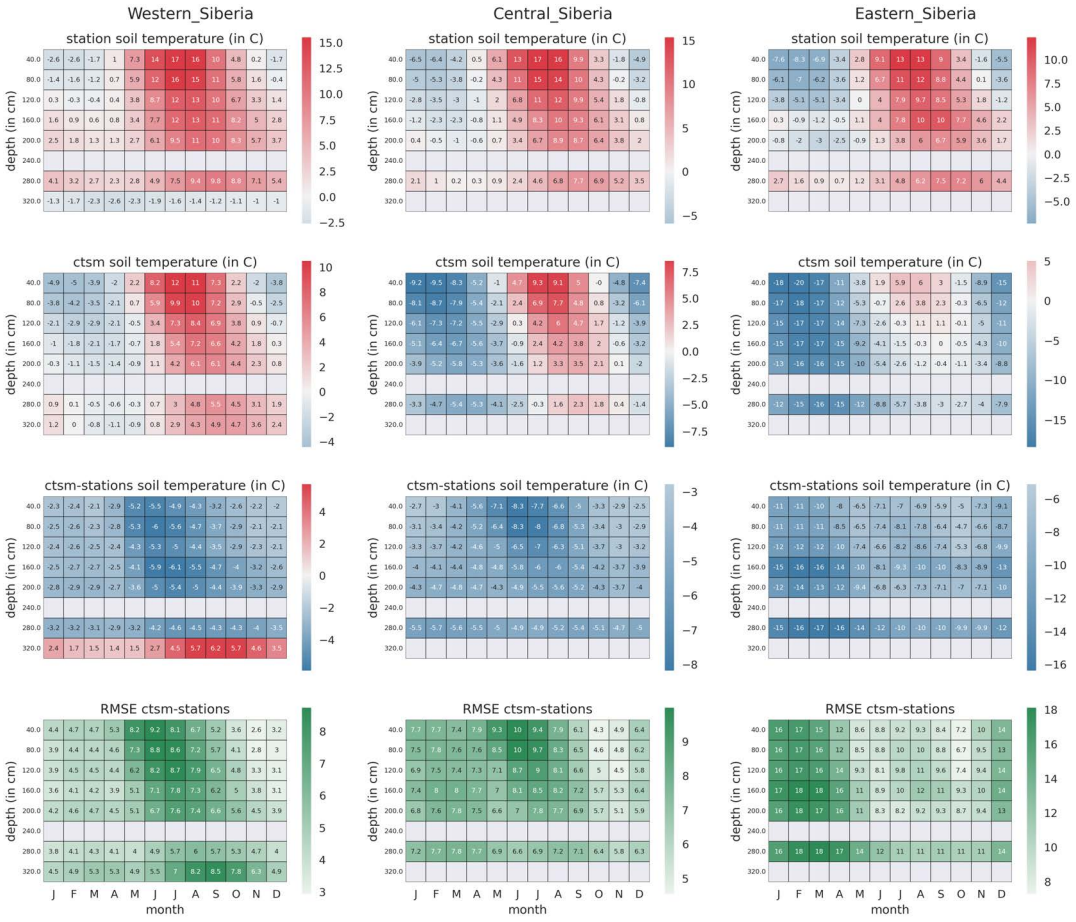
(a) Alaska **(b)** Canada **(c)** European Siberia

Fig. 3.9 Heatmaps of the period average (1980-2021) soil temperature for CLM5 and 295GT. Months are on the x-axis and the depth bins are on the y-axis. The first, second, third and fourth rows show the CLM5 averages, 295GT averages, their difference and the RMSE, respectively.

below a depth of 3.20m, the correlation coefficient drops from 0.4 to 0, indicating that the model is no longer able to accurately match the observations. Furthermore, when looking at the correlation by month, there is only a significant correlation during the summer, and no correlation between the observations and the model during the rest of the year.

The next figures, Fig. 3.9 and 3.10, present heatmaps of the period average soil temperature for CLM5 and 295GT for six regions: Alaska, Canada, Eastern Russia, Northern Europe, Western Russia, and the Tibetan Plateau.

The analysis reveals that the model is generally colder than observations across all months, depths, and regions. However, there are some specific discrepancies worth highlighting. Over Canada, the model shows winter temperatures that are warmer than observations, especially at the surface. This discrepancy is likely due to the model's poor representation of snow depth over certain islands in the Arctic, where many Cana-



(a) Western Siberia **(b)** Central Russia **(c)** Eastern Siberia

Fig. 3.10 Heatmaps of the period average (1980-2021) soil temperature for CLM5 and 295GT. Months are on the x-axis and the depth bins are on the y-axis. The first, second, third and fourth rows show the CLM5 averages, 295GT averages, their difference and the RMSE, respectively.

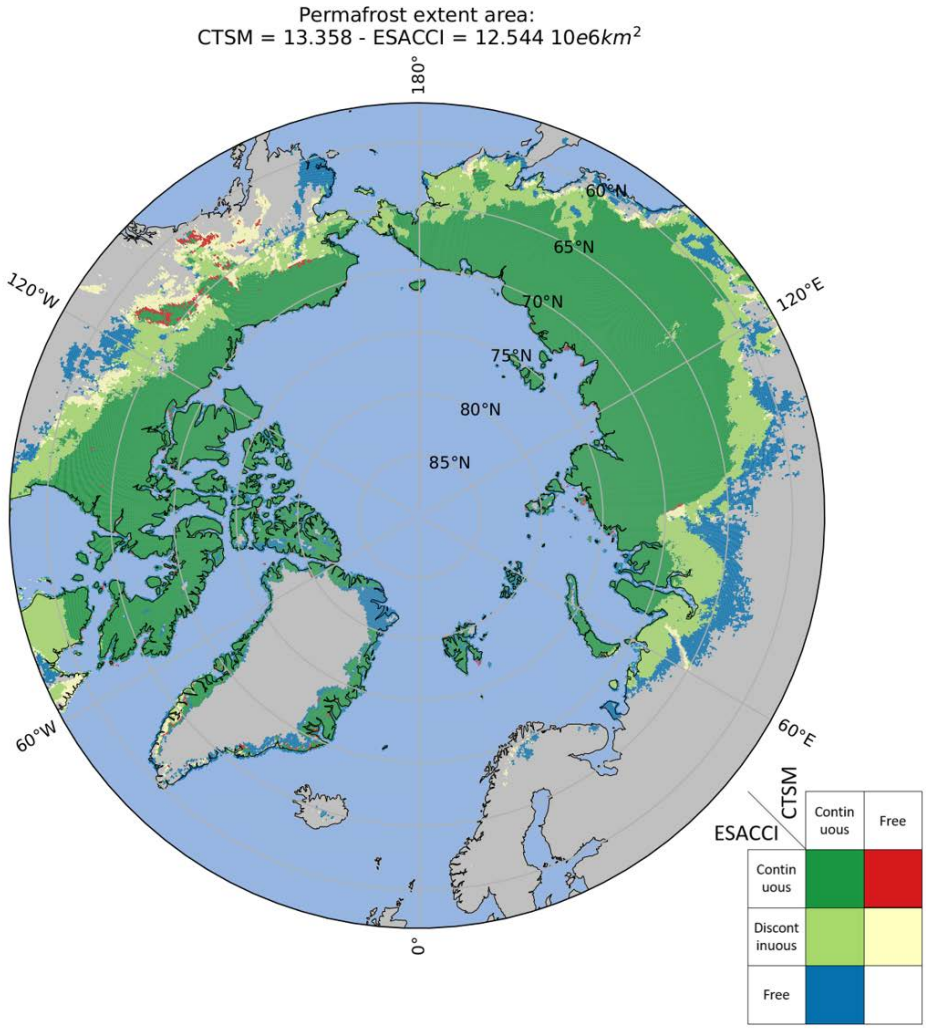


Fig. 3.11 Permafrost extent area mask difference between CLM5 and ESA-CCI. Refer to the legend for colour meanings. ESA-CCI data are aggregated on the CLM5 grid using a conservative projection method.

dian observation stations are located. Over Alaska, the model exhibits warmer temperatures than observed in summer, although this disparity is less prominent than the one observed in Canada.

Appendix C provides a supplementary analysis that complements the regional study of soil temperature presented in the preceding chapter. This appendix offers a detailed comparison of temperature data from a selection of 10 stations, each carefully chosen to enhance our understanding of specific aspects related to soil temperature variations.

3.5.2 Permafrost extent

There is strong agreement between the CLM5 and ESA-CCI permafrost extents, with 93% of the two datasets overlapping, including the discontinuous Arctic permafrost regions (Fig. 3.11), with a total extent of 13.358 and 12.544 $\times 10^6$ km^2 for CLM5 and ESA-CCI, respectively. However, there is a slight overestimation of permafrost extent

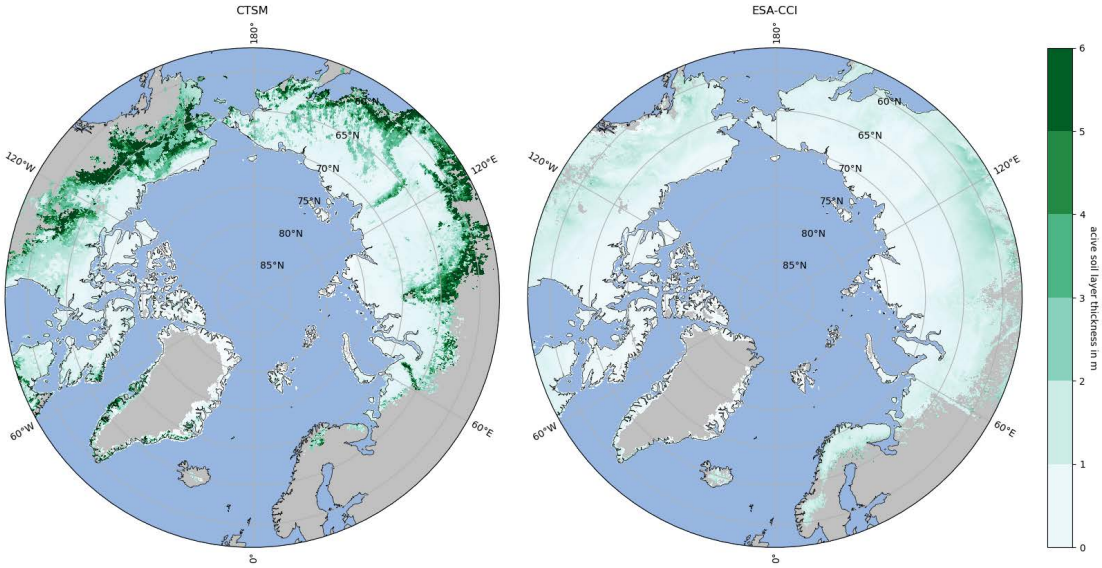


Fig. 3.12 Period (1997 to 2019) annual averages of Active Layer Thickness (ALT) in meters for CLM5 (left) and ESA-CCI (right). Darker green indicates deeper ALT. ESA-CCI data are aggregated on the CLM5 grid using a conservative projection method.

in the CLM5 results in the southern regions of Alaska, Canada, and particularly Siberia.

3.5.3 Active Layer Thickness (ALT)

The period annual averages and their difference of ALT between the CLM5 climate model and the ESA-CCI remote sensing product are shown in Fig. 3.12 and Fig. 3.13, respectively. Both figures exhibit stark contrasts at the borders of the domain, which we consider to be warm Arctic permafrost regions. Specifically, the CLM5 model shows large ALT values in these areas, with depths reaching up to 6 meters and even deeper locally. Additionally, some mountainous areas also display large differences of up to 4 m in ALT. However, in the majority of the domain above 65 °N, the CLM5 model is within 1 meter of the ESA-CCI product. These findings suggest that while the CLM5 model performs well in some areas, there are notable discrepancies in the warm Arctic permafrost regions and mountainous areas. However, ESA-CCI products are known to be underestimated in these areas (Heim et al., 2021).

Fig. 3.14 compares the annual period difference of ALT between the CLM5 model and in-situ observations from the CALM network. Similar to the comparison with the ESA-CCI product, there is a strong agreement between the CLM5 model and CALM observations above 65°North.

However, in the warm Arctic permafrost regions over Alaska, there were large differences between the two datasets. Unfortunately, due to a lack of local CALM stations, I was unable to confirm whether this pattern extends to the warm Arctic permafrost regions over Siberia. Overall, these results confirm that the CLM5 model performs well in

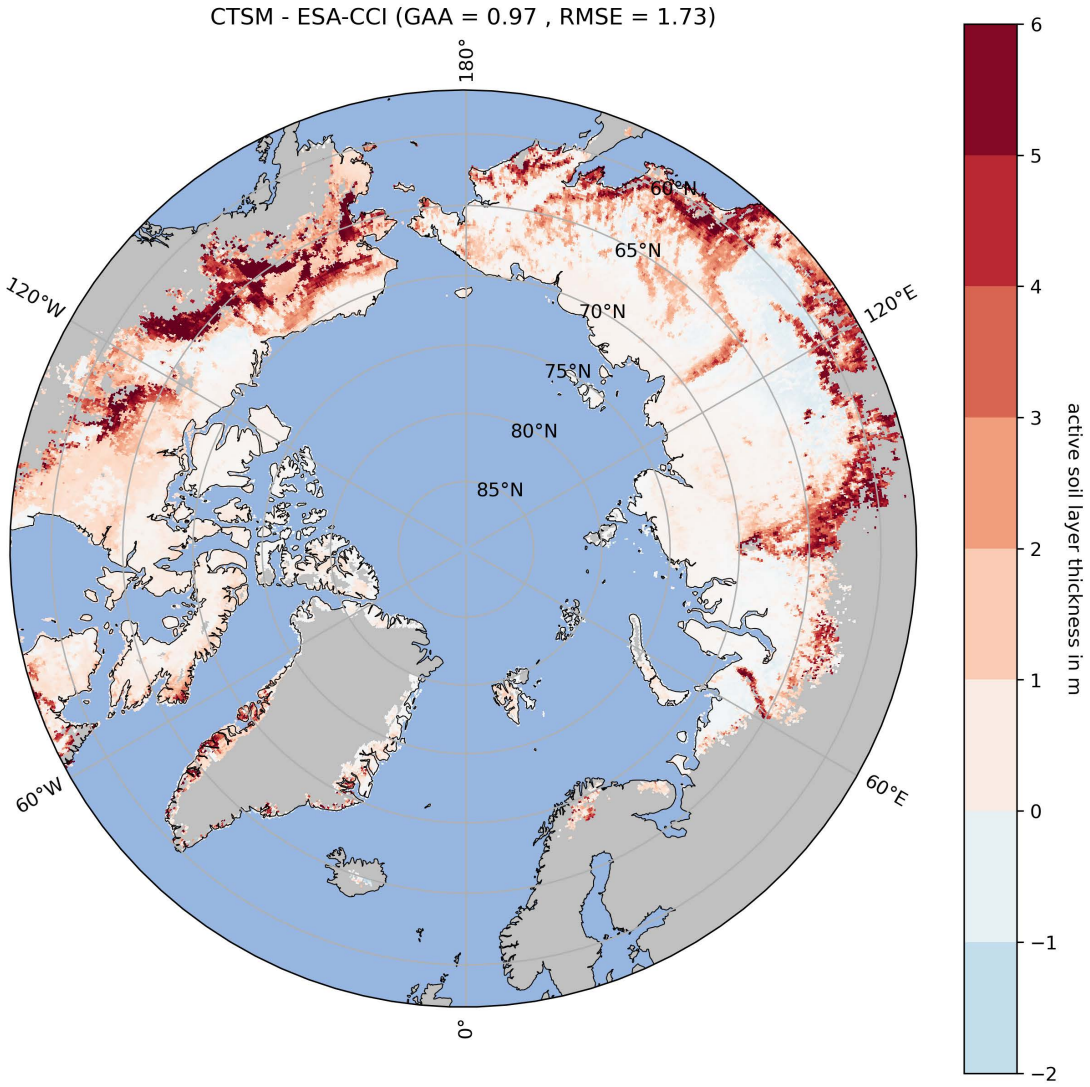


Fig. 3.13 Period (1997 to 2019) annual difference of Active Layer Thickness (ALT) in meters between CLM5 and ESA-CCI. Darker red indicates that CLM5 ALT is deeper than ESA-CCI ALT. ESA-CCI data are aggregated on the CLM5 grid using a conservative projection method.

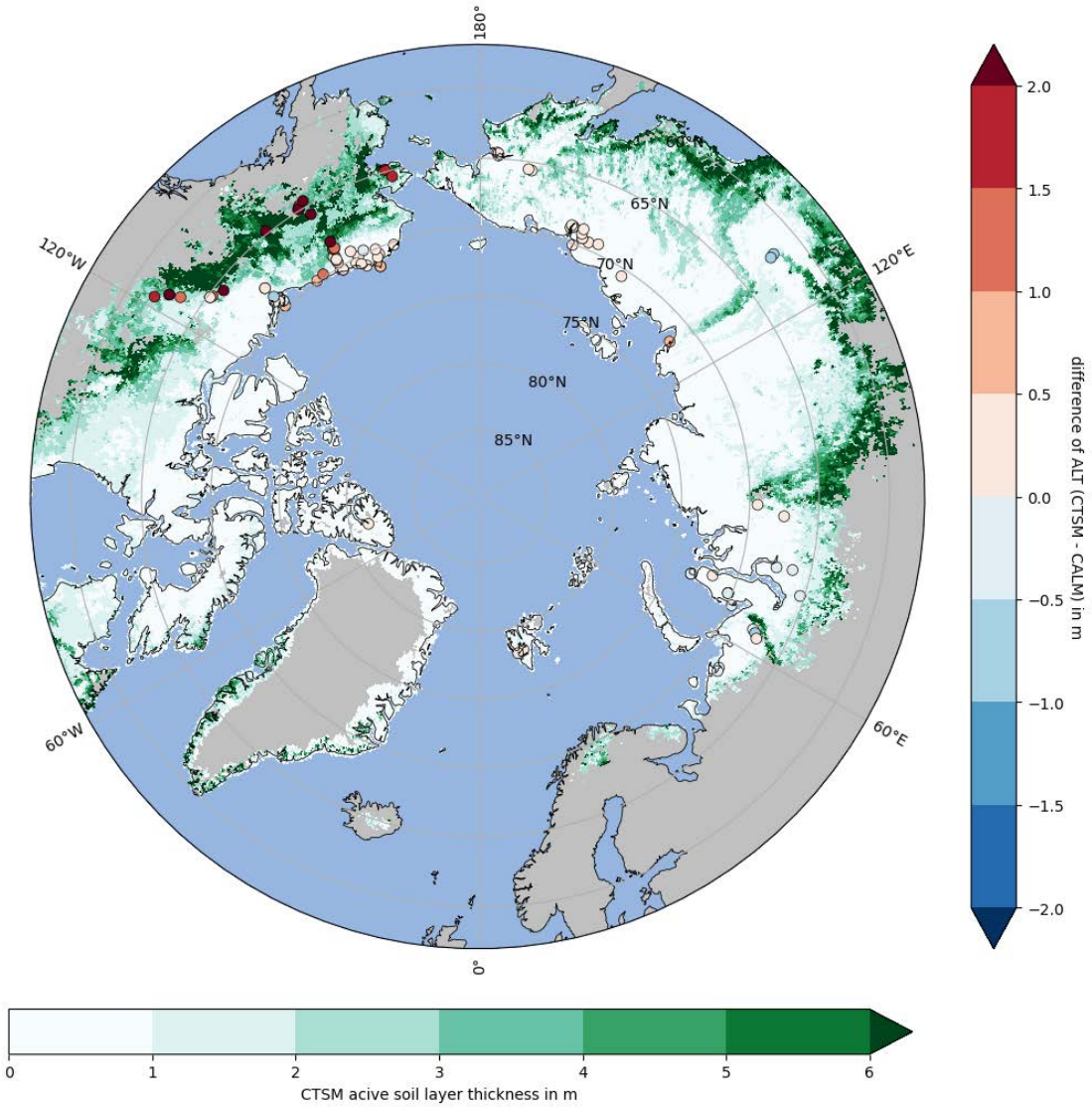


Fig. 3.14 Period (1990 to 2021) annual difference of Active Layer Thickness (ALT) in meters between CLM5 and CALM network. The background represents the ALT period average of CLM5. Coloured dots represent the ALT difference for each location between CLM5 and CALM. Darker red indicates CLM5 ALT is deeper, darker blue means CALM ALT is deeper.

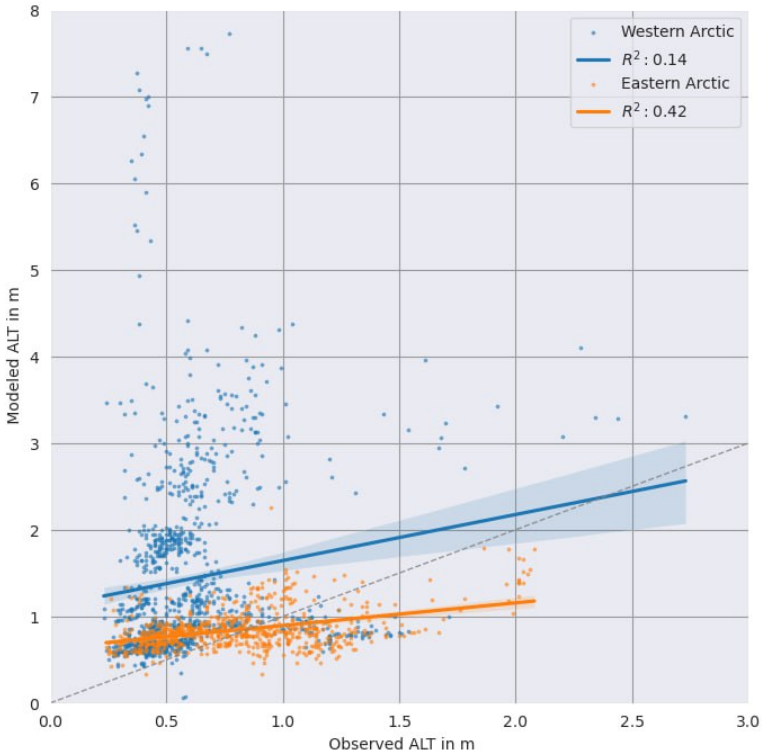


Fig. 3.15 Active Layer Thickness (ALT) year averages in meters of CLM5 against CALM for the western (blue) and eastern (orange) Arctic. The blue line represents an attempt of correlation between the observed and modeled values.

certain regions but struggles in others, particularly in warm Arctic permafrost regions.

Fig. 3.15 offers an alternative comparison of annual average ALT between CLM5 and the CALM network, without averaging over time periods. This allows us to capture the full variability of the ALT at the CALM sites. To further analyze the results, I divided the domain into two regions because of their distinct patterns: the West Arctic, which includes Greenland, Canada, and Alaska, and the East Arctic, which includes Europe and Siberia. The scatter plot in the West Arctic shows a poor performance of the CLM5 model with no significant correlation found with the in-situ data. This is likely due to most of the CALM stations in this region being located in the warm permafrost areas where we previously observed large differences between the CLM5 model and the ESA-CCI product. On the other hand, the comparison in the East Arctic was slightly better. However, the model tends to overestimate ALT above 1 meter.

3.6 Discussion and conclusions

This chapter presents the first evaluation of a regional land surface model to use three distinct datasets of in-situ observations and remote sensing products over the Arctic covering a period of 42 years (from 1980 to 2021). This aligns with the recommendations of previous studies in high-latitude regions calling for model evaluations that involve

multiple sources of observations and cover temporal periods of a minimum of 10-30 years (Mooney et al., 2020). Furthermore, my approach tackles the limitations of traditional evaluations by addressing depth, seasonality, and regional variations individually, instead of sacrificing one for the others. This allows for a comprehensive assessment of permafrost and soil temperature dynamics. Furthermore, to enable accessibility and facilitate future research, all algorithms are publicly available online.

As anticipated from my literature review, both comparisons against borehole station datasets and the European Space Agency Climate Change Initiative products (ESA-CCI) reveal a significant cold temperature bias across all months, depths, and regions, particularly over Siberia. However, the model displays significantly warmer values compared to ESA-CCI over mountainous areas, which is likely linked to the difference in resolution between the two products. This pattern could not be confirmed with the borehole dataset, as there are few stations situated above 1000 meters.

The results demonstrate that CLM5 is in strong agreement with ESA-CCI in simulating permafrost extent, albeit with a slight overestimation in the southern parts of the domain. Concerning the active layer thickness (ALT), there is significant agreement between CLM5 and both ESA-CCI products and the CALM network above 65°N.

However, this evaluation reveals a substantial overestimation of ALT over warm Arctic permafrost regions. Interestingly, the presence of the cold bias may potentially obscure areas where the ALT evaluation is accurate. It is plausible to anticipate a higher degree of overestimation in ALT if the model's cold bias is eliminated. Nevertheless, ALT primarily depends on the maximum annual ground temperature (MaxAGT), which might not be significantly impacted by the observed cold bias.

One major issue with CLM5 is its inability to properly represent seasonality of soil temperature, which leads to a spurious high yearly correlation, as the months tend to offset each other. This issue is particularly pronounced in Eastern Siberia. Taking the period and year average, as is typically done in the literature, smooths the results, but the actual comparison between the model and the observations is likely to be much worse. Similar trends emerge when examining the correlation between ALT annual averages, which is notably weaker than when averaging over specific time periods.

It should be noted that CLM5 is run at a relatively coarse horizontal resolution (12 km^2), while observations are mainly point-scale or representative of a much smaller area (1 km^2), and therefore not highly representative of large area means, particularly for complex terrain (Dankers et al., 2011). Moreover, grid-averaged soil properties and coarse-resolution atmospheric forcings might also introduce biases in the comparison between observed and simulated variables. Therefore, comparisons between observed and simulated results should be viewed as an indication of the model performance rather than a true in-situ validation.

In conclusion, the CLM5 model exhibited both strengths and weaknesses in sim-

ulating permafrost properties and dynamics. While it showed good agreement with ESA-CCI in simulating permafrost extent, its significant cold temperature bias, particularly over Siberia, is a major limitation. Moreover, the model's overestimation of the ALT over warm Arctic permafrost regions, coupled with its inability to properly represent seasonality, calls for caution when interpreting the model results. The findings of this study have important implications for guiding future model development, and for understanding the impacts of permafrost dynamics on climate and ecosystems.

I conduct two sensitivity experiments to address the cold bias observed:

- In Chapter 4, the first experiment (Research question 3) focuses on the summer snow-free period, where insulation is primarily influenced by soil texture and organic matter. By introducing a novel approach to derive these variables, I aim to enhance the representation of summer conditions.
- In Chapter 5, the second experiment (Research question 4) concentrates on the winter period, where snow thermal insulation dominates the heat transfer processes. By modifying the snow scheme, I aim to improve the representation of winter conditions.

CHAPTER 4

Sensitivity experiment on soil texture and soil organic matter

4.1 Introduction

In order to predict the future impacts of climate change on permafrost, it is crucial to develop accurate models that can simulate permafrost dynamics. LSMs have been developed to simulate the land surface and soil layers and predict future climate change. Soil thermal and hydrologic conductivities are key variables in these models, and they are greatly influenced by soil texture (proportions of sand, clay, and silt (SCS) in a soil) and soil organic matter (SOM), represented by the organic matter density (OMD) in LSMs (see Chapter 2). The latter emerges as a pivotal factor, particularly considering that soils in Arctic permafrost regions store substantial quantities of carbon (Schuur et al., 2008).

Typically, most LSMs derive the SCS and OMD proportions for each grid point through coarse resolution grid maps that are subsequently interpolated on the grid. Two prominently employed databases tailored to the high-latitude regions are the global Harmonized World Soil Database (ITPS, 2020) and the Northern Circumpolar Soil Carbon Database (NCSCD, Hugelius et al. 2013). The NCSCD gathers 1778 pedons from Arctic permafrost regions for an extensive circumpolar coverage. However, Hugelius et al. (2013) have highlighted various challenges associated with employing the NCSCD database:

- **Generalization and Ecological Fallacy**: Maps at smaller scales tend to be more generalized, thereby introducing the modifiable areal unit problem (MAUP) and ecological fallacy. MAUP underscores the imposition of arbitrary spatial divisions rooted in the analysis scale, while ecological fallacy occurs when findings derived from aggregated data are erroneously assumed to be valid for individual data points within that aggregation. In the context of the NCSCD, this pertains to using pixel values from gridded datasets for localized analyses.

- **Gridded Data Limitations**: Gridded NCSCD datasets consist of high-resolution pixels derived from polygon units of varying resolutions. Consequently, prudence

is essential when applying these datasets to local-scale analyses involving a limited number of pixels. Pixel values reflect their immediate spatial context and the surrounding conditions.

- **Uncertainties and Errors**: Quantitative assessments of uncertainties/errors linked to using NCSCD for SOC storage estimation are presently lacking. Distinct confidence levels have been ascribed to circumpolar SOC estimates, predicated on the availability of regional data. The extent of uncertainties hinges on landscape intricacy and the scale of study.

- **Landscape Complexity and Scale**: Endeavors conducted on a circumpolar scale are predisposed to heightened variability and uncertainty due to the intricate nature of landscapes. In contrast, regional investigations may yield lower uncertainties in SOC estimates.

This chapter addresses these caveats through a sensitivity experiment, wherein soil texture and SOM are derived using a novel approach. This sensitivity experiment seeks to assess the influence of this new method on permafrost dynamics by juxtaposing its outcomes with those attained through the original method.

Following this introduction, the second section provides a brief overview of soil texture and SOM, and how they affect soil thermal and hydrologic conductivities in CLM5. The third section describes the new method used in this experiment. The fourth section presents the results of the sensitivity experiment. Finally, the last section discusses implications for permafrost dynamics.

4.2 Soil texture and soil organic matter in CLM5

Model soil thermal and hydraulic conductivities are defined via mineral soil parameterizations that are dependent on soil texture and SOM, which are then used to simulate the soil temperature and moisture content. As explained in Chapter 2, soil texture represents the relative percentages of sand (particle size of 0.05 to 0.2 mm), silt (particle size of 0.002 to 0.05 mm), and clay (particle size below 0.002 mm), while soil organic carbon is represented as a parameter, called organic matter density (OMD) in kg/m^3.

In CLM5, soil texture and OMD are derived from Bonan et al. (2002a) and Hugelius et al. (2013), respectively. This dataset reports carbon down to 1 meter depth and is partitioned across the top 10 CLM5 layers (approximately 1m36 in depth) as in Lawrence et al. (2019).

The following subsections use Lawrence and Slater (2008) demonstration to link soil texture and SOM with soil thermal and hydraulic conductivities, and heat capacity.

4.2.1 Soil thermal conductivity

In CLM5, soil thermal conductivity is a parameter that is used to calculate the temperature of the soil at each depth. Soil heat conduction is solved numerically in CLM5 via the heat transfer equation 2.5 explained in Chapter 2:

$$c\frac{dT}{dt} = \frac{d}{dz}[\lambda_s \frac{dT}{dz}] \tag{4.1}$$

For each of the top ten soil layers, soil thermal conductivity is calculated as a combination of the saturated λ_{sat} and dry λ_{dry} thermal conductivities weighted by a normalized thermal conductivity:

$$\lambda = K_e \lambda_{sat} + (1 - K_e)\lambda_{dry} \tag{4.2}$$

The Kersten number K_e, is a function of the degree of saturation S_r and phase of water for unfrozen and frozen soils in the equations:

$$K_e = log(S_r) + 1 \text{ (unfrozen soil)} \tag{4.3}$$

$$K_e = log(S_r) \text{ (frozen soil)} \tag{4.4}$$

Below a degree of saturation S_r of 10^{-7}, the soil thermal conductivity is equal to:

$$\lambda = \lambda_{dry} \tag{4.5}$$

Below the top ten soil layers, λ is equal to $\lambda_{bedrock} = 3$ W m^{-1} K^{-1} which is the thermal conductivity assumed for the deep ground layers (typical of saturated granitic rock).

I focus exclusively on the saturated thermal conductivity, with no consideration for the dry thermal conductivity, as the latter remains independent of soil texture or SOM.

4.2.1.1 Saturated thermal conductivity

The saturated thermal conductivity λ_{sat} depends on the thermal conductivities of the soil solid, liquid water, and ice constituents:

$$\lambda_{sat} = \lambda_{soil}^{1-\theta_{sat}} \lambda_{liq}^{\frac{\theta_{liq}}{\theta_{liq}+\theta_{ice}}\theta_{sat}} \lambda_{ice}^{\theta_{sat}\left(1-\frac{\theta_{liq}}{\theta_{liq}+\theta_{ice}}\right)} \tag{4.6}$$

where λ_{soil}, λ_{liq} and λ_{ice} are the soil, liquid water and ice thermal conductivities, respectively, and θ_{sat}, θ_{liq} and θ_{ice} are the volumetric saturated, liquid and ice water contents, respectively.

In this experiment I focus exclusively on soil solid thermal conductivity λ_{soil}, which varies with the soil texture and SOM:

$$\lambda_{soil} = (1 - f_{om})\lambda_{st} + f_{om}\lambda_{om} \tag{4.7}$$

While this study primarily modifies soil solid thermal conductivity, the alterations can indirectly affect soil moisture due to its dependence on soil texture. This results in implicit changes in water content, even though there have been no adjustments to the calculation of thermal conductivity related to liquid water content.

In equation 4.7, there are three key variables. First, f_{om} is the soil organic matter fraction:

$$f_{om} = omd/omd_{max} \tag{4.8}$$

where omd is the organic matter density and omd_{max} = 130 kg m^{-3} is the maximum soil carbon density, equivalent to a standard bulk density of peat (Lawrence and Slater, 2008).

Second, λ_{st} is soil texture thermal conductivity and is empirically derived as:

$$\lambda_{st} = \frac{8.80\ \ (\%sand) + 2.92\ \ (\%clay)}{(\%sand) + (\%clay)}, \tag{4.9}$$

where $\%sand$ and $\%clay$ are the percentages of sand and clay.

Finally, λ_{om} is the thermal conductivity of SOM solid and has been empirically derived as equal to 0.25 W m^{-1} K^{-1} (Lawrence and Slater, 2008).

4.2.2 Soil heat capacity

The volumetric soil heat capacity introduced in equation 4.1 is a function of the heat capacities of the soil solid, liquid water, and ice constituents:

$$c = c_{soil}\,(1 - \theta_{sat}) + \frac{w_{ice}}{\Delta z_i}c_{ice} + \frac{w_{liq}}{\Delta z_i}c_{liq} \tag{4.10}$$

where c_{soil}, c_{liq} and Cc_{ice} are the specific heat capacities (J kg^{-1} K^{-1}) of soil solid, liquid water and ice, respectively. As with the thermal conductivity, this study focuses only on the soil solid. c_{soil} is derived at each of the first ten soil layers as:

$$c_{soil} = (1 - f_{om})c_{st} + f_{om}c_{om} \tag{4.11}$$

where c_{om} is the the heat capacity of SOM and equals 2.5×10^6 J m^{-3} K^{-1} (Lawrence and Slater, 2008). c_{st} is the heat capacity of the soil texture and is empirically derived as:

$$c_{st} = \left(\frac{2.128\ \ (\%sand)_i + 2.385\ \ (\%clay)_i}{(\%sand)_i + (\%clay)_i}\right) \times 10^6 \tag{4.12}$$

Below the first ten layers, c_{st} is simplified as the heat capacity of the bedrock $c_{s,bedrock} = 2 \times 10^6$ J m^{-3} K^{-1}.

4.2.3 Hydraulic conductivity

The soil hydraulic properties are also influenced by the soil texture and organic content. As the hydraulic properties of soil texture and organic soil can vary significantly, the overall hydraulic properties of each soil layer are calculated as weighted averages of the properties of the soil texture and organic content, similar to how thermal conductivity is determined.

The hydraulic conductivity is defined at the depth of the interface of two adjacent layers and is a function of the saturated hydraulic conductivity k_{sat}, the liquid volumetric soil moisture of the two layers, and an ice impedance factor. The scope of this study is limited to saturated hydraulic conductivity.

The saturated hydraulic conductivity is defined in the model as:

$$k_{sat} = (1 - f_{om})k_{st} + f_{om}k_{om} \tag{4.13}$$

where f_{om} is the soil organic matter fraction defined in Chapter 2. k_{st} is the soil texture hydraulic conductivity and is empirically derived as:

$$k_{st} = 0.0070556 \times 10^{-0.884+0.0153(\%sand)} \tag{4.14}$$

Finally, k_{om} is the hydraulic conductivity of SOM and is derived as:

$$k_{om} = max(0.28 - 0.2799 \times z_i/zsapric, k_{st}) \tag{4.15}$$

where z_i is the soil depth at the layer i and $zsapric$ = 0.5 m is the depth that SOM takes on the characteristics of sapric peat (Lawrence et al., 2018).

4.3 New method to derive soil texture and soil organic carbon

4.3.1 Description of the Obu method and experiment

The method presented here was initially developed for the ESA-CCI products (Heim et al., 2021) to shift away from using a global grid map for deriving soil texture and SOM. In the default configuration, CLM5 relies on data from Bonan et al. (2002a) and Hugelius et al. (2013). Instead, the model calculates the soil texture and SOM for each individual grid cell within the domain, depending on the distribution of plant functional types (PFTs) of each grid cell. For each PFT, a certain amount of SCS and OMD is given. This section explains the linkages between PFTs, SCS and OMD used in Heim et al. (2021).

"Stratigraphies" were generated from approximately 6500 samples collected from 700 pedons distributed across Arctic regions, including Siberia, Northern America, Scandinavia, Greenland, and Svalbard. These samples were meticulously collected with known volumes and were weighed both before and after the drying process. Subsequently, they use an elemental analyzer to determine their organic carbon content. The process involved several steps:

- Mass calculation for soil components: The initial step involved calculating the mass of various soil components.
- Water mass calculation: The mass of water in the samples was determined as the difference between the wet sample and dry sample mass.
- SOM mass calculation: To ascertain the mass of SOM, the dry sample weight was multiplied by the organic matter mass, which was then doubled. This multiplication by two served as a conversion factor between Soil Organic Carbon (SOC) and SOM, as proposed by Pribyl (2010).
- Mineral material mass calculation: Finally, the mass of mineral material was derived by subtracting the dry sample mass from the calculated SOM mass.

Fractions of volumetric contents for each of the components were calculated by dividing volume of component with a total volume of sample. Component volumes were calculated from the mass by assuming the following densities:

- $1g/cm^3$ for water
- $1.3g/cm^3$ for SOM (Farouki, 1981)
- $2.65g/cm^3$ for mineral material in soil (Hillel, 2003)

The volumetric fraction of air was calculated as the remainder after subtracting the sum of the other fractions from one.

Pedons were initially categorized based on the field-recognized land cover classes. Subsequently, averages were computed for each centimeter of the soil column, and standard deviation was employed to establish ranges of potential values associated with each land cover category. These generated soil columns were then further grouped into seven distinct depths, using a hierarchically constrained clustering approach. This depth categorization was designed to optimize the data for suitability as input in LSMs, ensuring that the resulting datasets are structured to enhance modeling accuracy and relevance.

The application of a PFT-based approach to derive soil texture and SOM has not yet been assessed within the realm of land surface modeling. To address this knowledge

gap, I conducted a sensitivity experiment using CLM5. Two distinct simulation runs were conducted: the "control run," which shares the configuration discussed and evaluated in Chapter 3, and the "Obu run," where the conventional method for deriving soil texture and SOM is replaced with the Obu method.

4.3.2 Differences in SCS and OMD between the control run and the Obu run

Figure 4.1 shows the column average (first 10 layers) differences between the control run and Obu run in terms of the distribution of SCS.

In the control run, sand content exhibits a stark contrast between western and eastern Arctic regions, with values reaching around 80% in parts of Canada but only around 20-30% in Siberia. The Obu run shows a more linear distribution of sand, with values decreasing as one moves further north. However, the Canadian archipelago is as an exception to this pattern, as the Obu run shows a significant increase in sand content in areas outside of the permafrost region, such as southern Russia, Europe, and southern Canada. Within the permafrost region, there is a general trend of decreased sand content, especially in inland Canada and the Victoria Island, with the exception of Alaska where sand content has increased. Furthermore, the distribution of sand in the Obu run is much more heterogeneous compared to the control run, which exhibited more distinct patches.

Clay content in the control run is generally around 20-30% throughout the domain, except for parts of northern Canada where values are below 10%. In contrast, the Obu run shows much more homogeneous distribution of clay with values around 10%, resulting in an overall reduction in clay content, particularly in northern Canada.

Silt content increased over the entire domain in the Obu run, except in southern Russia and southern Canada, with a significant increase in northeastern Canada. Overall, these changes in the distribution of soil minerals suggest that there may be significant changes in the hydraulic properties of the soil, which can impact the water cycle and other related Arctic processes.

Figure 4.2 shows the column average (first 10 layers) distribution of OMD in the soil for the control run, the Obu run, and their difference. In general, OMD is more heterogeneous in the control run, with values ranging from 10 kg/m^3 to 80 kg/m^3 across the domain. In contrast, the Obu run shows a more homogeneous distribution of OMD, with values ranging from 40 kg/m^3 to 50 kg/m^3 across most of the domain.

Comparing the two runs, OMD values decreased in the Obu run for most of Canada and Siberia, with only the Alaskan region showing an increase. This reduction in OMD values is particularly evident in the northern regions of Canada, where OMD values in the control run are generally higher than in the Obu run. Overall, these results suggest that the changes in soil properties due to the Obu experiment have led to a more ho-

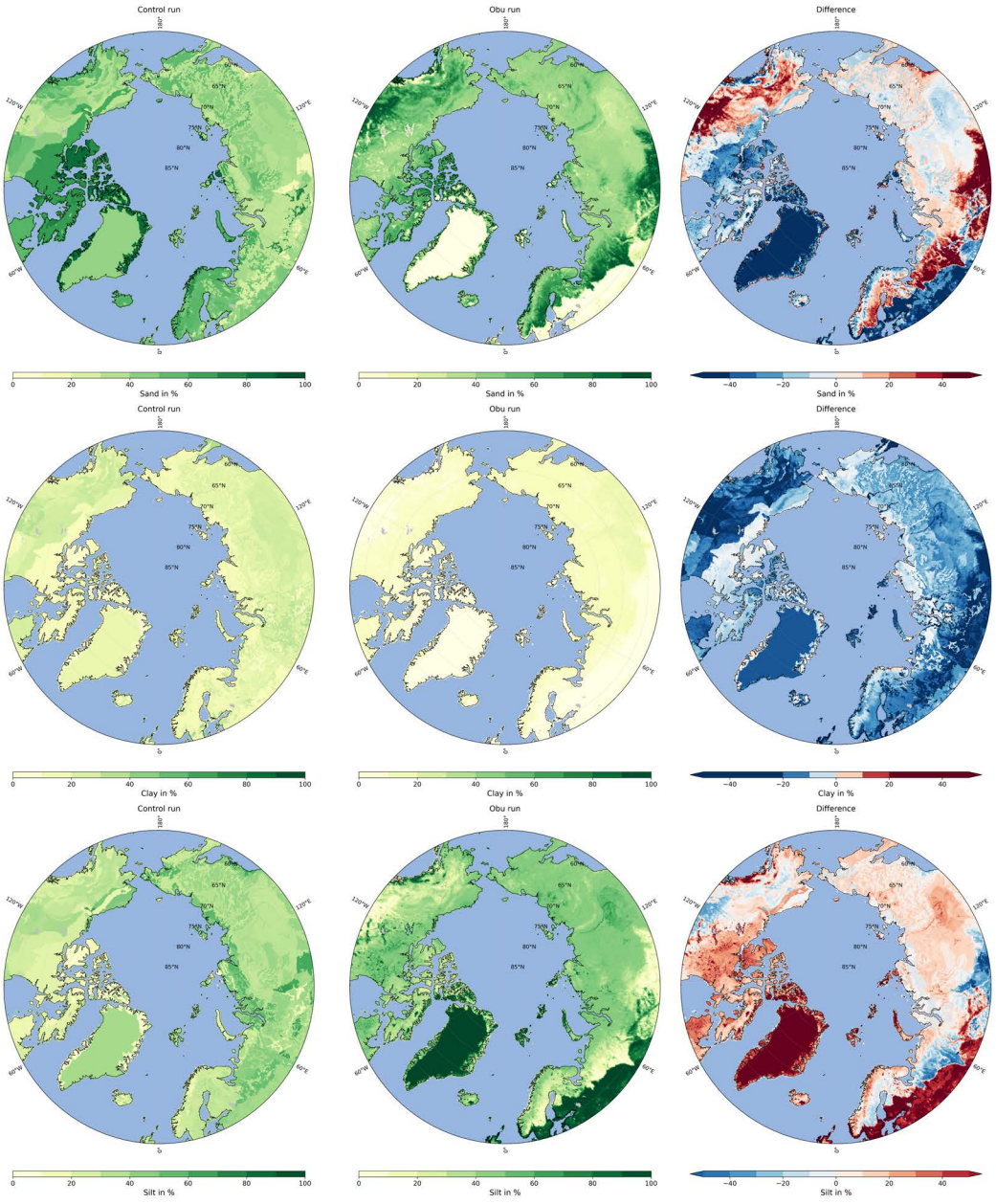


Fig. 4.1 Column average (first 10 layers) of soil texture: sand (first row), clay (second row) and silt (third row) in the control run (first column), Obu run (second column) and their difference (third column). Darker red indicates higher values in the Obu run. Unit is %.

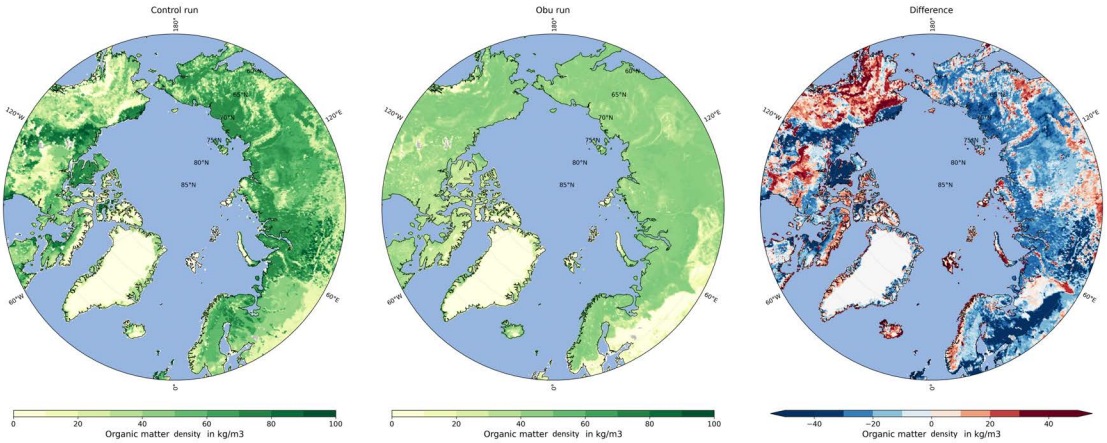


Fig. 4.2 Column average (first 10 layers) of organic matter density in the control run (first column), Obu run (second column) and their difference (third column). Darker red indicates higher values in the Obu run. Unit is kg/m^3, assuming a carbon content of 0.58 gC per gOM.

mogeneous distribution of OMD across the domain, with a decrease in OMD values in most regions.

4.4 Results

4.4.1 Soil temperature

4.4.1.1 Comparison between the Obu and control runs

Figure 4.3 shows the four seasons averaged difference in soil temperature between the Obu and control runs for the 1980-2021 period. In winter, there is a significant increase in temperature over most areas, up to +2°C, except over the Taymir Peninsula and the northwest where temperatures are similar. The only significant decrease in temperature is observed in the center of Canada. Autumn and spring share similar results to winter.

In summer, the increase in temperature is more pronounced over northern Europe and southern Siberia, with an increase of up to +3°C. The rest of Siberia shows increases between +1 to +2°C. In contrast to the January results, the western Arctic experiences an increase exclusively in the center of Canada. Over northwest Canada and especially over Alaska, there is a decrease in temperature of up to -3°C.

Figure 4.4 presents the year-averaged difference in soil temperature at a depth of 1 meter between the Obu and the control runs. In general, the Obu run exhibits an overall increase in soil temperature, ranging from +0.5 to +2°C. The most substantial temperature increases are observed in central Siberia and in proximity to the Mackenzie River in Canada. Conversely, certain regions in Canada and Alaska display a decrease in temperature, particularly notable in the western areas of Hudson Bay.

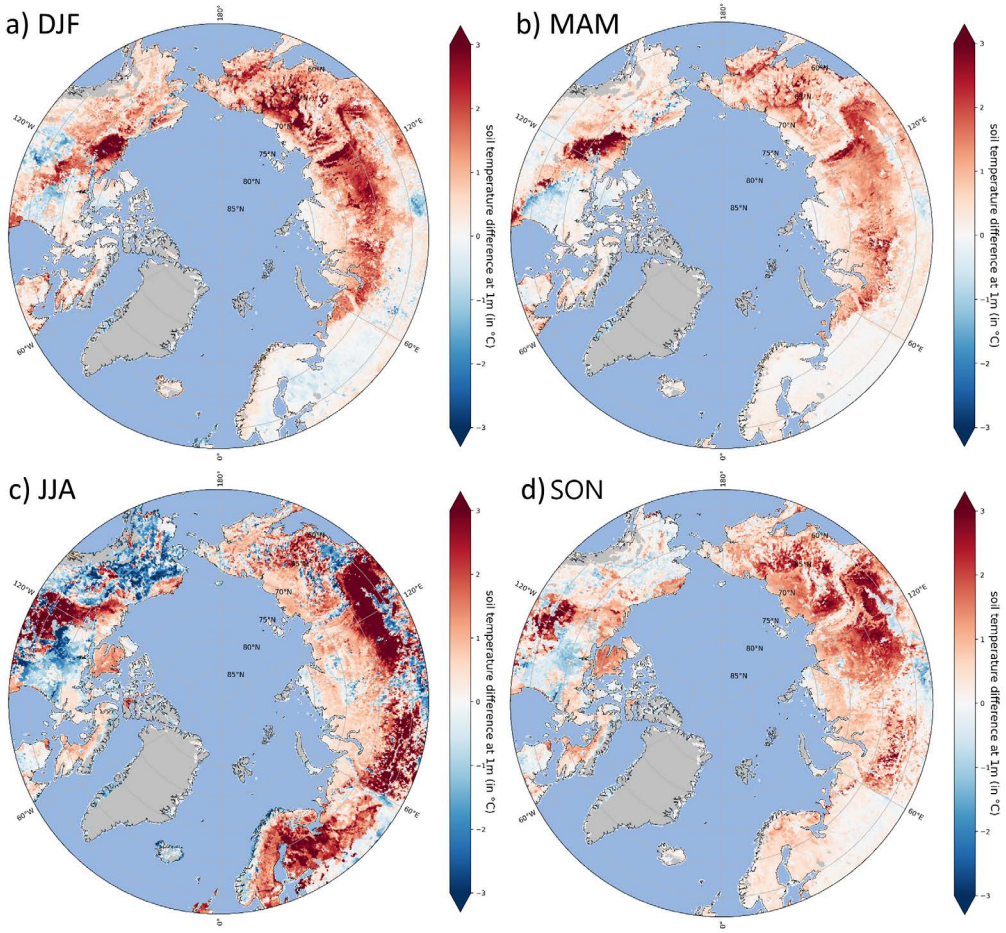


Fig. 4.3 Period averaged (1980-2021) soil temperature difference between the Obu and control runs at 1 m depth in°C for four seasons: a) December, January, February (DJF), b) March, April, May (MAM), c) June, July, August (JJA) and d) September, October, November (SON). Darker red indicates the Obu run is warmer than the control run.

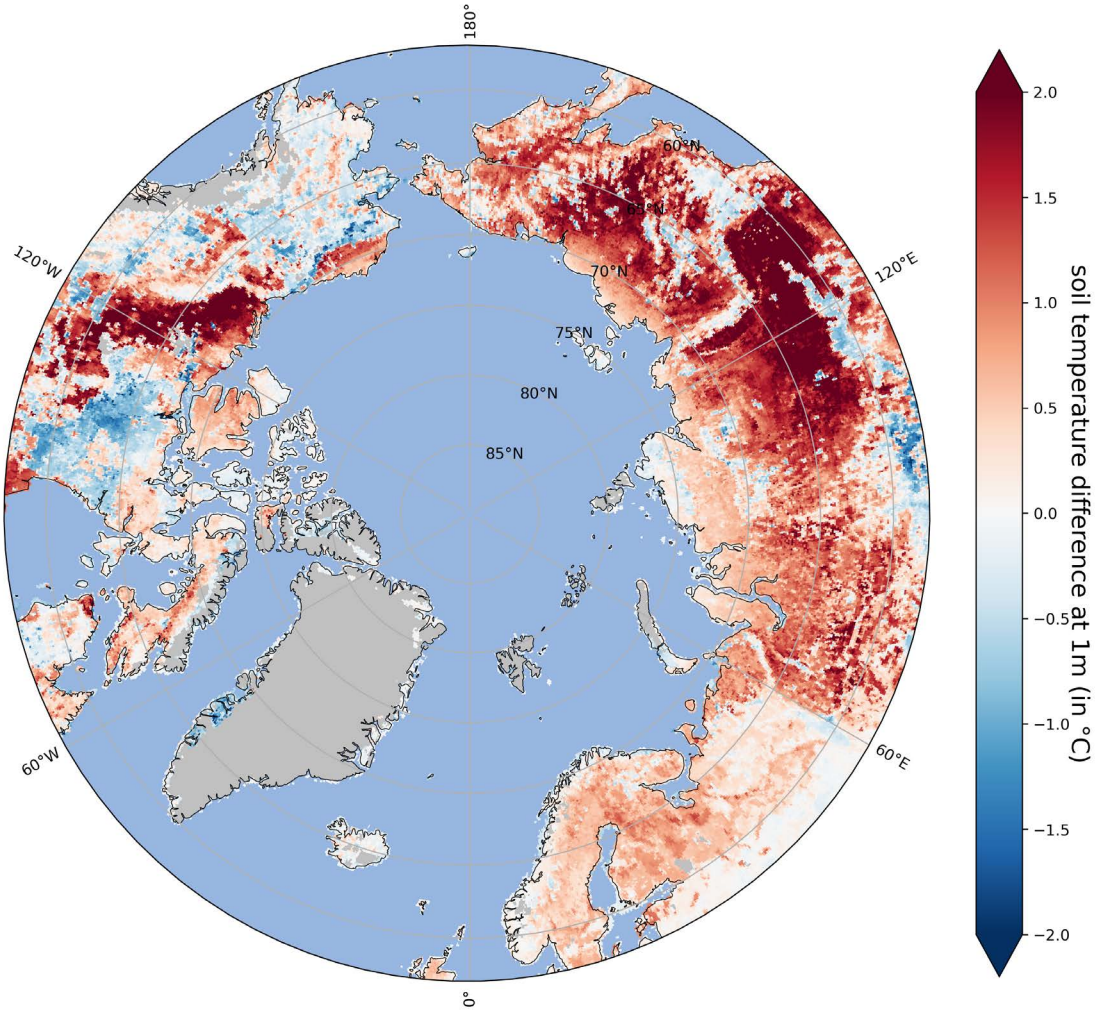


Fig. 4.4 Yearly period averaged (1980-2021) soil temperature difference between the Obu and the control runs at 1 m depth in°C. Dark red indicates that the Obu run is warmer than the control run.

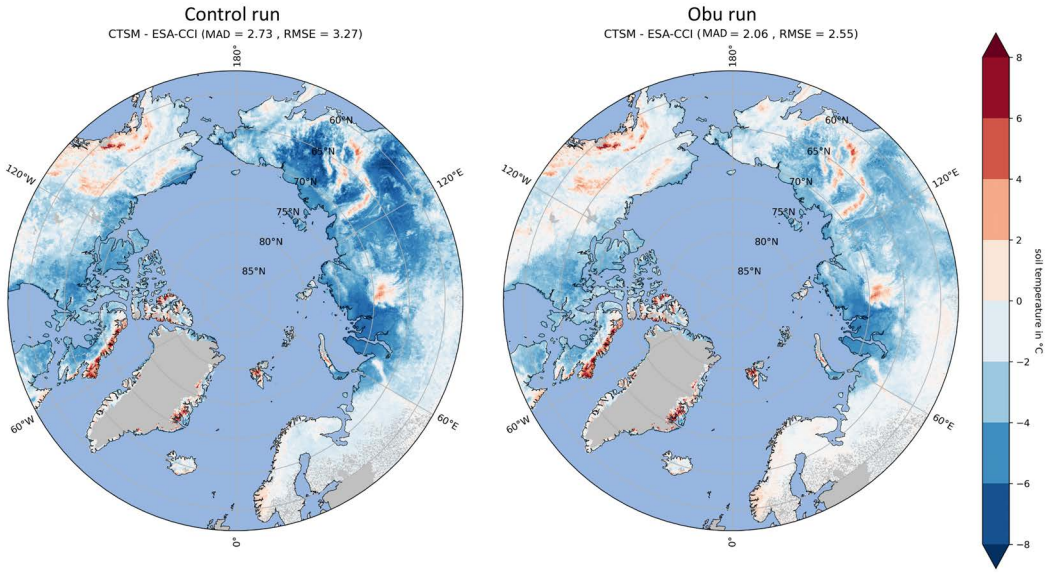


Fig. 4.5 Period (1997 to 2019) MAGT at 1 m difference between CTSM and ESA-CCI in°C for the control run (left) and Obu run (right). Dark blue means CTSM soil temperature is colder than ESA-CCI. ESA-CCI data are aggregated on the CTSM grid using a conservative projection method.

4.4.1.2 Comparison between the Obu run and ESA-CCI

Figure 4.5 compiles an evaluation of the 1-meter soil temperature year-averaged, comparing the results from the control run against the ESA-CCI dataset and the same evaluation with the Obu run. The Obu run has effectively mitigated a part of the cold bias observed in the control run. While the spatial heterogeneity of the bias stays identical, we observed a reduction in the cold bias of up to 4°C. This decrease is greater over areas like Siberia where the cold bias is strong in the control run. Overall, this level of bias aligns reasonably well with what has been observed in the ESA-CCI evaluation (Heim et al., 2021). Only the Ural region show a bias above -4°C. The warm bias observed over high-altitude regions is still present. The MAD and RMSE show a strong improvement, decreasing from 2.73 in the control run to 2.06 in the Obu run, and from 3.27 to 2.55, respectively.

4.4.1.3 Comparison between the Obu run and the 295GT dataset

Figure 4.6 illustrates the period-averaged monthly soil temperatures at various depths (-20 cm, -80 cm, -160 cm, and -320 cm) for the observations (in black), control run (in blue), and Obu run (in green) in°C. These values are derived from an average of 295 stations over the entire period (1980-2021). With a focus on seasonal and column-depth variations, this comparative analysis emphasizes the improvements realized through the Obu simulation in contrast to the control run.

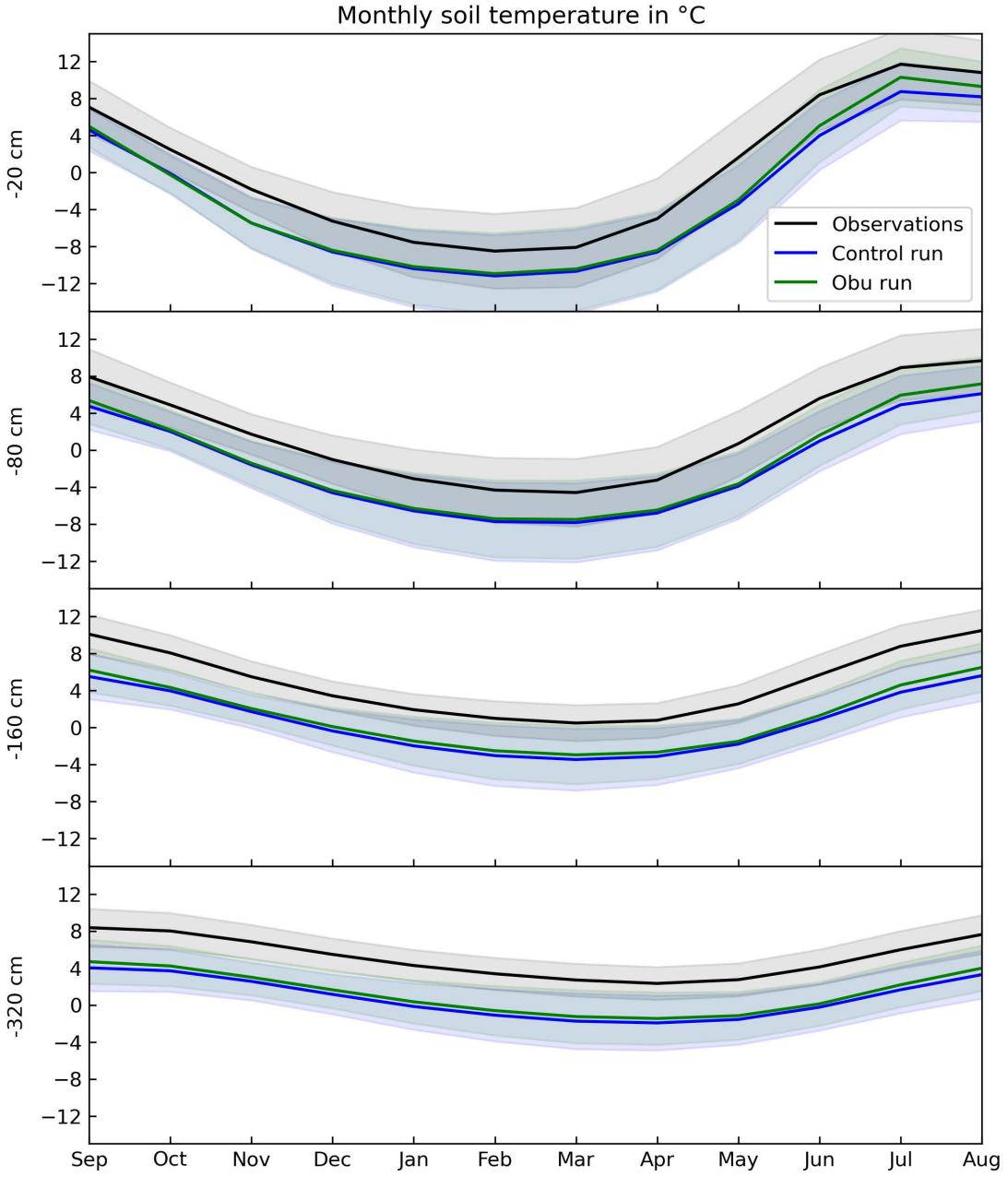


Fig. 4.6 Period averaged (1980-2021) of monthly soil temperature at 4 different depths (-20, -80, -160 and -320 cm) for the observations (black), control run (blue) and Obu run (green) in°C. For each depth, an average depth range was taken as follows: -20 = [0, -40], -80 = [-41, -120], -160 = [-121, -200], -320 = [-201, -440]. The area represents the RMSE over all years. All values come from an average of the 295 stations through the full period.

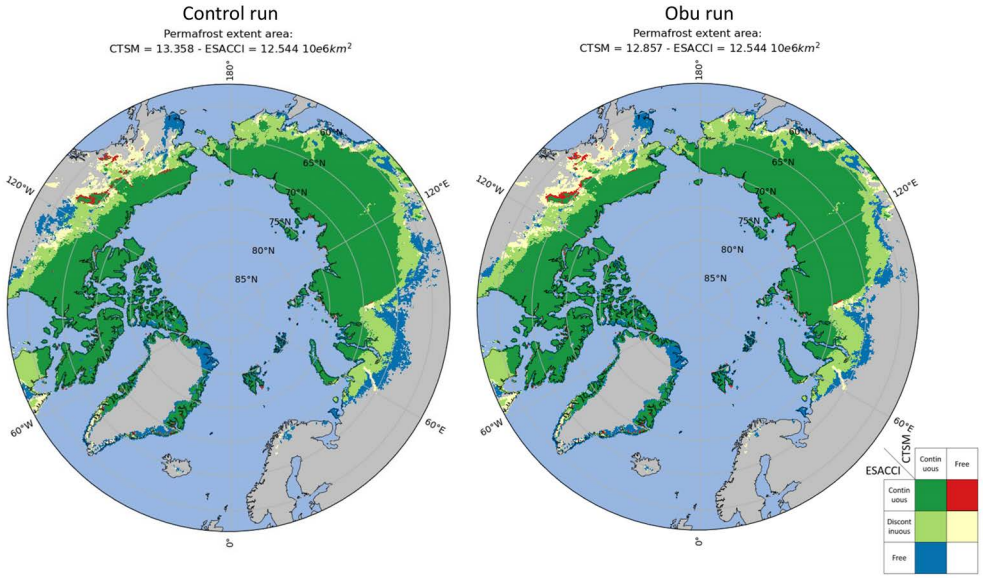


Fig. 4.7 Permafrost extent area mask difference between CLM5 and ESA-CCI for the control run (left) and Obu run (right). Refer to the legend for colour meanings. ESA-CCI data are aggregated on the CLM5 grid using a conservative projection method.

On average, the model exhibits a tendency to underestimate soil temperatures at all depths, with a bias of approximately -4°C. However, the Obu run presents a slight reduction in this bias compared to the control run, notably during the summer months and primarily within the uppermost soil layer. However, this effect becomes less pronounced at greater depths, as the properties of soil increasingly overlook the insulating properties of SOM. These observations suggest that while the model consistently underestimates soil temperatures, the Obu run introduces some improvements, particularly in mitigating the bias during summer and within the shallow soil layers. Nevertheless, further refinements may be necessary to achieve a closer alignment with observed soil temperature values.

Appendix C presents an analysis that supplements the regional soil temperature study discussed in this chapter and compares temperature data from 10 selected stations, using the Obu run.

4.4.2 Permafrost extent

Similar to the soil temperature in Figure 4.5, the permafrost extent difference between CLM5 and ESA-CCI remote products is depicted in Figure 4.7, comparing the control run (left) and the Obu run (right). In general, the Obu run demonstrates a closer alignment to ESA-CCI, with the total permafrost area measuring 12.857 $\times 10^6$ km^2, compared to 12.544 $\times 10^6$ km^2 for ESA-CCI, while the control run recorded a larger extent of 13.358 $\times 10^6$ km^2. Significant changes are observed in the southern region of

Canada, where all permafrost areas simulated by CLM5 but not captured in ESA-CCI have been eliminated. Additionally, the extensive permafrost area in the Southwestern Europe, previously absent in ESA-CCI, experienced a substantial reduction. Despite the improvements, disparities between ESA-CCI and the model persist, particularly in northern Greenland, western Alaska, and southern regions of western Siberia and Europe. These findings indicate that the Obu run successfully enhances the model's representation of permafrost extent, aligning it more closely with ESA-CCI, although some discrepancies still exist in specific geographic regions.

4.4.3 Active Layer Thickness (ALT)

4.4.3.1 Comparison between the Obu and control runs

Figure A.6 in Appendix A, illustrates the difference in period-averaged (1980-2021) active layer thickness (ALT) between the Obu and the control runs. The results indicate that ALT generally exhibits greater depths in Siberia and Europe, while it is comparatively shallower in Alaska and in western Hudson Bay in the Obu run as compared to the control run. Specifically, the southern regions of Siberia experience an increase in ALT of up to 2 meters deeper, while these changes become less pronounced over mountainous areas. Conversely, the ALT in Alaska and Canada shows a decrease of around 0.5 to 1 meter, with some localized exceptions in the southern regions and along the northern coasts. These findings suggest that the Obu run yields substantial modifications to the ALT distribution, deepening the active layer in Siberia and Europe, while reducing its depth in Alaska and Canada, though local variations are evident.

4.4.3.2 Comparison between the Obu run and ESA-CCI

Similar to the soil temperature in Figure 4.5, ALT difference between CLM5 and ESA-CCI remote products are shown in Figure 4.8, comparing the control run (left) and the Obu run (right). As anticipated, the Obu run leads to an overall increase in bias across the domain. Notably, the bias becomes more pronounced moving southward. However, the rise in MAD remains modest, transitioning from 0.97 to 1.24, while the RMSE remains relatively consistent, shifting only slightly from 1.73 to 1.93.

4.4.4 Soil liquid and ice water

Figure A.7 in Appendix A showcases substantial differences between the Obu and the control runs regarding the sum of soil liquid and ice water in the top 1-meter column. These differences exhibit both positive and negative values across various regions. Firstly, significant decreases exceeding 150 kg/m^2 of soil water are observed in southern parts of Europe, as well as western and central Siberia. Areas near the Mackenzie River

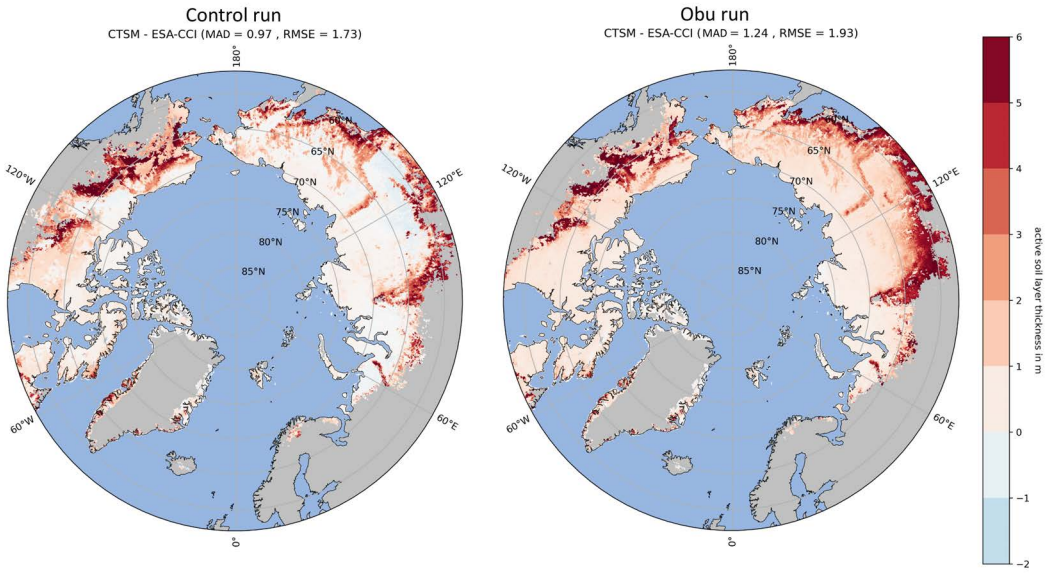


Fig. 4.8 Active layer thickness difference between CTSM and ESA-CCI in meters for the control run (left) and Obu run (right). Darker red means CTSM ALT is deeper than ESA-CCI. ESA-CCI data are aggregated on the CTSM grid using a conservative projection method.

also displayed similarly pronounced decreases of more than -150 kg/m^2. Conversely, certain regions exhibit an increase in the sum of soil water. Notably, the western region of Hudson Bay demonstrates values around +100 kg/m^2, while Alaska shows increases around +50 kg/m^2. The remaining areas within the domain exhibit differences ranging between -50 to +50 kg/m^2. These findings indicate substantial alterations in the distribution of soil liquid and ice water within the 1-meter top column in the Obu run compared to the control run.

4.5 Discussion and conclusions

In land surface modelling, soil texture and soil organic matter (SOM) are recognized as two of the most influential factors that shape soil properties, physics, and thermal dynamics in Arctic permafrost regions. Understanding the interplay between these variables is crucial to accurately simulate permafrost dynamics in LSMs.

Deviating from the conventional use of global coarse-resolution maps (e.g., Hugelius et al. 2013), this study pioneers the use of a Plant Functional Type (PFT)-based approach to derive soil texture and SOM, called the Obu method. While PFT-based methods have shown promise in recent remote sensing products like ESA-CCI (Heim et al., 2021), their effectiveness in LSM remains untested. This study aims to bridge this gap by evaluating the performance of PFT-based approaches to understand their potential contributions to permafrost dynamics and soil physics modeling.

To achieve this, I conducted a sensitivity experiment using CLM5 with two distinct runs: the "control run" with the configuration discussed in Chapter 3, and the "Obu run" where the conventional method for deriving soil texture and SOM is replaced by the Obu method. This investigation centered on assessing the repercussions of this experiment on key parameters, including soil temperature, permafrost extent, ALT, and soil liquid and ice content.

Overall, the Obu method led to a more homogeneous distribution of soil organic matter density (OMD) across the domain, marked by a decrease in OMD values in most regions. However, changes in soil texture exhibit a more complex spatial pattern.

Key findings indicate that the Obu run induces significant modifications to the soil temperature profile, characterized by a notable warming trend in most regions. First, the increase in soil and content has been associated with summer temperature decreases in Alaska, primarily due to enhanced water drainage and lower water retention capacity compared to soils with higher clay or silt content, leading to higher porosity and lower thermal conductivity. Conversely, an increase in sand content results in autumn temperature increases in the Mackenzie Basin, attributed to a higher amount of latent heat released during freezing. Secondly, decreases in SOM content are linked to summer temperature increases in the Mackenzie Basin and the eastern Siberian regions. This effect is attributed to the insulating properties of SOM, characterized by high porosity and low thermal conductivity. Although a warming effect is also observed in the northern regions, it is less pronounced due to the insulating snow cover in winter and in deeper layers where soil properties increasingly overlook the insulating properties of SOM. Nevertheless, localized cooling effects are evident in specific areas of Canada and Alaska.

Comparing the Obu run to the observation datasets reveals a significant decrease of the cold bias observed in the control run. In most regions, the Obu run demonstrates an improvement compared to the control run, with a cold bias below -4°C. However, an exception is observed in the Ural region, which exhibits a more substantial cold bias, necessitating further investigation. The summer season is where the highest increase emerges. During the winter season, no clear relationship can be drawn, primarily due to the insulating properties of the snow cover, which dampens thermal dynamics (Zhu et al., 2019). Regarding year average, the global mean absolute deviation and the spread of temperature values improved (from 2.73 to 2.06 for MAD, 3.27 to 2.55 for RMSE), indicating a better overall representation of soil temperatures.

For permafrost extent, the Obu experiment greatly solved the overestimation of permafrost present in the control run over the southern regions in Siberia (from 13.358 to 12.857 $\times 10^6$ km^2, compared to 12.544 $\times 10^6$ km^2 for ESA-CCI).

Regarding active layer thickness (ALT), the comparative analysis between the Obu and the control runs reveals significant increases in ALT over most regions, with the

greatest impact observed in the southern areas and a marginal effect in the northern regions. Consequently, we observe an slight increase in the MAD and RMSE, from 0.97 to 1.24 and 1.73 to 1.93, respectively. This observed increase in ALT can be attributed to the fact that ALT strongly depends on the maximum annual ground temperature (MaxAGT). As the Obu experiment increases temperatures in the summer, it directly impacts the maximum thaw depth, leading to greater ALT values. However, the impact of these changes in ALT is limited when compared to the overall improvements in soil temperature and permafrost extent achieved in the Obu run.

It is important to emphasize that my experiment incorporates changes in both soil texture and SOM, making it challenging to dissect the exact impact of each factor individually. Nonetheless, my findings highlight the high sensitivity of models to soil texture and SOM and the pivotal role of soil composition in governing heat transfer processes and shaping the seasonal variation of soil temperatures in Arctic permafrost regions. Furthermore, it is crucial to acknowledge that the quality of input datasets also plays a substantial role in modeling accuracy, and addressing uncertainties in these datasets remains an ongoing challenge in permafrost research.

CHAPTER 5

Sensitivity experiment on snow thermal conductivity

5.1 Introduction

Snow is a crucial component of Arctic's climate system, exerting significant influence on various aspects of the region's surface conditions. The snow cover acts as a thermal insulator by limiting heat loss in the winter (Lawrence and Slater, 2010; Li et al., 2021; Royer et al., 2021), but its insulating properties are highly variable and insufficiently detailed in Earth System Models (ESMs) (Barrere et al., 2017). With the Arctic region experiencing rapid warming, especially in the winter (Post et al., 2019), a significant reduction in snow cover has been observed (Meredith et al., 2019). Model studies suggest that the recent reduction in the extent of snow cover plays a significant role in the thawing of permafrost in northern regions and may become increasingly important as Arctic warming continues (Park et al., 2015; Mudryk et al., 2018; Overland et al., 2019).

The Arctic has long been recognized as a difficult region to study due to its inherent remoteness, the lack of observations (Domine et al., 2019; Royer et al., 2021), and human errors (Menard et al., 2021). The lack of information on snow properties in Arctic permafrost regions places a major limitation on permafrost and climate modelling more generally (Domine et al., 2016; Gouttevin et al., 2018).

The depth and density of the snowpack are fundamental in determining its insulation capacity, and are inherently intertwined with the temperature conditions of the ground surface. The efficiency of snow cover as an insulator intensifies with thickness, with optimal insulation observed at an approximate depth of 25 cm (Slater et al., 2017). Beyond this point, further increases in snow depth do not yield a significant enhancement of insulation properties.

The diversity of snowpack density properties can be categorized into two main snow types in Arctic environments: depth hoar and wind slab (Domine et al., 2018).

- Depth hoar refers to a type of snow crystal structure that forms within the snowpack due to temperature gradients and water vapor fluctuations. It typically develops in cold, dry conditions with large temperature differences between the ground and the overlying snow layer. The temperature gradient, especially at

the beginning of the snow season, is affected by soil moisture, as greater moisture maintains the soil at 0°C longer, allowing large temperature gradients to persist, driving large upward water vapor fluxes, and favoring depth hoar formations. Depth hoar crystals are characterized by their large, faceted, and often cup-shaped grains. These crystals are highly porous and have a low density, making them poor heat conductors. Due to their porosity, depth hoar crystals trap a significant amount of air within their structure, enhancing their insulating properties.

- Wind slab is a distinct layer of snow, the commonly called "tundra snow", that forms due to wind transport and deposition. Wind can transport snow particles, causing them to accumulate in locations with low vegetation and high wind speeds. Wind slab layers have a relatively high density compared to other snow layers. The compacted nature of wind slabs reduces their overall porosity and air content, resulting in better heat conductivity and decreased insulation properties compared to other snow types.

Models such as CLM5 (Dutch et al., 2022), Crocus (Barrere et al., 2017) and SNOWPACK (Domine et al., 2019; Gouttevin et al., 2018) struggle to represent these two phenomena. Vertical density profiles simulated by these models exhibit noteworthy discrepancies, with observed snow density deviating significantly from the model's estimations, both in the top wind slab and bottom depth hoar layers of the snowpack (Dutch et al., 2022). A paper under review is trying to solve this issue (Fourteau et al., 2023).

Studies show that neglecting the role of depth hoar in providing thermal insulation properties to Arctic snowpacks can have significant consequences for soil temperature (Gouttevin et al., 2018; Royer et al., 2021; Dutch et al., 2022). For instance, CLM5 tends to overestimate the thermal conductivity of snowpack and, thus, underestimate its insulating capacity (Dutch et al., 2022), which is consistent with the cold bias identified in Chapter 3.

The crux of the challenge lies in the representation of the thermal property variability within Arctic snow, aiming to rectify the prevalent cold bias found in most LSMs. I hypothesize that a modification to the snow thermal conductivity schemes, which quantify snow's insulation in models, could effectively capture the sensitivity inherent in Arctic snow, thereby restoring an accurate thermal insulating function to the snowpack. To realise this endeavor, I compare a sensitivity experiment using a alternative snow thermal conductivity scheme, to the default scheme used in CLM5.

After this introduction, the second section of this chapter elaborates on snow thermal conductivity studies and its implementation in snow modules. The third section describes the CLM5 snow module and the sensitivity experiment. The fourth section presents the results. Finally, the last section discusses implications of the experiment

for permafrost dynamics.

5.2 Snow thermal conductivity

The insulation capability of snow is commonly quantified through its thermal conductivity. In the presence of snow, snow thermal conductivity determines the rate of heat exchange between the soil and the atmosphere. Some research suggests that it is one of the most critical model parameters in snow physics (Hu et al., 2023) and a significant source of uncertainty in LSMs (Langer et al., 2013).

Heat transfer processes in snow exhibit greater complexity compared to those in soil. In snow, heat is transferred by conduction, convection, and radiation across the air space, and by the the movement of vapor by sublimation and condensation (discussed in the next section). However, as a result to low temperatures, the effect of radiation transfer is usually not significant (Yen, 1981). Due to these complications, the thermal conductivity of snow is expressed as an effective thermal conductivity K_{eff}, which is intended to embrace all of the heat-transfer processes.

Snow exhibits a low K_{eff}, generally falling within the range of 0.01-0.7 $Wm^{-1}K^{-1}$ (Gouttevin et al., 2018). Researchers (Sturm et al., 1997; Domine et al., 2016; Dutch et al., 2022) have found that tundra snowpacks typically display K_{eff} values toward the lower end of this range.

Numerous investigations have measured snow thermal conductivity (Sturm et al., 1997; Dutch et al., 2022) and establish statistical relationships (Anderson 1976, Mellor 1977, Yen 1981, Jordan 1991) to formulate an equation between K_{eff} and snow density, based on experiments made in laboratories or on different snowpacks around the world. It is therefore reasonable to assume that a statistical relationship derived from alpine snowpack experiments will differ greatly from the actual K_{eff} of an Arctic snowpack or on ice sheet, and inversely. Interestingly, these equations are still used in the most advanced snow models like Crocus and SNOWPACK (Domine et al., 2019).

Nonetheless, Sturm et al. (1997) contend that these models lack in-depth critical analysis or evaluation of the original data, presenting fitted curves without providing the actual data points (e.g., in Yen 1981). Over time, this unintended consequence has given the illusion of greater regularity and less dispersion in the compilation of measurements than is actually true. Furthermore, this approach has hindered readers from assessing the quality of the data. Curves established on the foundation of robust experiments with extensive datasets are treated equally to those built on insufficient data. Similarly, curves developed using indirect measurement techniques (such as deriving conductivity from variations in snow temperature) and those with lower accuracy are given the same weight as curves derived from data acquired using more precise laboratory methodologies.

In a concerted effort to address these concerns, Sturm et al. (1997) introduced what we consider to be the most expansive dataset available to date. Encompassing approximately 488 measurements collected mostly in the Arctic and Antarctica across various types of seasonal snow, this dataset includes meticulously recorded snow sample temperatures and comprehensive descriptions, facilitating a comprehensive exploration of the influence of attributes beyond density on thermal conductivity. Utilizing this dataset, they derived refined regression equations for thermal conductivity based on density, outlined as follows:

$$K_{\text{eff}} = \begin{cases} 0.023 + 0.234 \cdot \rho_{sno}, & \text{if } \rho_{sno} < 0.156 \\ 0.138 - 1.01 \cdot \rho_{sno} + 3.233 \cdot \rho_{sno}^2, & \text{if } 0.156 \leq \rho_{sno} \leq 0.6 \end{cases} \tag{5.1}$$

where ρ_{sno} is the snow density in gcm^{-3}. Sturm et al. (1997) highlighted six advantages of their dataset compared to others:

- They employed a needle probe and a dynamic test method known for its accuracy with low-conductivity materials (Sturm and Johnson, 1992), which is particularly advantageous for materials prone to metamorphism.
- They presented a unified dataset, where all data were collected using a consistent methodology.
- This dataset includes measurements on standard materials, enabling the assessment of systematic errors and accuracy.
- Measurements were conducted at stable, relatively low temperatures, minimizing sensitivity to temperature dependence.
- The dataset uniquely provides detailed descriptions of the type of snow.
- The dataset encompasses a significantly larger amount of data compared to other combined datasets.

The variance in mean thermal conductivity across different datasets can largely be attributed to variations in the temperature at which the tests were conducted. The mean value in the dataset is -14.6°C, while the authors estimated that other data still in use were collected at an average of approximately -5°C. This discrepancy makes the dataset from Sturm et al. (1997) particularly well-suited for simulating Arctic permafrost regions.

Nonetheless, Sturm et al. (1997) also acknowledged certain limitations in their study. Primarily, the dataset overrepresents two snow types frequently found in high

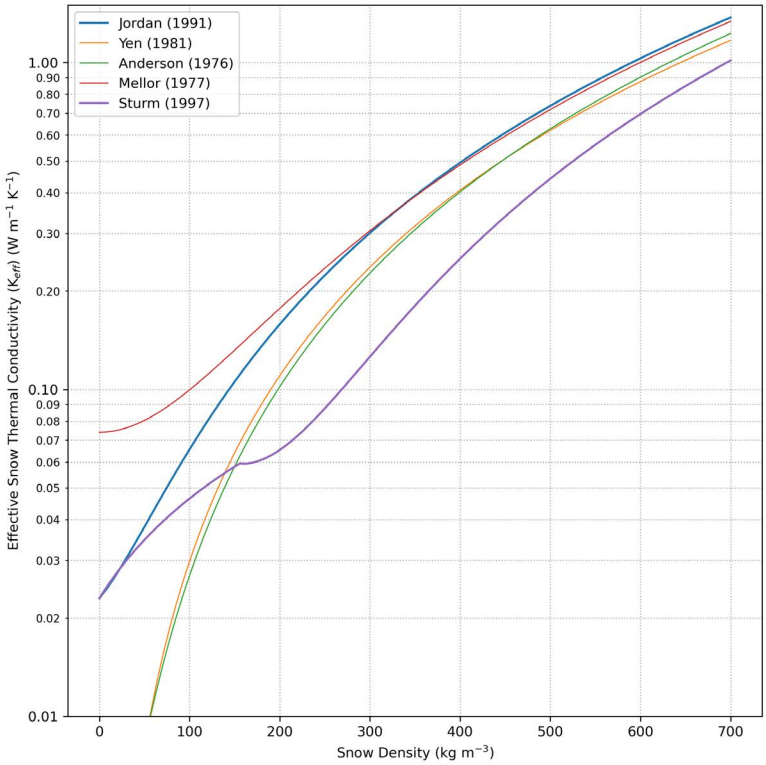


Fig. 5.1 Comparison of four schemes for K_{eff} from 0 to 700 kgm^{-3} for snow density. Note that the y-axis is logarithmic.

latitudes: depth hoar and wind-slab, to the detriment of snowpack that can be found in mountainous areas. Consequently, they advocate for the adoption of distinct schemes based on different locations and snow conditions. Furthermore, other studies have criticized the method used by Sturm et al. (1997) to measure snow thermal conductivity and suggest that more recent methods would be better options (Riche and Schneebeli, 2010; Fourteau et al., 2022).

Sturm et al. (1997) equation is compared to the previous equations discussed in Chapter 2 in Figure 5.1, and extended above the 0.6 gcm^{-3} threshold. As expected, there is a distinct offset between Sturm et al. (1997) and others, with differences of up to a factor of 3 compared to Mellor (1977) and Jordan (1991). Above 150 kgm^{-3}, it is distinguished by its significantly smaller K_{eff}. Its dual-equation system ensures that unrealistically low K_{eff} values are avoided, in cases involving small snowpack densities, unlike Yen (1981) and Anderson (1976).

Dutch et al. (2022) recently applied the alternative formulation proposed by Sturm et al. (1997) in their investigation, which involved a comparative analysis between CLM5 and in-situ measurements acquired in Trail Valley Creek, Northwest Territories, Canada. Their study unveiled that the initial version of CLM5 overestimated snow thermal conductivity by a factor of 3 compared to observations, consequently inducing a cold bias in the wintertime soil temperature simulations. When applying the Sturm

et al. (1997) scheme to CLM5, a significant improvement of wintertime soil temperatures simulations was discerned. This study thereby provides valuable insights into the propensity of most LSMs to neglect the implications of the Sturm et al. (1997) analysis in their framework.

While the contribution by Dutch et al. (2022) constitutes a notable advancement within the realm of snow science, their investigation was limited to a single location assessment. To understand the full potential of applicability of the Sturm et al. (1997) scheme, it is essential to extend their experimentation to a larger regional context. This chapter is dedicated to bridging this knowledge gap by conducting an Arctic domain simulation and sensitivity experiment with CLM5 using the Sturm et al. (1997) snow thermal conductivity scheme.

5.3 Description of snow module in CLM5

The snow module in CLM5, described in van Kampenhout et al. (2017) and Lawrence et al. (2019), includes physical processes such as snow accumulation, compaction (due to overburden pressure and drifting snow), refreezing, melting, and sublimation, allowing for a comprehensive representation of snow's behavior. However, the snow module does not take into account water vapor.

The model uses a multi-layer approach that discretizes the snowpack into a maximum of 12 layers. Vertical discretization is performed in an Eulerian method, employing prescribed maximum (Δz_{max}) and minimum (Δz_{min}) layer thicknesses that adapt based on temperature and grain size gradients, leading to merging or creating layers as needed to maintain these thresholds. Any mass in excess of the snow cap is routed to the river component. Four prognostic variables are represented: snow density, water content, temperature, grain size, and aerosol concentration.

Radiation and atmospheric-related processes are described in van Kampenhout et al. (2017) and Lawrence et al. (2019). Here, I focus on the important processes for a standalone land model simulation.

The thermal state is dependent on the skin temperature, the divergence of penetrated shortwave radiation, and the latent heat released through refreezing, which is determined by three factors: snow depth, thermal capacity, and porosity.

van Kampenhout et al. (2017) introduced a new fresh snow density (ρ_{fs}) parameterizsation by combining a temperature term ρ_T with a linear wind-dependent density term ρ_w (Appendix B, Equations B.8, B.9, B.10).

As stated in Chapter 2, snow can densify via four distinct processes: compaction by overburden pressure, compaction by drifting snow, destructive metamorphism, or melt metamorphism. Furthermore, snow thermal conductivity is solely dependent on snow density and calculated following Jordan (1991) scheme using equation 2.12.

5.4 Sturm experiment with CLM5

In general, improvements to the CLM5 snow module have led to increased snow density across most of the Arctic compared to CLM4.5, as observed by Lawrence et al. (2019) and in our results (Appendix A, Figure A.9). There is a theoretical positive correlation between snow density and thermal conductivity: as snow density increases, the thermal conductivity also tends to increase (Adams and Sato, 1993). Denser snow has fewer air voids, resulting in fewer insulating air pockets. Consequently, heat is transferred more efficiently through the solid ice matrix, leading to higher thermal conductivity.

My hypothesis is that the underestimation of snowpack density by CLM4.5, combined with the high thermal conductivity scheme from Jordan (1991), artificially resulted in an acceptable winter offset over Arctic permafrost regions, as shown in Koven et al. (2013). Meanwhile, the introduction of the new fresh snow density function by van Kampenhout et al. (2017) may have unintended consequences, making the snow too dense in Arctic permafrost regions, particularly because specific tundra snowpack features like depth hoar are not represented by the model (Dutch et al., 2022). As the snow thermal conductivity scheme remains unchanged from CLM4.5 to CLM5, higher snow density means that cold winter conditions can penetrate the snow more effectively. This results in reduced snow insulation and could explain the significant cold bias observed in CLM5 in Chapter 3 and in Dutch et al. (2022).

Unfortunately, it appears unlikely that the issue of better snow density representation in LSMs will be resolved in the near future, and recent developments in this area seem premature (Fourteau et al., 2023). In response, Dutch et al. (2022) suggest using the Sturm et al. (1997) snow thermal conductivity scheme, which is better suited for tundra snowpack properties. This study seeks to expand upon their initial research conducted in the Trail Valley Creek by conducting a sensitivity experiment across the entire Arctic domain.

To achieve this objective, I conducted two simulation: (1) the ”control run”, which shares the configuration discussed and evaluated in Chapter 3, and (2) the ”Sturm run”, where the conventional snow thermal conductivity scheme is replaced with the scheme proposed by Sturm et al. (1997). For practical reasons, I modified the equation 5.1 to compute snow density above 0.6 gcm^{-3}. By analyzing the differences in model outputs between these two runs, I investigate the consequences of adopting this new approach and its implications for LSMs' representation of Arctic permafrost regions.

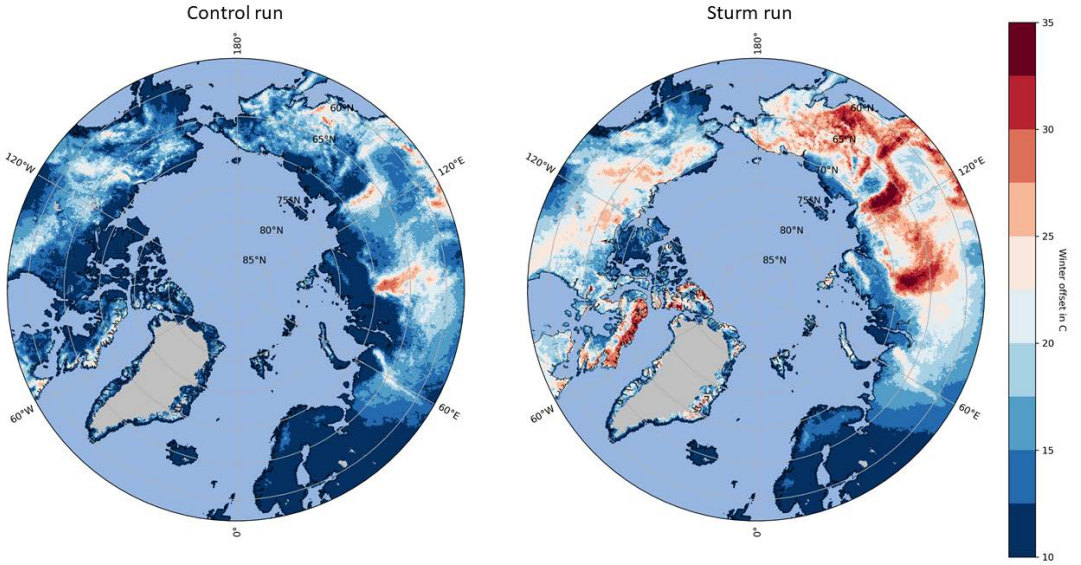


Fig. 5.2 Period (1980 to 2021) winter offset for the control run (left) and Sturm run (right).

5.5 Results

5.5.1 Soil temperature

5.5.1.1 Comparison between the Sturm and control runs

An effective method for assessing the wintertime heat transfer from the atmosphere to the soil involves the computation of the ”winter offset.” I adopted the definition proposed by Burke et al. (2020), wherein the winter offset represents the difference between the mean soil temperature at 0.2 m and the mean air temperature for the December to February period. Burke et al. (2020) argue that this index serves as a valuable tool for gauging the insulating capacity of snow and shows the largest uncertainty in representing permafrost dynamics among CMIP6 models.

Figure 5.2 presents a comparative analysis of the winter offset between the control and Sturm simulations for the 1980-2021 period. As expected, the offset values show a significant increase throughout the entire Arctic domain. Notably, the values ranging from 20 to 30°C are expected from a tundra snowpack (Wang et al., 2016b). This observation is particularly remarkable, considering that there is no difference in effective snow depth between the Sturm and control runs (Appendix A - Figure A.8). This suggests that the snow insulation in the Sturm simulation is considerably diminished and, in general, more realistic.

However, it is crucial to note that the figure also reveals unrealistic offset values, above 30°C, particularly in high-altitude regions. These anomalies are particularly prominent over the central Siberian Plateau, the Verkhoyansk Range, most of Eastern

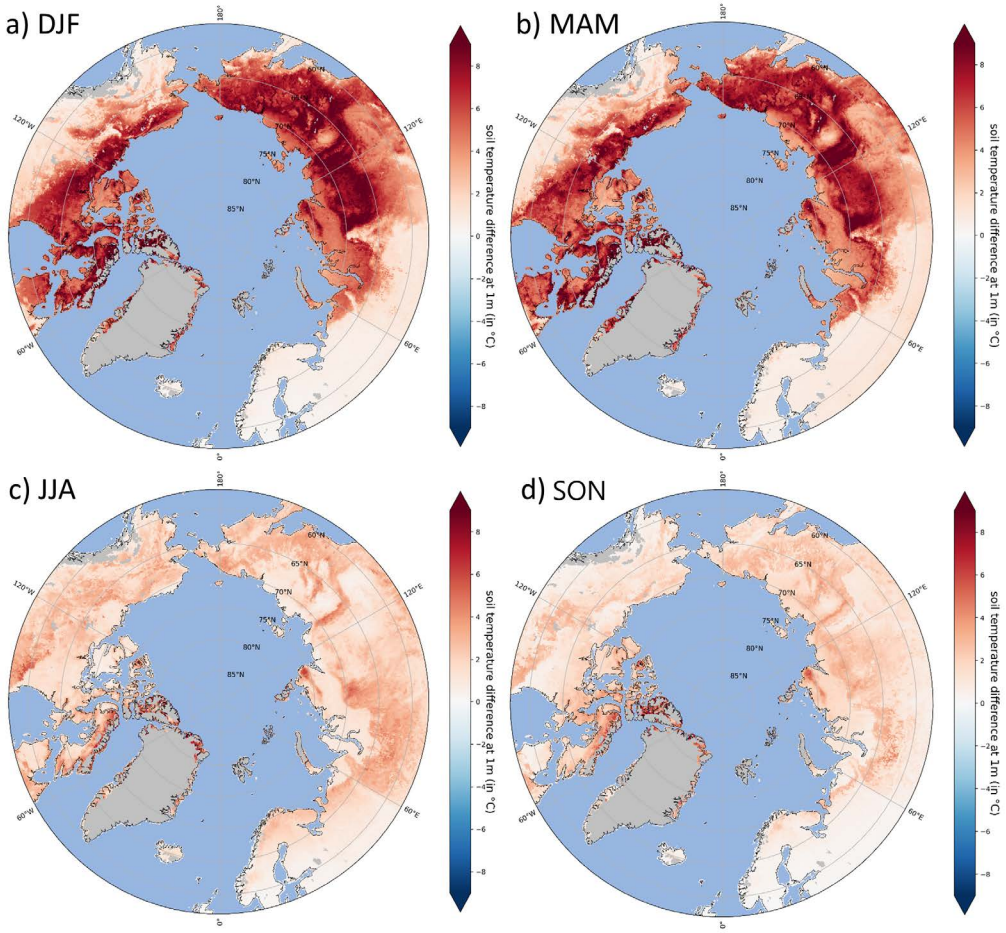


Fig. 5.3 Period averaged (1980-2021) soil temperature difference between the Sturm and control runs at 1m depth in°C for four seasons: a) December, January, February (DJF), b) March, April, May (MAM), c) June, July, August (JJA) and d) September, October, November (SON). Darker red indicates that the experiment is warmer than the control run. The grey mask represents glaciers. Note that the colour scale is not the same for all maps.

Siberia, and the northern regions of the Baffin Island.

Figure 5.3 shows the four seasons averaged difference in soil temperature between the Sturm and control runs for the 1980-2021 period. In winter, there is an extremely significant increase in temperature, from +4 to +10°C over the Siberian permafrost region, and up to 5°C in northern Canada and Alaska. It is especially prevalent in mountainous areas. While in spring, there is an increase up to 3°C is found mostly over high-altitude areas over the whole domain and on the southwestern Hudson Bay. In summer, the increase in temperature is much less marked over the whole domain with an increase of temperature from +1 to +2°C, except for mountains areas and southwestern Hudson Bay. In spring, an increase of up to 3°C is found mostly over mountainous areas across the whole domain and on the south-west of Hudson Bay.

Figure 5.4 presents the year-averaged difference in soil temperature at a depth of 1 meter between the Sturm and the control run. As expected, there is a global temperature

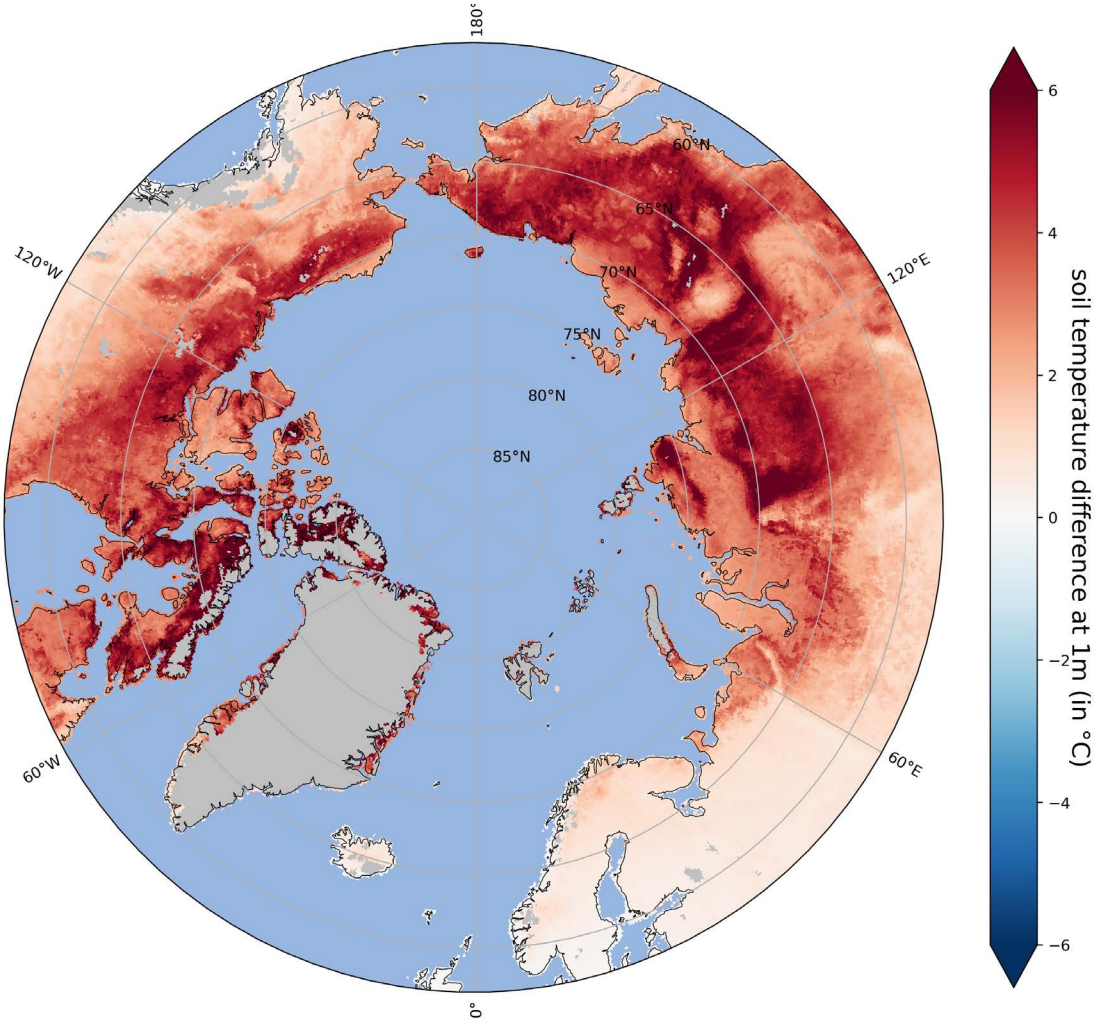


Fig. 5.4 Yearly period averaged (1980-2021) soil temperature difference between the Sturm and control runs at 1m depth in°C. Darker red indicates that the Sturm run is warmer than the control run. The grey mask represent the glaciers.

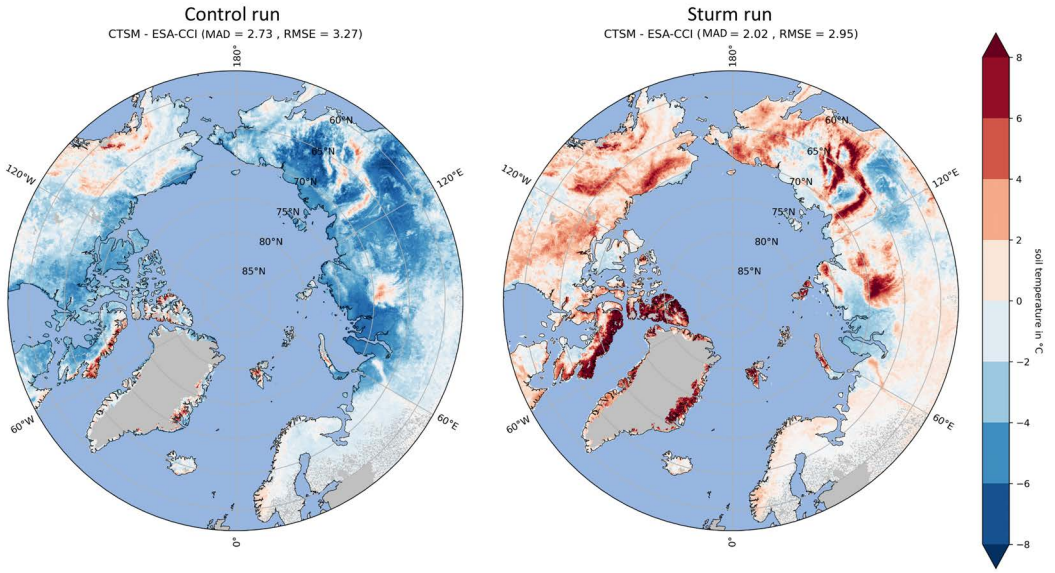


Fig. 5.5 Period (1997 to 2019) MAGT at 1m difference between CTSM and ESA-CCI in°C for the control run (left) and Sturm run (right). Darker blue indicates that CTSM soil temperature is colder than ESA-CCI. ESA-CCI data are aggregated on the CTSM grid using a conservative regridding method.

increase of up to +6°C. The greatest increase is observed over Siberia and high-altitude areas like the Brooks Range and northern Baffin Island. Only the south of Sweden, Finland and West Russia show no significant increase.

5.5.1.2 Comparison between the Sturm run and ESA-CCI

Figure 5.5 compiles an evaluation of the 1-meter soil temperature year-averaged, comparing the results from the control run against the ESA-CCI dataset and the same evaluation with the Sturm run for the period spanning 1997-2019, similar to Fig. 4.5. It is evident that the Sturm run effectively mitigated a significant portion of the cold bias observed in the control run. Most regions only have a small cold bias, typically reaching up to -2°C. This level of bias aligns well with what has been observed in the ESA-CCI evaluation, as noted by Heim et al. (2021).

The MAD shows a noteworthy improvement, decreasing from 2.73 in the control run to 2.02 in the Sturm run. However, the spread of temperature values, represented by the RMSE, only reduced slightly, from 3.27 to 2.95. This is probably linked the pronounced warm bias observed over high-altitude areas, a feature that was present in the control run but was significantly amplified in the Sturm simulation. These high-altitude regions encompass the central Siberian Plateau, the Verkhoyansk Range, most of Eastern Siberia, the northern regions of Baffin Island, and the Brooks Range.

5.5.1.3 Comparison between the Sturm run and the 295GT dataset

Figure 5.6 illustrates the period-averaged monthly soil temperatures at various depths (-20 cm, -80 cm, -160 cm, and -320 cm) for the observations (depicted in black), control run (in blue), and Sturm run (in red) in °C. These values are derived from an average of 295 stations over the entire period (1980-2021), similar to Fig. 4.6.

Notably, the Sturm run effectively minimized the bias gap introduced by the control run, particularly during winter and within the uppermost layers. Nonetheless, the degree of improvement is less pronounced in summer and for deeper layers, as the properties of soil increasingly overlook the insulating properties of snow with depth. Furthermore, there is a significant overshoot observed in the top -20 cm layer during the winter season.

Appendix C presents an additional analysis from ten distinct stations that complements the regional soil temperature in this chapter, using the Sturm run.

5.5.2 Permafrost extent

Similar to the soil temperature in Figure 5.5, Figure 5.7 illustrates the difference in permafrost extent between CLM5 and ESA-CCI remote products, with the control run (left) and the Sturm run (right). Here, the overestimation of permafrost made by the control run has been resolved to the detriment of mountainous regions (in red) that have been reclassified as non-permafrost by the Sturm run. In addition, the Sturm run shows a significant loss of discontinuous permafrost (in yellow). This induces a significant decrease of 9.489 $\times 10^6$ km^2, compared to 12.544 $\times 10^6$ km^2 for ESA-CCI.

5.5.3 Active Layer Thickness (ALT)

5.5.3.1 Comparison between the Sturm and control runs

Figure A.11 in Appendix A, illustrates the difference in period-averaged (1980-2021) active layer thickness (ALT) between the Sturm and control runs. The results show a significant increase in depth of up to +4 m. However, it is important to note that most of these changes occur in mountainous areas, while tundra plains exhibit no significant alterations. Additionally, large portions classified as non-permafrost in the Sturm run are not displayed on this map.

5.5.3.2 Comparison between the Sturm run and ESA-CCI

In alignment with the soil temperature data presented in Figure 5.5, Figure 5.8 illustrates the differences in ALT between CLM5 and ESA-CCI remote products. This comparative analysis encompasses both the control run (left) and the Sturm run (right).

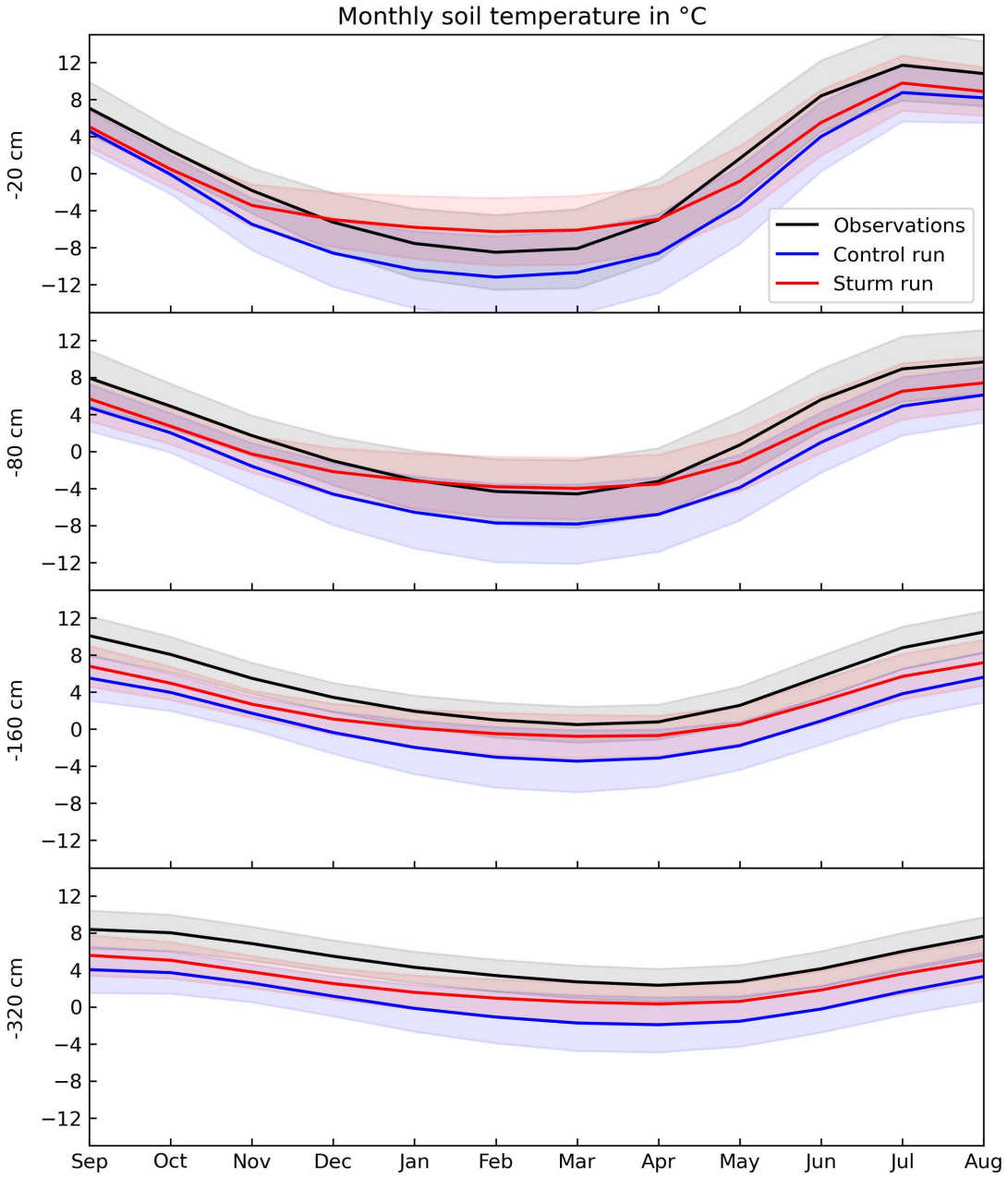


Fig. 5.6 Period averaged (1980-2021) of monthly soil temperature at 4 different depths (-20, -80, -160 and -320 cm) for the observations (black), control run (blue) and Sturm run (red) in °C. The area represents the RMSE over all years. All values come from an average of the 295 stations through the full period.

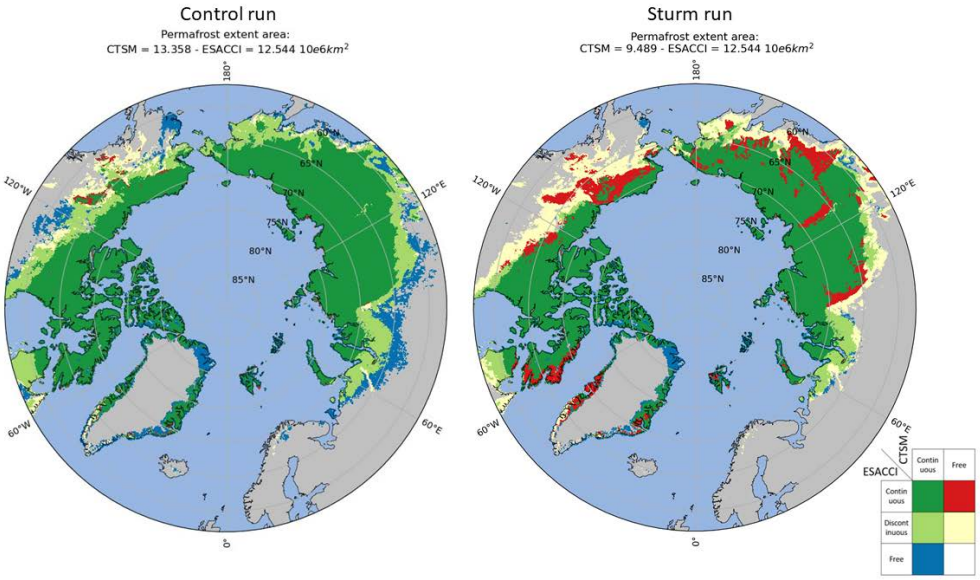


Fig. 5.7 Permafrost extent area mask difference between CTSM and ESA-CCI for the control run (left) and Sturm run (right). Refer to the legend for colour meanings. ESA-CCI data are aggregated on the CTSM grid using a conservative regridding method.

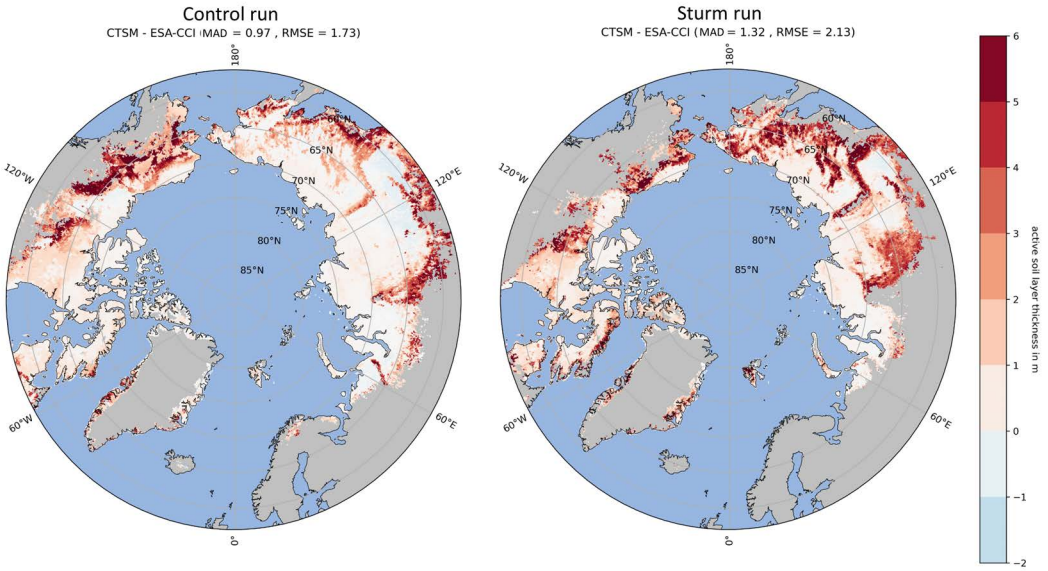


Fig. 5.8 Active layer thickness difference between CTSM and ESA-CCI in meters for the control run (left) and Sturm run (right). Darker red colour indicates that CTSM ALT is deeper than ESA-CCI. ESA-CCI data are aggregated on the CTSM grid using a conservative regridding method.

The results indicate a significant bias increase in mountainous areas, while the bias in plain tundra regions remains relatively similar. MAD and RMSE scores shift from 0.97 to 1.32 and from 1.73 to 2.13, respectively. Again, these statistics do not consider large regions classified as non-permafrost in the Sturm run.

5.6 Discussion and conclusions

Snow plays a critical role over permafrost regions, providing thermal insulation during winter which has substantial implications for heat exchange between the atmosphere and the soil. My investigation delved into the complexities of snow thermal conductivity, an essential parameter for understanding snow's insulation capabilities.

The study discussed the limitations of existing models in adequately representing the thermal properties of Arctic snow, highlighting the discrepancies between model simulations and observational data. I drew attention to the pioneering work of Sturm et al. (1997) and their comprehensive dataset, which provides underestimated insights into the relationship between snow density and thermal conductivity. I suggest that they have provided a snow thermal conductivity scheme under-utilized by state-of-the-art LSMs.

Building upon Dutch et al. (2022)'s site experiment in Trail Valley Creek , this chapter applies Sturm et al. (1997)'s snow thermal conductivity scheme to the entire Arctic domain, as it is better suited to the snow density profile found over Arctic permafrost regions. My aim was to study the impact of this scheme on simulated soil temperatures and permafrost dynamics, thereby improving the model's performance in reproducing snow physics over Arctic permafrost regions.

Comparing the Sturm simulation to the control run, I observed a substantial increase in soil temperature throughout the entire Arctic domain, particularly in winter and spring. This outcome aligns with my hypothesis that the diminished snow insulation in the Sturm run would result in higher winter soil temperatures. However, similar to observations in Chapter 4, this effect becomes less pronounced at greater depths, as the properties of soil increasingly overlook the insulating properties of snow. Furthermore, results show unrealistic offset values in high-altitude regions.

Comparing the Sturm run to the observation datasets reveals a significant mitigation of the cold bias observed in the control run. In most regions, the Sturm run shows only a small cold bias, generally up to -2°C; however, there is a distinct warm bias of up to +4°C over some regions, mostly over mountainous areas. Notably, the winter cold bias over Siberia has been significantly reduced and the global mean absolute deviation (MAD) has improved (from 2.73 to 2.02). However, the spread of temperature values (RMSE) remains relatively constant (from 3.27 to 2.95), probably due to the strong observed positive bias in high-altitude regions.

Regarding permafrost extent, the Sturm experiment addressed the overestimation of permafrost present in the control run over southern Siberia. However, it is to the detriment of many regions which are no longer considered permafrost in the domain, most of which are located over mountainous areas at the southern borders of the domain. Additionally, there is a notable reduction in the extent of discontinuous permafrost. Consequently, the total permafrost extent area decreases from 13.358 to 9.489 $\times 10^6$ km^2, in contrast to the 12.544 $\times 10^6$ km^2 extent as estimated by ESA-CCI.

In terms of active layer thickness (ALT), the comparative analysis between the Sturm and control runs show similar bias over flat areas. However, the bias in the Sturm run is significantly amplified over mountainous regions. Overall, the mean regional bias and its variability slightly increased (from 0.97 MAD and 1.73 RMSE to 1.32 MAD and 2.13 RMSE), with the caveat that large regions designated as non-permafrost were not included in these calculations.

The distinct features observed over mountainous areas can be attributed to two main factors. Firstly, the comparisons between ESA-CCI and the control run highlight a significant difference between low and high-altitude regions, with the latter showing either no bias or a warm bias. This discrepancy arises from variations in atmospheric forcing resolution between CLM5 and ESA-CCI, as discussed in Chapter 3. Removing the "cold mask" in the Sturm simulations amplified these temperature differences. Secondly, the Sturm scheme is recognized to be less suited for mountainous regions (Sturm et al., 1997; Dutch et al., 2022).

While the Jordan (1991) scheme might improve results over regions with high-density snowpacks, it is important to note that the column-averaged snow density represented by the model does not differ between the mountain and tundra regions (Appendix A - Figure A.10). This implies that the primary factor contributing to these distinct features is the resolution-dependent atmospheric forcings. However, I believe that a significant portion of this warm bias observed over mountainous areas could be resolved through a more accurate representation of high-density mountain snowpacks. Such an enhancement would involve the utilization of different snow schemes tailored to specific types of snow (e.g., tundra or alpine) and require further research.

In summary, the integration of the Sturm et al. (1997) snow thermal conductivity scheme within CLM5 has led to significant enhancements in the modeling of soil temperature in tundra regions. These improvements are especially notable in reducing cold biases and aligning model outputs more closely with observations datasets. However, this enhancement has the unintended result of misclassifying significant mountainous areas as non-permafrost. These findings underscore the importance of refining snow-related processes in LSMs to enhance our understanding of permafrost dynamics in the context of climate change.

This study highlights potential avenues for future research. One notable limitation

in my study, as in many others, is the exclusive focus on presenting thermal conductivity solely as a function of snow density, neglecting other intrinsic snow characteristics such as grain size, bonding, and temperature, which can also significantly influence thermal conductivity (Sturm et al., 1997). Barrere et al. (2017) demonstrates that relying solely on density to determine snow thermal conductivity is particularly inadequate in the Arctic region. Another key limitation is my focus on thermal conductivity without delving into the densification processes and parameters within the model, even though snow thermal conductivity is entirely dependent on snow density. Figure A.10 (Appendix A) reveals a relatively uniform distribution of snow density across the domain. This contradicts real-life observations where mountain snowpack typically exhibits higher density (Zhao et al., 2023) compared to tundra snowpack. Scholars have argued that achieving reliable permafrost modeling is not feasible in the absence of a proper representation of snow density (Langer et al., 2013; Dutch et al., 2022).

CHAPTER 6

Conclusion

6.1 Introduction

Arctic permafrost regions are critical for anthropogenic climate change and constitute a central focus within climate science (Overland et al., 2019). Permafrost thaw and its subsequent greenhouse gas emissions profoundly influence various societal, ecological, hydrological, and biogeochemical processes in the Arctic (Ramage et al., 2021; Yang et al., 2010; Hinzman et al., 2013) and stand as one of the most significant and uncertain positive feedback loops in the context of climate change (Schuur et al., 2008; Schaefer et al., 2014; Schuur et al., 2015; Biskaborn et al., 2019). For each additional 1°C of warming, the global volume of permafrost is projected to decrease by about 25% relative to the present volume. However, these decreases may be underestimated due to an incomplete representation of relevant physical processes in ESMs (IPCC, 2021). This underscores the urgent need for providing accurate predictions concerning changes in permafrost coverage in response to future global climate change (Miner et al., 2022).

To address these critical questions and challenges, the development of LSMs began incorporating physical soil processes, enabling us to explore the complex dynamics of permafrost and its interactions with the broader climate system.

This thesis is dedicated to enhancing our comprehension and developing precise representations of permafrost dynamics within LSMs, with a particular emphasis on accurate modeling of heat and water fluxes, which are crucial components for simulating permafrost physics (Zhu et al., 2019; Yang et al., 2022). At the outset of this study, I formulated four research questions to set the framework for my investigation into fundamental model prerequisites, model performance evaluations, the influence of soil properties, and the critical role of snow thermal conductivity in permafrost physics. The following sections summarize the conclusions drawn from each research question.

6.2 Research question 1

Which physical processes and variables determine a model's capability to represent permafrost soils? Which methods do land surface models use to simulate heat

and water transfer in permafrost regions?

Chapter 2 summarizes the essential requirements for effectively representing permafrost soils within LSMs. In addition, this study is the first review within the literature to make a comparison between LSMs that participated in CMIP6 to reveal their difference, performance, and shortcomings on soil physics, soil hydrology, snow physics and vegetation representation.

This study underscores the profound importance of factors like soil discretization, thermal conductivity, hydrology schemes, critical snow physics parameters (e.g. layering, densification, water phases, and thermal conductivity), and vegetation representation in representing permafrost.

While each LSM has a distinct approach to these processes, selecting the most suitable model necessitates careful consideration of spatial and temporal scales, research objectives, and available computational resources. Importantly, the divergence in model outcomes emphasizes the need for a holistic and synthetic evaluation of various aspects, highlighting the inherent complexity of permafrost dynamics.

6.3 Research question 2

How accurately does the Community Land Model (CLM5) simulate permafrost and soil temperature dynamics in the Arctic region, in comparison to observed data on permafrost depth and temperature profiles? In particular, to what extent can CLM5 effectively capture the spatial and seasonal variabilities inherent in permafrost physics across the Arctic region?

The evaluation of CLM5 provides valuable insights into simulating Arctic permafrost regions. My approach overcomes traditional evaluation limitations by individually addressing depth, seasonality, and regional variations, providing a comprehensive assessment of permafrost and soil temperature dynamics. I compare CLM5's results with three extensive datasets: (a) soil temperatures from 295 borehole stations, (b) active layer thickness (ALT) from Circumpolar Active Layer Monitoring Network (CALM) data, and (c) soil temperatures, ALT, and permafrost extent against the ESA Climate Change Initiative (ESA-CCI).

Comparisons with both borehole station datasets and ESA-CCI reveal a strong cold temperature bias across the majority of the domain. These findings align with a persistent challenge identified in numerous studies: a systematic "cold bias" in soil temperature over Arctic permafrost regions. Concerning permafrost extent and ALT, the results indicate that CLM5 aligns closely with ESA-CCI and CALM, albeit with a substantial overestimation of ALT over warm Arctic permafrost regions. One significant limitation of CLM5 is its inability to accurately capture the seasonality of soil temperature and inter-annual variability of ALT.

To promote the ease of use and foster future research in the field, all algorithms employed in this evaluation are publicly accessible and documented online: `https://github.com/AdrienDams/cegio`.

6.4 Research question 3

What is the impact of using a plant functional type (PFT)-based approach to derive soil texture and soil organic matter in CLM5 on permafrost dynamics, and how does it compare to traditional approaches that use fixed soil properties?

In the context of permafrost physics modeling, soil texture and soil organic matter (SOM) are fundamental parameters for representing soil physics and thermal dynamics. This work is the first to depart from the conventional use of coarse-resolution global data in LSMs by applying a Plant Functional Type (PFT)-based approach to derive soil texture and SOM.

This novel method results in a more uniform distribution of SOM across the domain, marked by reduced SOM values in most regions. However, changes in soil texture exhibit a more intricate spatial pattern. Comparing these results to observations reveals a significant reduction in the cold bias observed in the control run. This reduction is attributed to the increase in sand content observed in some western regions and the regional decrease in SOM, particularly during the summer season. During the winter season, no clear relationship can be established, primarily due to the insulating properties of the snow cover, which dampen thermal dynamics (Zhu et al., 2019).

This method provided significant improvements in permafrost extent with the trade-off of an overestimation in ALT. This increase can be attributed to the strong dependence of ALT on the MaxAGT. As the Obu experiment results in higher summer temperatures, it directly influences the maximum thaw depth, resulting in increased ALT values. Nevertheless, it is crucial to emphasize that the influence of these ALT changes remains limited when juxtaposed with the comprehensive enhancements in soil temperature and permafrost extent achieved in the Obu run.

6.5 Research question 4

To what extent can modifications to the parameterization of snow thermal conductivity improve the performance of CLM5 in reproducing snow cover dynamics over Arctic permafrost regions?

Inspired by the site experiment conducted by Dutch et al. (2022) in Trail Valley Creek, I implemented the snow thermal conductivity scheme initially proposed by Sturm et al. (1997) across the Arctic region. This adaptation aims to use a snow scheme that aligns more effectively with the observed snow density profile characteristic of

Arctic permafrost regions. My objective is to investigate implications of this modification on simulated soil temperatures, with the ultimate goal of improving the model's accuracy in replicating variations in snow cover dynamics over Arctic permafrost areas.

The outcomes of this experiment reveal a notable alleviation of the cold bias previously identified in the control run. In most regions, the Sturm run displays only a modest cold bias. However, certain areas exhibit a distinctive warm bias, particularly those with high-elevation terrain.

Regarding the representation of permafrost extent, this experiment effectively addresses the overestimation of permafrost observed in the control run, particularly in southern Siberia, to the detriment of many high-altitude regions which are no longer considered permafrost in the domain and a loss of discontinuous permafrost. In addition, we observed a significant increase in ALT bias, primarily in mountainous areas.

I attribute the bias observed over high-altitude regions to two factors: (1) differences in the resolution of the atmospheric forcing data used between ESA-CCI and CLM5 and (2) the newly implemented snow scheme may not be ideally suited for mountainous regions (Sturm et al., 1997; Dutch et al., 2022).

These findings underscore the critical necessity of investigating the treatment of snow-related processes within LSMs, thereby advancing our comprehension of permafrost dynamics within the broader context of climate change.

6.6 Outlook

This section highlights potential avenues for future research:

- As discussed in Chapters 4 and 5, my experiments primarily focused on the physics near the surface of the soil column, specifically addressing SOM within the top meter of the soil column and snow physics. Consequently, I did not specifically address the cold bias in the deeper layers, where soil properties gradually dominate over the insulating effects of SOM and snow. Addressing biases in these deeper layers would require further research. One potential solution may involve investigating various soil physics parameters examined in Chapter 2, such as the Wang et al. (2016a) scheme used in ORCHIDEE. This model, being the only one that did not exhibit a cold bias in its evaluation (Guimberteau et al., 2018), warrants further investigation to uncover the reasons behind this unique characteristic.

- One aspect that was not thoroughly examined in the two sensitivity experiments was the issue of poor seasonality discussed in Chapter 3. Exploring how these experiments address the challenge of poor seasonality could be an intriguing avenue for future research.

- The Obu experiment (RQ3) introduced a novel approach by simultaneously altering both soil texture and soil organic matter (SOM) in the model. While this approach revealed intriguing insights into their combined influence, it also posed a challenge in understanding the exact impact of each factor individually. Future research could delve deeper into isolating the effects of soil texture and SOM on permafrost dynamics. This dissection could provide a more nuanced understanding of their respective roles and guide model improvements tailored to each component.

- While my exploration of snow thermal conductivity (RQ4) marked a significant improvement, it was limited by focusing solely on snow density as the key parameter influencing heat transfer. In the future, it would be advantageous to broaden the scope of this parameterization. Incorporating additional factors such as grain size, bonding, and temperature, could lead to a more accurate representation of snow's thermal properties (Sturm et al., 1997).

- While the Sturm et al. (1997) scheme proved to be effective in many regions, its applicability diminishes in mountainous areas and may prove impractical for different types of snowpack, such as those found in ice sheets. To address this challenge, the development of a snow model that incorporates various snow thermal conductivity schemes based on snow types appears promising. However, this goal can only be achieved if the snow model is capable of accurately representing diverse snow types. As observed, the distribution made by CLM5 of column-averaged snow density across the domain appears relatively uniform, which diverges from real-world observations where mountain snowpack typically exhibits higher density compared to tundra snowpack (Zhao et al., 2023). Consequently, achieving reliable permafrost modeling is unattainable without an accurate representation of snow density (Langer et al., 2013; Dutch et al., 2022). Nevertheless, recent efforts to solve this issue appear promising (Fourteau et al., 2023).

- Both the RQ3 and RQ4 experiments have yielded valuable insights, but neither is perfect on its own. As soil properties, snow characteristics, and permafrost dynamics are interdependent, future research could explore the potential benefits of combining these two approaches. For example, Figure 6.1 compares the in-situ stations evaluation between the control, Obu, and Sturm runs for different depths and seasons. Obu and Sturm runs are more prominent in summer and winter, respectively. A unified run incorporating both experiments could provide a more holistic perspective and improve the overall accuracy of LSMs.

- Finally, my simulations were limited to the use of a stand-alone LSM. It would be highly valuable to investigate the impact of the two sensitivity experiments

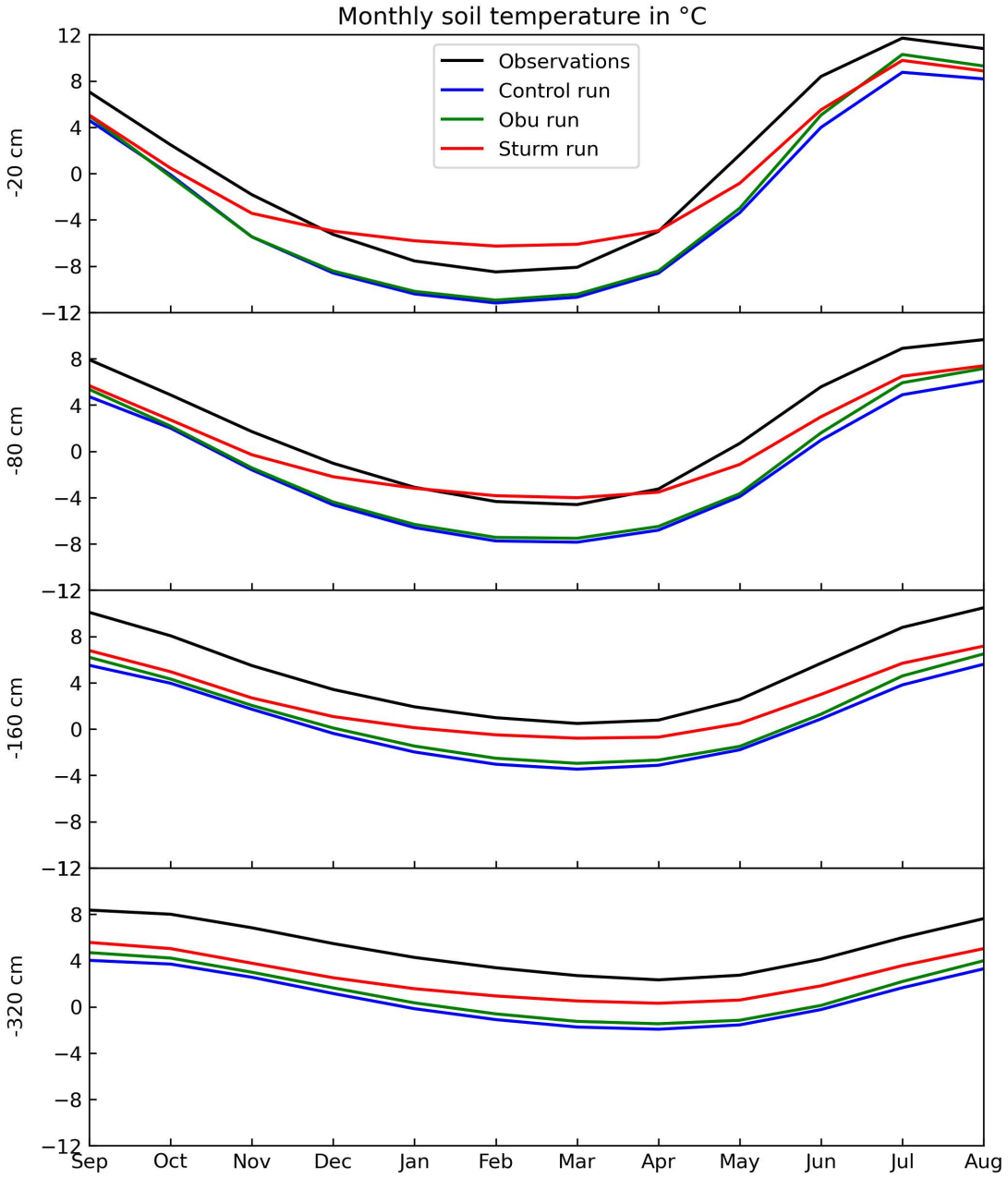


Fig. 6.1 Period averaged (1980-2021) of monthly soil temperature at 4 different depths (-20, -80, -160 and -320 cm) for the observations (black), control run (blue), Obu run (green), and Sturm run (red) in°C. The area represents the RMSE over all years. All values come from an average of the 278 stations through the full period.

proposed here within a fully-coupled ESM. This approach will enable us to gain insights into the interdependencies between the soil and snow with the atmosphere (e.g. the early snowmelt in the Sturm run could influence albedo and, consequently, have a significant impact on atmospheric conditions), the vegetation (e.g. mitigating extreme cold temperatures should improve vegetation survivability and the impact on transpiration [Oogathoo et al. 2022]) and the carbon cycle (e.g., underestimated soil temperatures may lead to an underestimation of the maximum permafrost thaw).

6.7 Conclusion

In summary, this research emphasizes the critical importance of judiciously selecting climate parameters, with a particular focus on soil texture, SOC, and snow thermal conductivity. My results emphasize the crucial roles these parameters play in controlling heat transfer processes and shaping the seasonal variations of soil temperatures as represented by LSMs in Arctic permafrost regions. This thesis represents a significant stride forward in our understanding of permafrost dynamics and offers fresh perspectives on its representation within LSMs.

As we move forward with this research, it is essential to acknowledge the broader implications of my work. The accurate representation of permafrost dynamics by LSMs is not only crucial for understanding the climate system but also for predicting the behavior of permafrost and its potential response to climate change. By bridging the gap between theoretical models and real-world observations, this thesis contributes to the advancement of climate science and our ability to make informed predictions about our changing environment.

Appendix A

Additional figures

A.1 Global Soil Organic Carbon Map

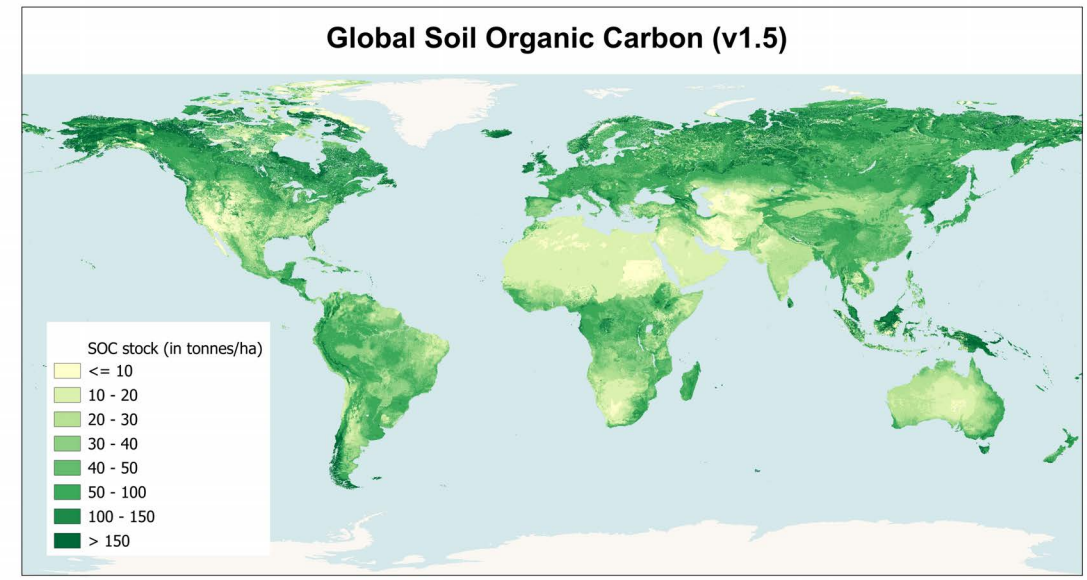


Fig. A.1 Global soil organic carbon in tonnes/ha derived from FAO and ITPS (2020). Darker indicates a higher SOM content per hectare.

A.2 Active Layer Thickness

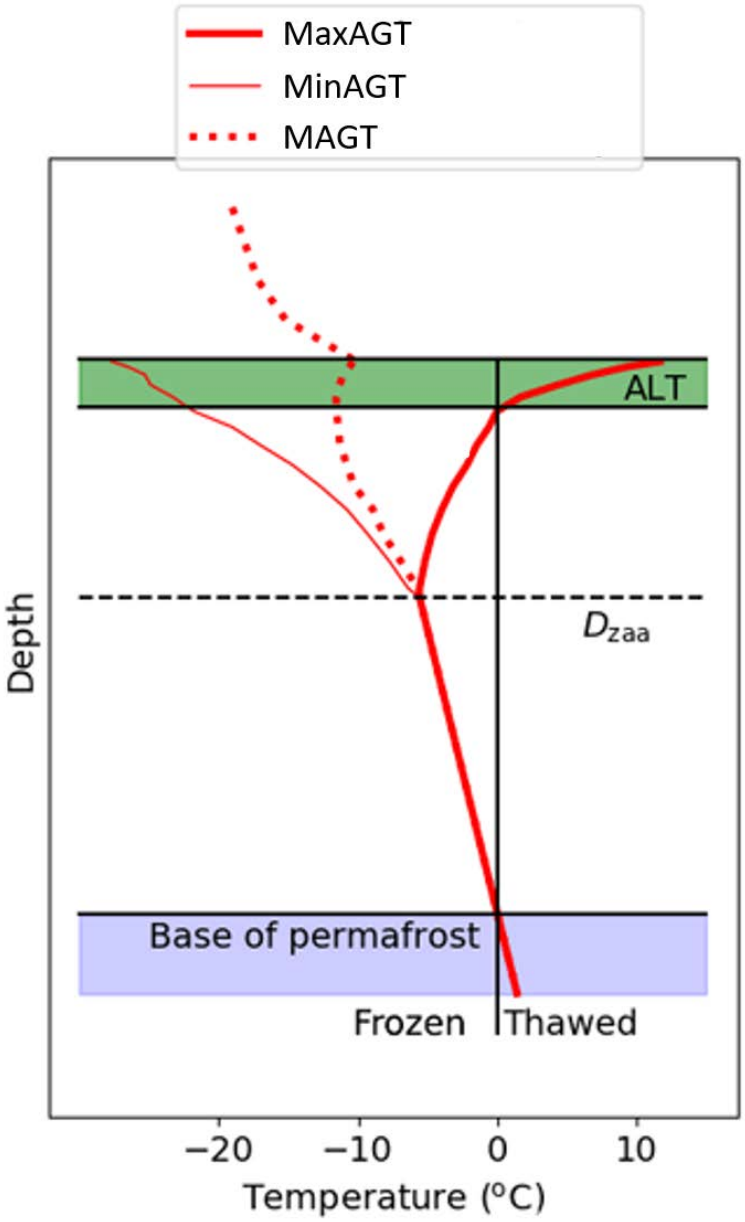


Fig. A.2 Schematic of the mean, minimum and maximum annual temperature profile from the surface boundary layer to below the bottom of the permafrost. MaxAGT is the maximum annual ground surface temperature, MinAGT is the minimum annual ground temperature, MAGT is the mean annual ground temperature, ALT is the active layer thickness, and D_{zaa} is the depth of zero annual amplitude of ground temperature. Modified from Burke et al. (2020).

A.3 Thermal conductivity vs. snow density for four schemes

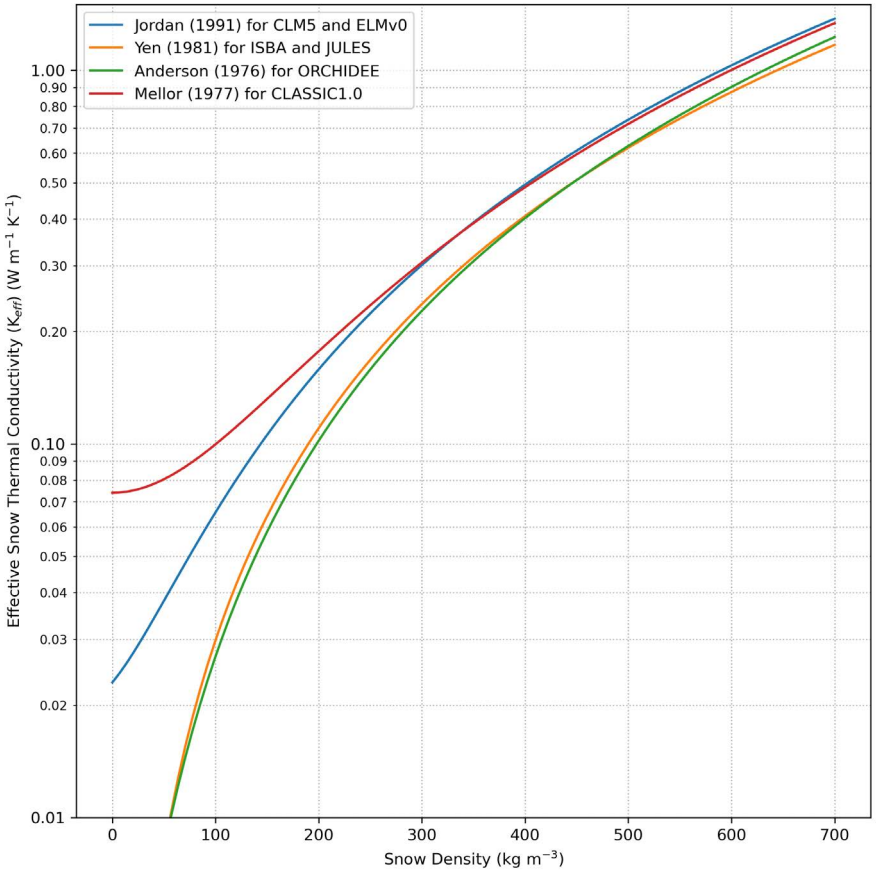


Fig. A.3 Comparison of four schemes for K_{eff} from 0 to 700 kgm^{-3} for snow density. Note that the y-axis is logarithmic.

A.4 CLM5 subgrid hierarchy

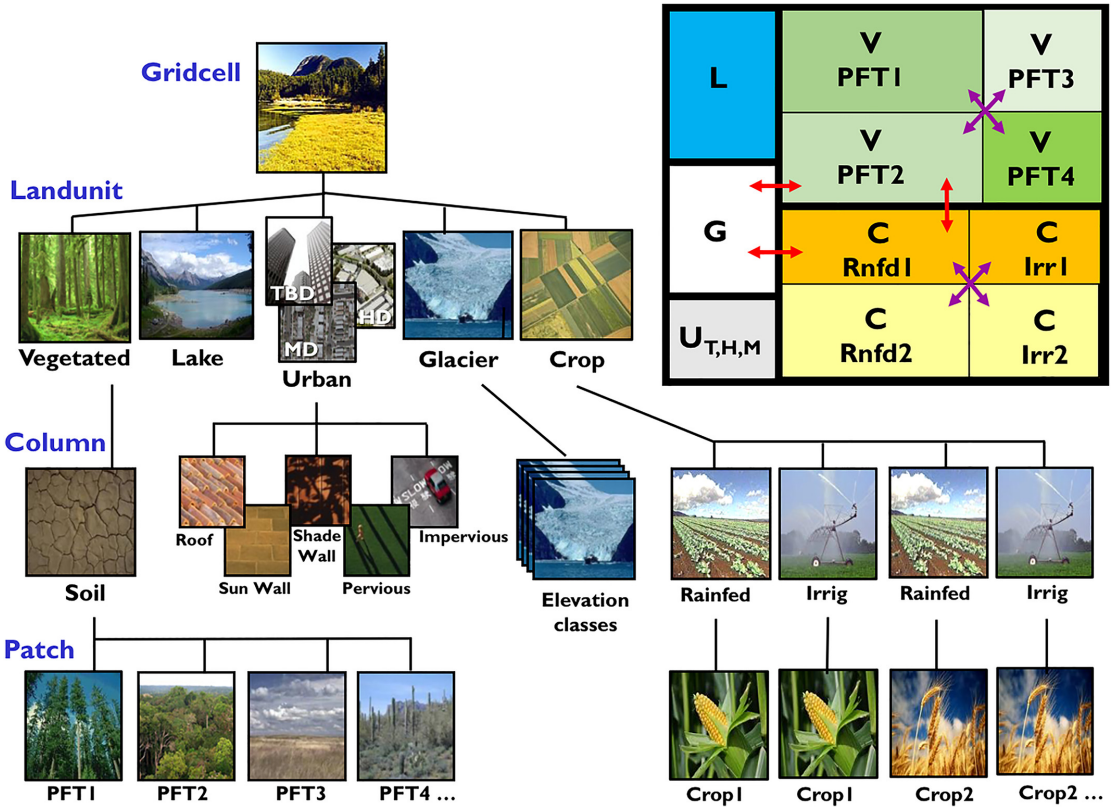


Fig. A.4 Standard configuration of the CLM5 subgrid hierarchy. The upper right box shows hypothetical subgrid distribution for a single grid cell. Note that the crop land unit is only used when the model is run with the crop model active. TBD = tall building district; HD = high density; MD = medium density; G = glacier; L = lake; U = urban; C = crop; V = vegetated; PFT = plant functional type; Irr = irrigated; Rnfd = rainfed. Red arrows indicate allowed land unit transitions. Purple arrows indicate allowed patch-level transitions (adapted from Lawrence et al. 2019).

A.5 Spin-up results of control run

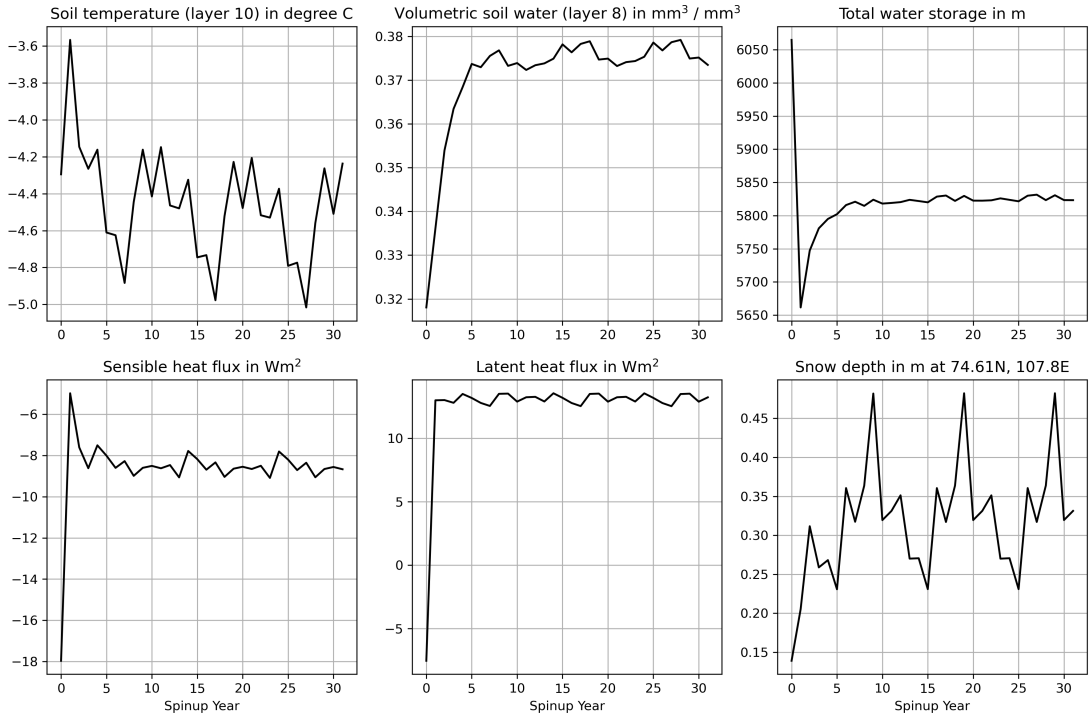


Fig. A.5 Results from the spin-up used in the control run of soil temperature at the tenth layer in °C (upper left), volumetric soil water at the eight layer in mm^3/mm^3 (upper middle), total water storage in meters (upper right), sensible heat flux in Wm^2 (bottom left), latent heat flux in Wm^2 (bottom middle), and snow depth at a location in Taymir - 74.61N, 107.8E (bottom right). All results represent a spatial average of all grid points in the domain.

A.6 Comparison of ALT between the Obu and control runs

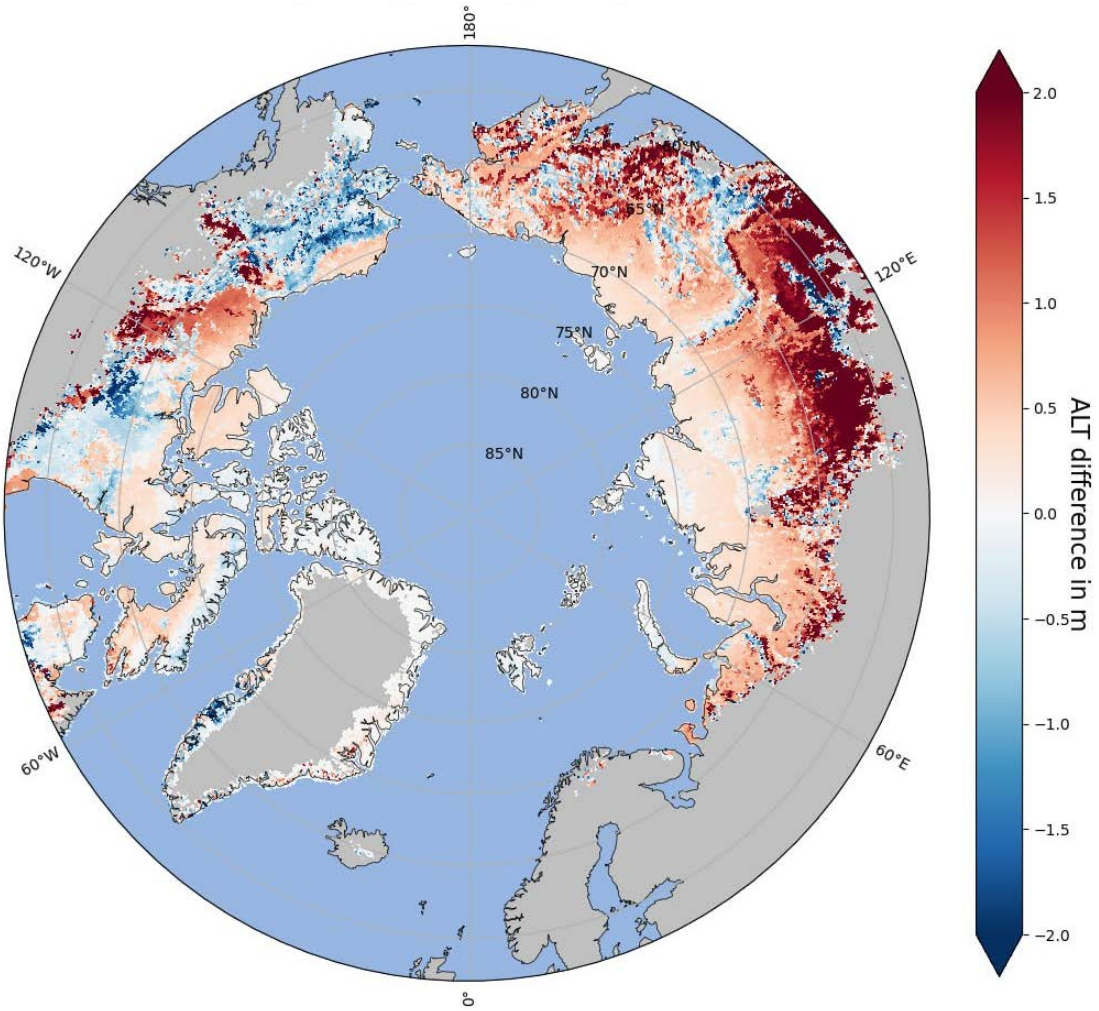


Fig. A.6 Period averaged (1980-2021) active layer thickness difference between the Obu and control runs in meters. Darker red indicates that the Obu run is deeper than the control run.

A.7 Soil liquid and ice water difference between the Obu and control runs

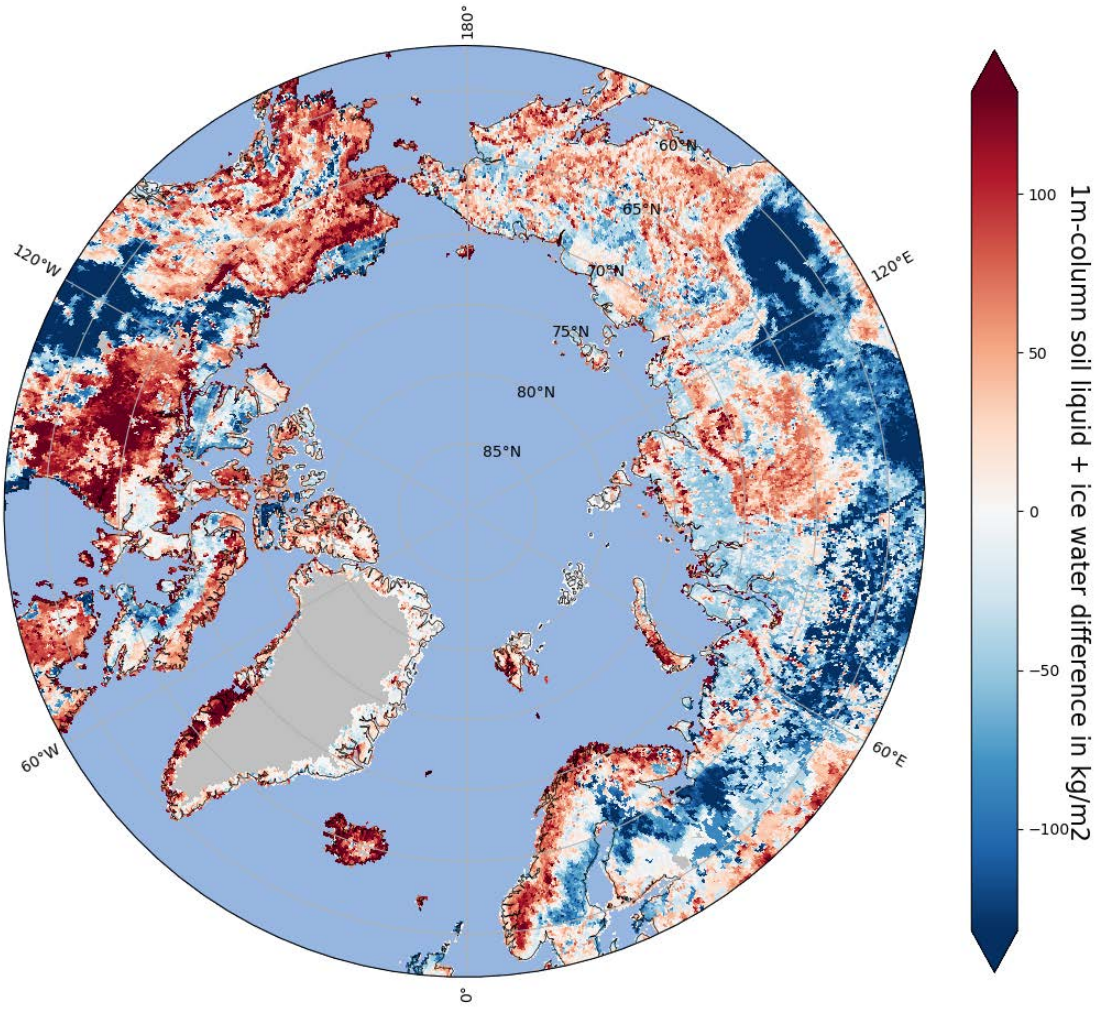


Fig. A.7 Yearly period averaged (1980-2021) 1m-column soil liquid + ice water difference between the Obu and control runs for July in kg/m^2.

A.8 Effective snow depth in the Sturm and control runs

Figure A.8 shows the period-average effective snow depth of the control and Sturm run. The effective snow depth characterizes the insulation provided by snow during the cold period (Burke et al., 2020). $S_{\text{depth,eff}}$ is a cumulative value where the average snow depth in each month, denoted as S_m in meters, is adjusted according to its duration:

$$S_{\text{depth,eff}} = \frac{\sum_{m=1}^{M} S_m(M+1-m)}{\sum_{m=1}^{M} m} \tag{A.1}$$

Snow can be present anytime from October (m = 1) to March (m = 6) with the maximum duration, M, being 6 months. his weighting approach favors early snowfall over late snowfall, as it contributes more to the overall insulating effect. When the effective snow depth, $S_{\text{depth,eff}}$, surpasses 0.25 meters, the insulating capacity of the snow remains relatively constant (Burke et al., 2020), and seasons with earlier snowfall typically exhibit higher $S_{\text{depth,eff}}$ than seasons with later snowfall.

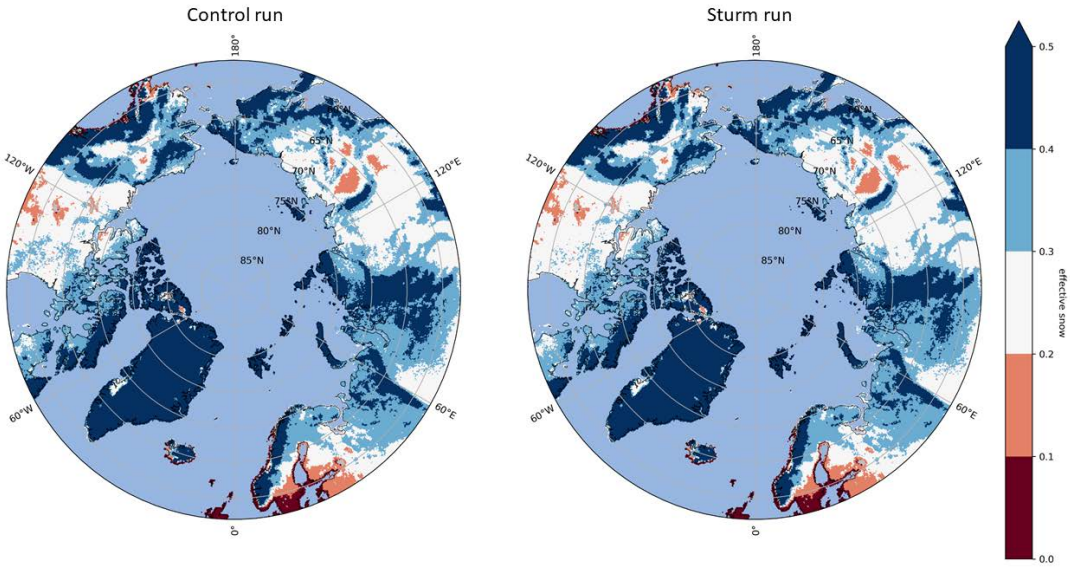


Fig. A.8 Period (1980 to 2021) average of effective snow for the control run (left) and Sturm run (right).

A.9 Snow density in CLM4.5 and CLM5

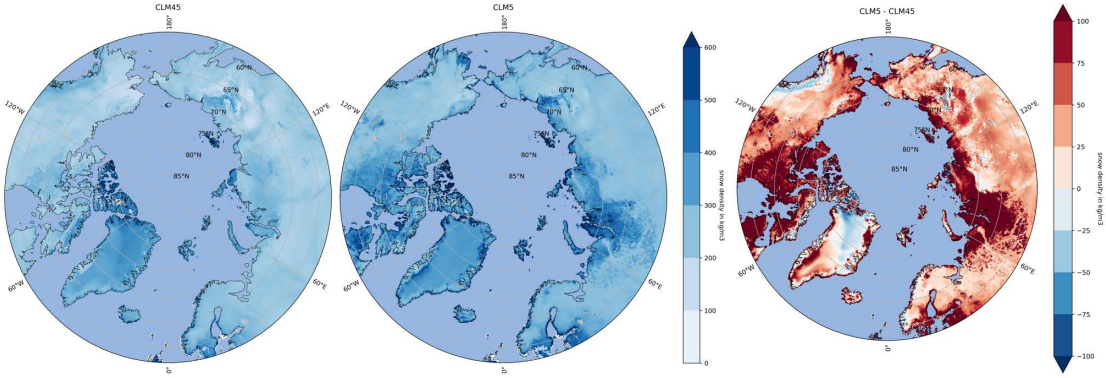


Fig. A.9 Monthly average of snow density (column-averaged) for January 2000 in kgm^{-3} for the CLM45 (left), CLM5 (center), and their difference (right). Darker read indicates that the snow density is higher in Sturm.

A.10 Snow density in the Sturm and control runs

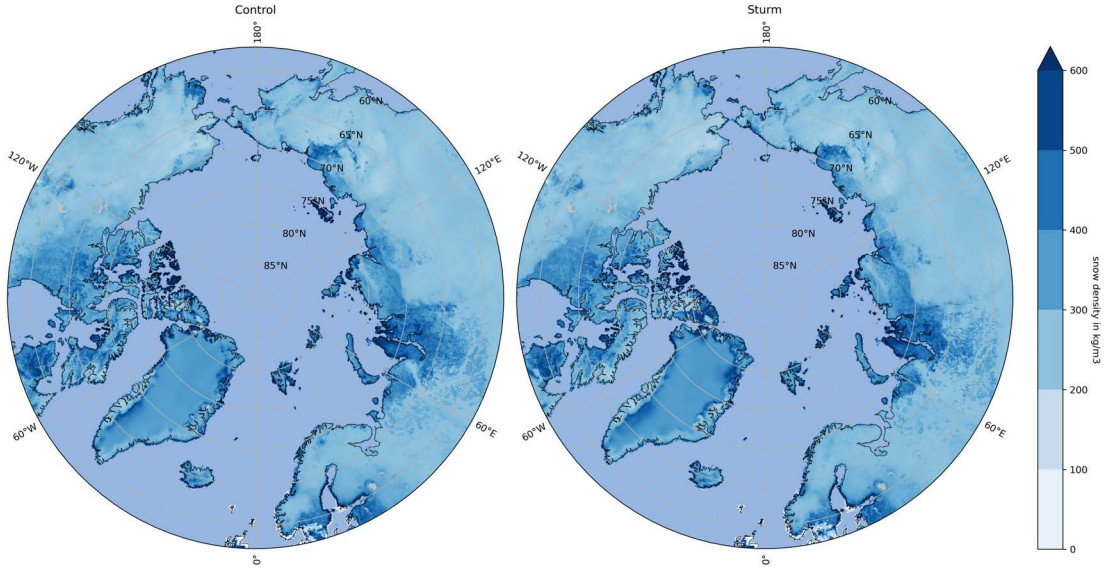


Fig. A.10 Period (1980 to 2021) monthly average of snow density (column-averaged) in January in kgm^{-3} for the control run (left) and Sturm run (right).

A.11 Comparison of ALT between the Sturm and control runs

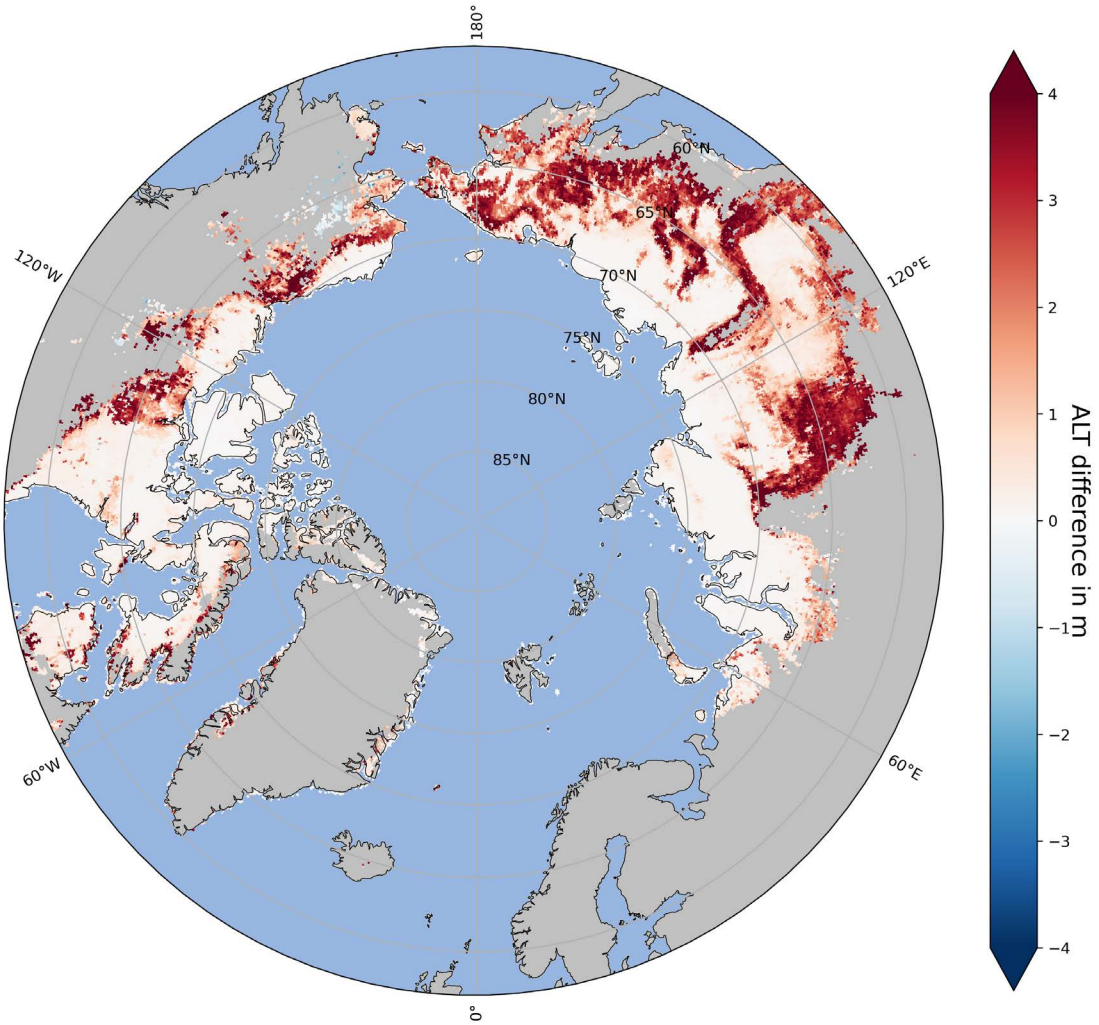


Fig. A.11 Period averaged (1980-2021) active layer thickness difference between the Sturm and control runs in meters. Darker red indicates that the Sturm run is deeper than the control run.

Appendix B

Additional equations

B.1 Particle density

The particle density is a fundamental soil physical property. Particle density is defined as the mass of soil particles divided by the volume occupied by the solids (i.e., excluding voids and water). Typical values for soils range from 2.5–2.8 Mg/m^3 (Warrick, 2002). The bulk density ρ_b is the state of compaction and the amount of pore space in a soil and can be derived as:

$$\rho_b = \frac{\text{mass of solids}}{\text{mass of solids and pore spaces}} \tag{B.1}$$

The solids are all soil particles, where as the pore spaces are spaces holding either air, water or other liquids. Typically, sands pack closely, with values ranging from 1.4 to 1.9 Mg/m^3 (Warrick, 2002). Clays tend to bridge and cannot pack as tightly, giving values from 0.9 to 1.4 Mg/m^3.

B.2 Brooks and Coorey, 1964 (BC) model

BC model formulae are:

$$\frac{\theta}{\theta_s} = \left(\frac{\Phi}{\Phi_s}\right)^{-\frac{1}{b}} \tag{B.2}$$

$$K_h = K_{hs}\left(\frac{\theta}{\theta_s}\right)^{2b+3} \tag{B.3}$$

where θ is the volumetric soil water content, θ is the volumetric soil water content at saturation, Φ is the soil water suction, Φ_s is the saturated soil water suction, b is the Clapp and Hornberger (1978) soil exponent, K_h is the hydraulic conductivity, and K_{hs} is the hydraulic conductivity for saturated soil.

B.3 van Genuchten, 1980 (VG) model

VG model formulae are:

$$S = \frac{\theta - \theta_r}{\theta_s - \theta_r} = \frac{1}{[1 + (\alpha_v \Phi)^n]^m} \tag{B.4}$$

$$K_h = K_{hs} S^{\xi} \left[1 - \left(1 - S^{\frac{1}{m}}\right)^m\right]^2 \tag{B.5}$$

where θ_r is the residual volumetric soil water content, n is the van Genuchten (1980) soil parameter, $m = 1 - 1/n$, and $\xi = 0.5$.

B.4 Root Mean Square Error (RMSE)

RMSE is metric used to measure the average magnitude of the errors between predicted and actual values. RMSE gives more weight to larger errors, which can be useful when you want to penalize large prediction errors more severely. RMSE provides a measure of how spread out the prediction errors are. RMSE is calculated as:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{B.6}$$

where n is the number of data points, y_i represents the observed values, and $\hat{y}_i$ represents the predicted (modeled) values.

B.5 Mean Absolute Deviation (MAD)

MAD is a metric used to measure the overall accuracy of a model's predictions. It quantifies how close the predicted values are to the actual values, regardless of whether the predictions are overestimates or underestimates. MAD is calculated as:

$$\text{MAD} = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i| \tag{B.7}$$

B.6 Van Kampenhout et al. (2017) functions

$$\rho_T = \begin{cases} 50 + 1.7(17)^{\frac{3}{2}} & \text{if } T > T_{\text{frz}} + 1, \\ 50 + 1.7(T - T_{\text{frz}} + 15)^{\frac{3}{2}} & \text{if } T_{\text{frz}} - 15 < T \leq T_{\text{frz}} + 2, \\ -3.8328(T - T_{\text{frz}}) - 0.0333(T - T_{\text{frz}})^2 & \text{if } T \leq T_{\text{frz}} - 15. \end{cases} \tag{B.8}$$

$$\rho_w = 266.861 \left(\frac{1}{2}(1 + \tanh(U/5))\right)^{8.8} \tag{B.9}$$

$$\rho_{\text{fs}} = \rho_T + \rho_w \tag{B.10}$$

where T denotes the atmospheric near-surface temperature (in Kelvin), T_{frz} the freezing temperature of water (273.158K), and U denotes 10 m wind speed in ms^{-1}.

Appendix C

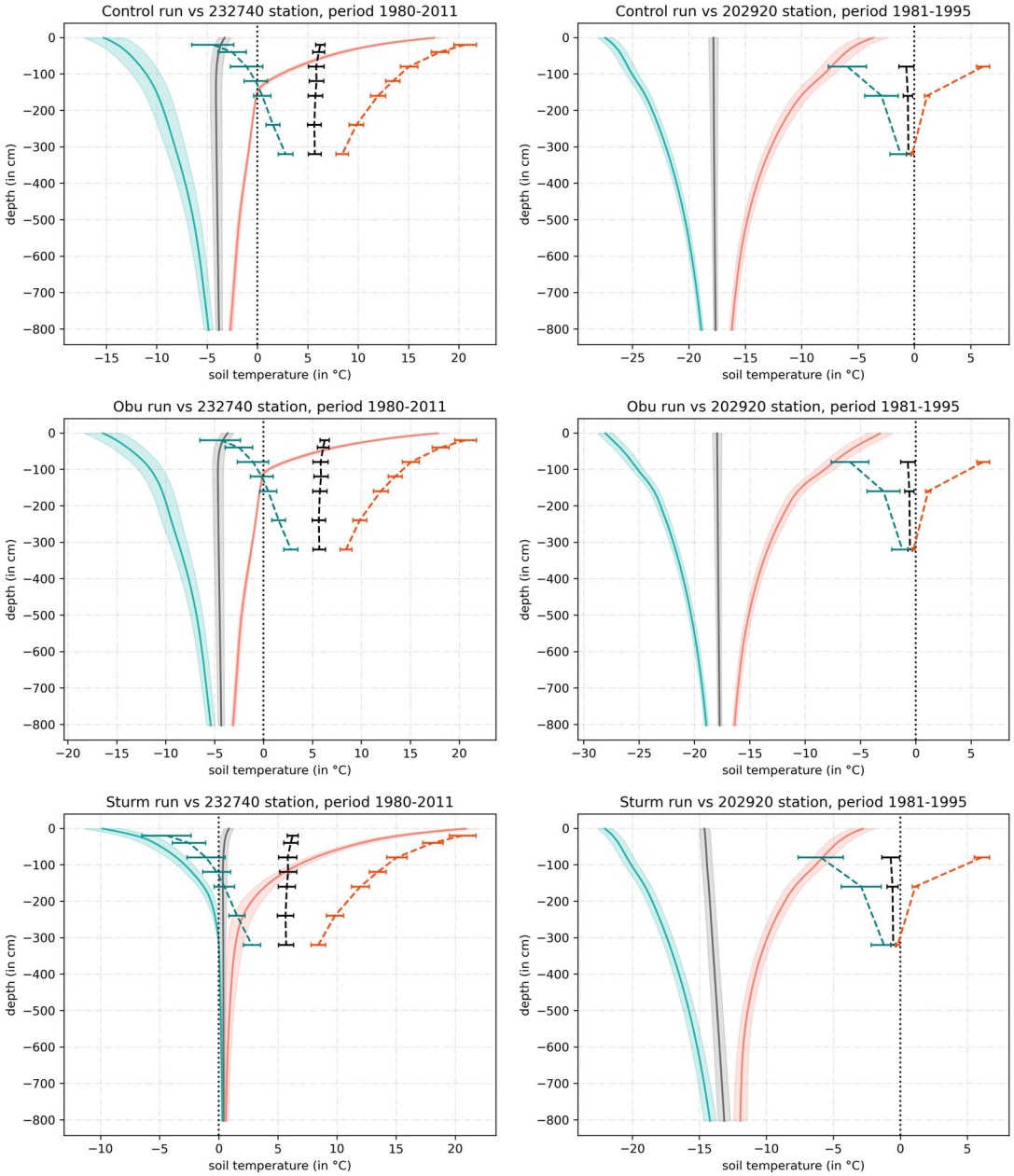
Local comparisons of a list of borehole stations

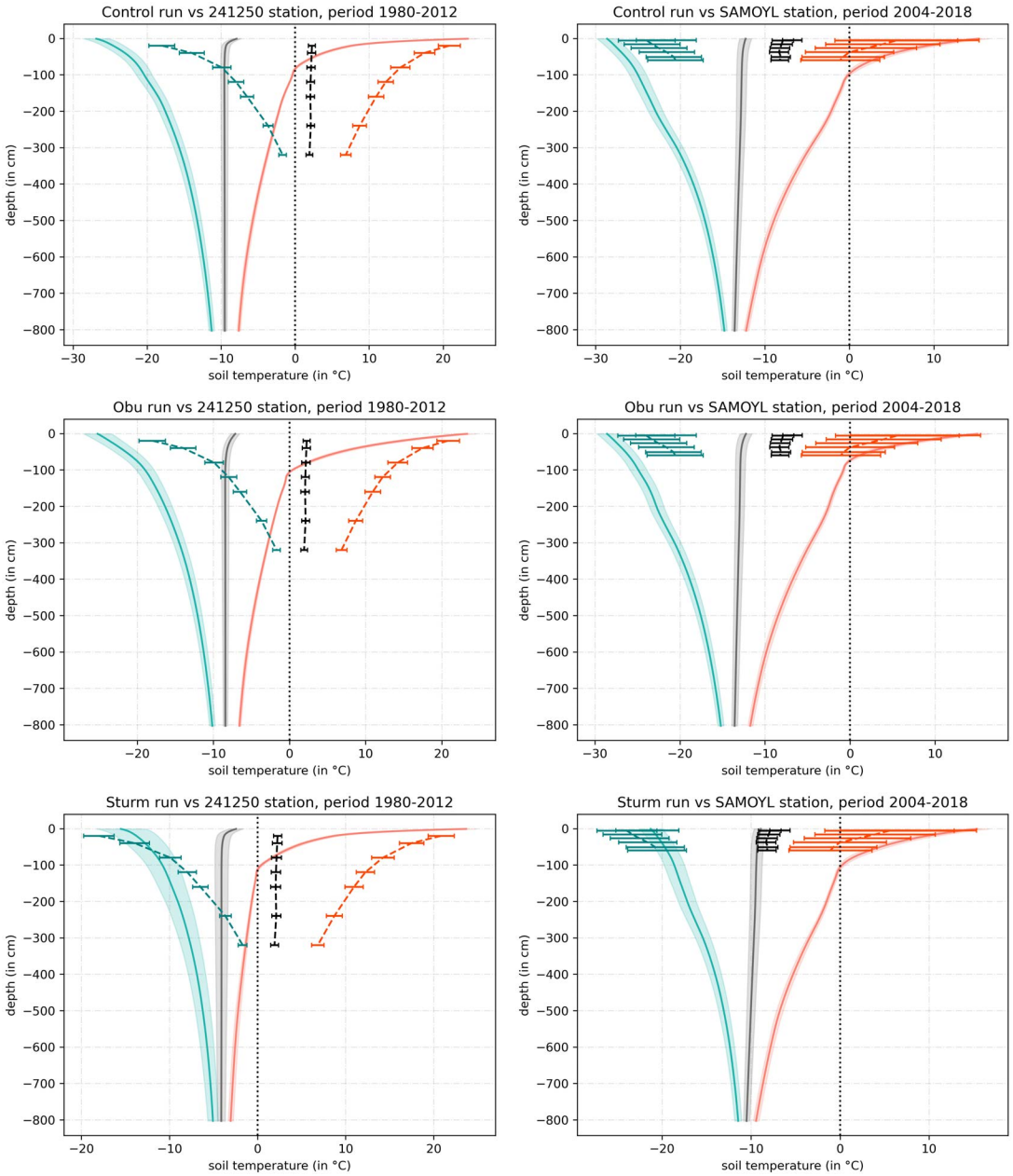
While the main thesis delved into the broader regional trends and patterns, this appendix examines the specific results of 10 selected stations and compare it to the 3 runs used in this study: control, Obu and Sturm runs. These stations were chosen based on the number of data records available, meeting criteria to be classified as permafrost by our model, and spread across diverse Arctic regions. The geographical locations of these stations are illustrated in Figure C.1. The dashed lines represents

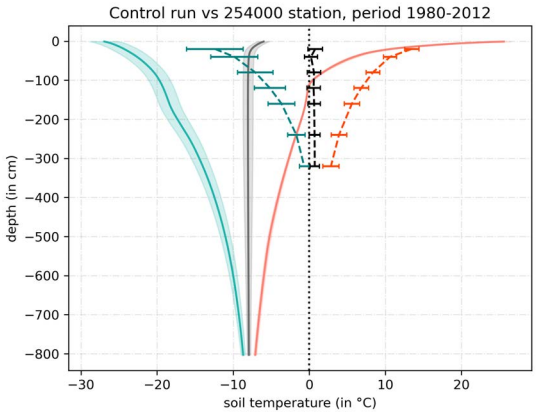
This complementary analysis (Fig. C.2) serves to provide a more comprehensive perspective on soil temperature dynamics, offering valuable insights that may not have been fully explored in the main discussion.

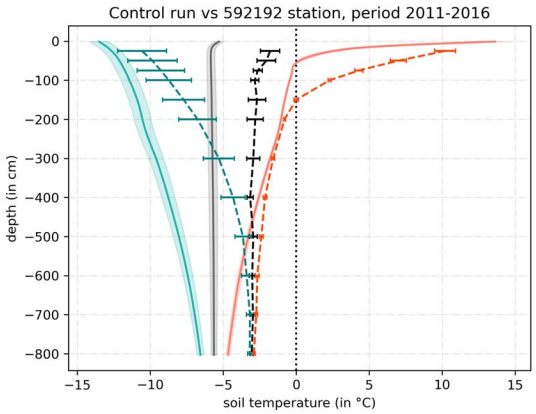
Fig. C.1 Location of the selected 10 borehole stations. U13xxx (68.63°N, -149.60 °E), SAMOYL (72.37°N, 126.47°W), IV4-1x (68.48°N, -155.74°E), 241250 (68.50°N, 112.43°W), 232740 (67.43°N, 86.62°W), 592192 (63.75°N, -68.55°E), ISAibx (78.78°N, -103.55°E), BISibx (73.22°N, -119.56°E), 254000 (65.73°N, 150.90°W), 202920 (77.72°N, 104.30°W).

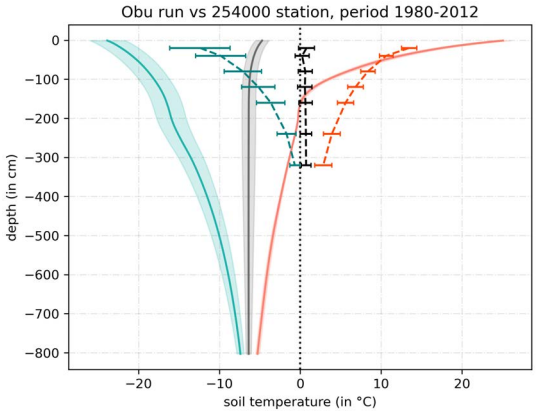
Fig. C.2 The dashed lines represent the station observations from 295GT, and the solid lines represent the model results. Red lines are MaxAGT, black lines are MAGT, and blue lines are MinAGT.

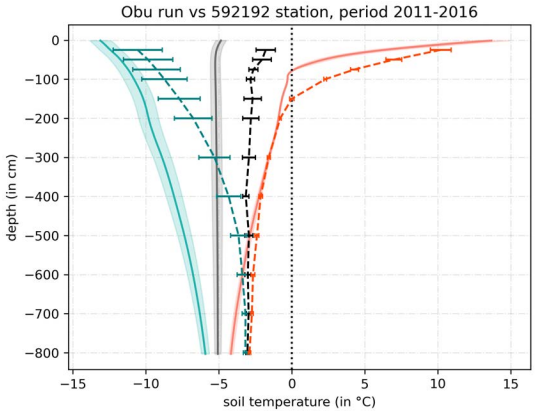


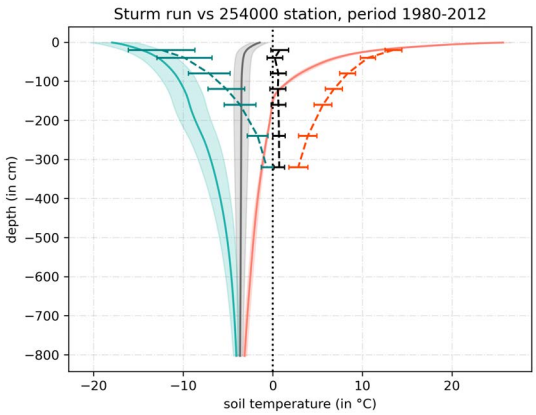

Control run vs 241250 station, period 1980-2012
Control run vs SAMOYL station, period 2004-2018
Obu run vs 241250 station, period 1980-2012
Obu run vs SAMOYL station, period 2004-2018
Sturm run vs 241250 station, period 1980-2012
Sturm run vs SAMOYL station, period 2004-2018
depth (in cm)
soil temperature (in °C)

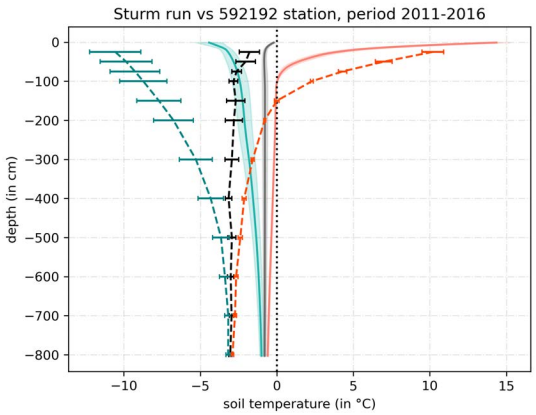

Control run vs 254000 station, period 1980-2012
depth (in cm)
soil temperature (in °C)

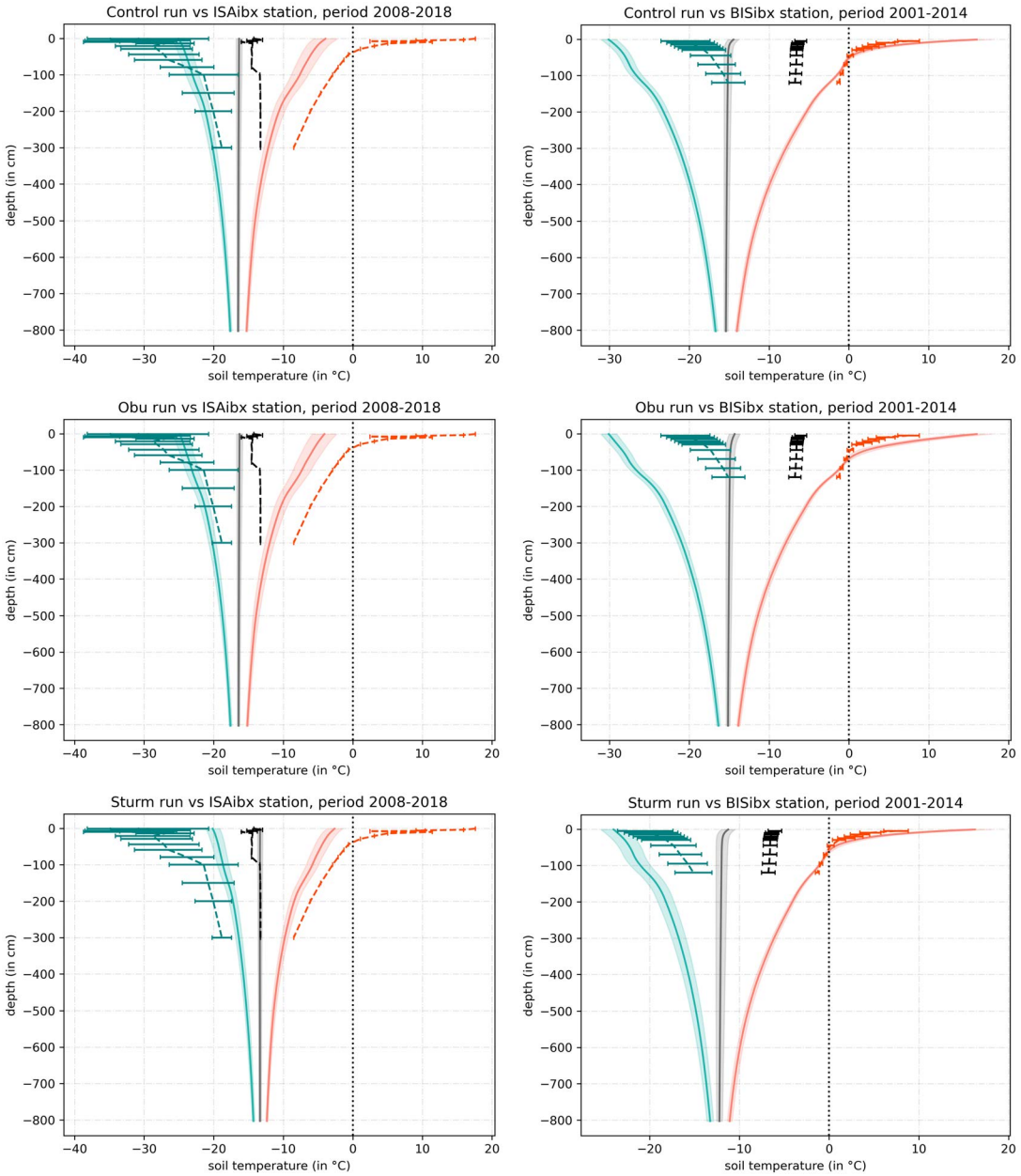

Control run vs 592192 station, period 2011-2016
depth (in cm)
soil temperature (in °C)

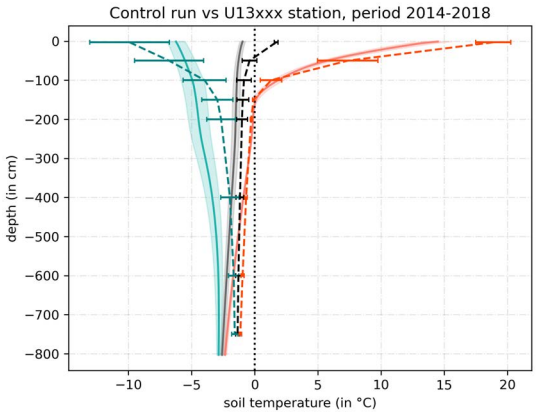

Obu run vs 254000 station, period 1980-2012
depth (in cm)
soil temperature (in °C)

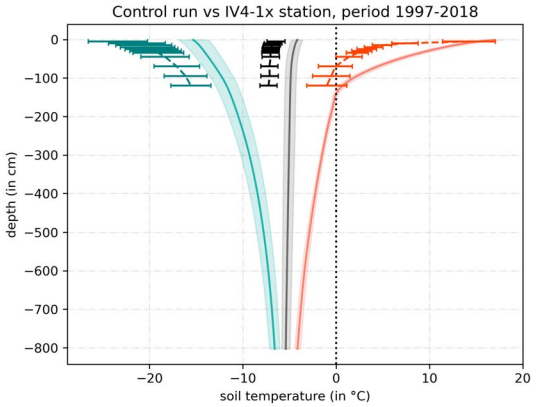

Obu run vs 592192 station, period 2011-2016
depth (in cm)
soil temperature (in °C)

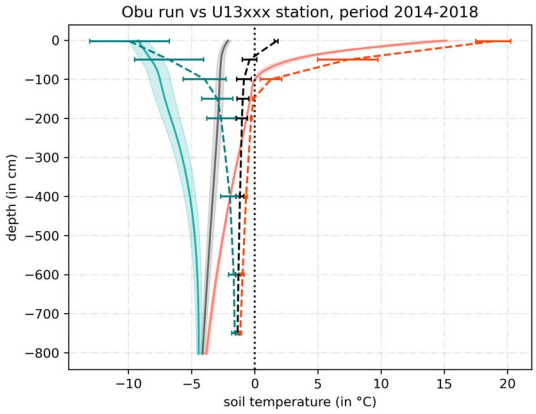

Sturm run vs 254000 station, period 1980-2012
depth (in cm)
soil temperature (in °C)

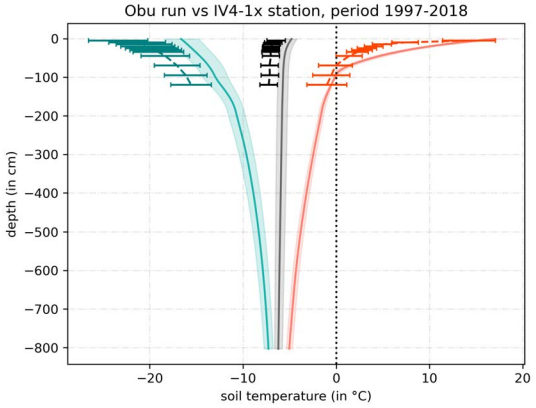

Sturm run vs 592192 station, period 2011-2016
depth (in cm)
soil temperature (in °C)

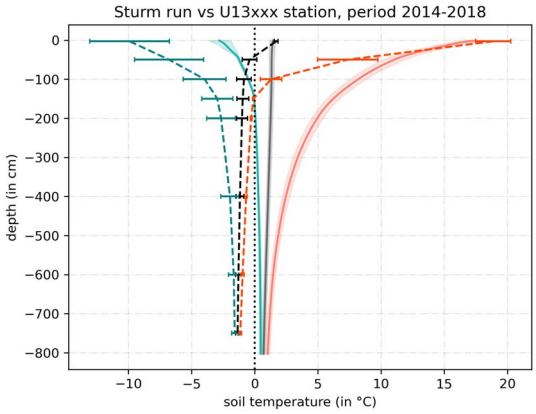

Control run vs ISAibx station, period 2008-2018
Control run vs BISibx station, period 2001-2014
Obu run vs ISAibx station, period 2008-2018
Obu run vs BISibx station, period 2001-2014
Sturm run vs ISAibx station, period 2008-2018
Sturm run vs BISibx station, period 2001-2014
depth (in cm)
soil temperature (in °C)

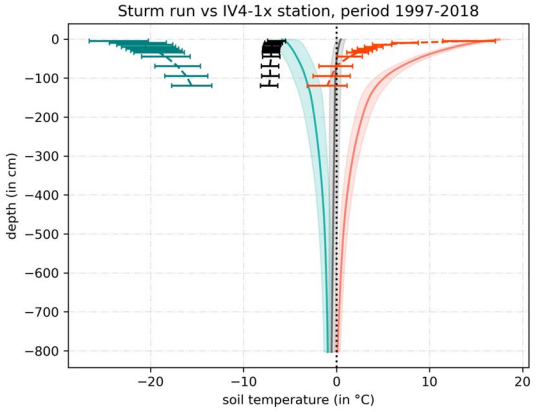

Control run vs U13xxx station, period 2014-2018
depth (in cm)
soil temperature (in °C)


Control run vs IV4-1x station, period 1997-2018
depth (in cm)
soil temperature (in °C)


Obu run vs U13xxx station, period 2014-2018
depth (in cm)
soil temperature (in °C)


Obu run vs IV4-1x station, period 1997-2018
depth (in cm)
soil temperature (in °C)


Sturm run vs U13xxx station, period 2014-2018
depth (in cm)
soil temperature (in °C)


Sturm run vs IV4-1x station, period 1997-2018
depth (in cm)
soil temperature (in °C)

REFERENCES

Adams, E. E. and Sato, A. (1993). Model for effective thermal conductivity of a dry snow cover composed of uniform ice spheres. *Ann. Glaciol.*, 18:300–304.

Albergel, C., Dutra, E., Munier, S., Calvet, J.-C., Munoz-Sabater, J., De Rosnay, P., and Balsamo, G. (2018). ERA-5 and ERA-Interim driven ISBA land surface model simulations: which one performs better? *Hydrol. Earth Syst. Sci.*, 22(6):3515–3532.

Alexeev, V. A., Nicolsky, D. J., Romanovsky, V. E., and Lawrence, D. M. (2007). An evaluation of deep soil configurations in the CLM3 for improved representation of permafrost: HOW DEEP SHOULD THE CLM3 SOIL LAYER BE? *Geophys. Res. Lett.*, 34(9).

Anderson, E. A. (1976). A point energy and mass balance model of a snow cover.

Balland, V. and Arp, P. A. (2005). Modeling soil thermal conductivities over a wide range of conditions. *Journal of Environmental Engineering and Science*, 4(6):549–558.

Barrere, M., Domine, F., Decharme, B., Morin, S., Vionnet, V., and Lafaysse, M. (2017). Evaluating the performance of coupled snow-soil models in SURFEXv8 to simulate the permafrost thermal regime at a high Arctic site. *Geoscientific Model Development*, 10(9):3461–3479.

Best, M. J., Pryor, M., Clark, D. B., Rooney, G. G., Essery, R. L. H., Ménard, C. B., Edwards, J. M., Hendry, M. A., Porson, A., Gedney, N., Mercado, L. M., Sitch, S., Blyth, E., Boucher, O., Cox, P. M., Grimmond, C. S. B., and Harding, R. J. (2011). The Joint UK Land Environment Simulator (JULES), model description – Part 1: Energy and water fluxes. *Geosci. Model Dev.*, 4(3):677–699.

Birch, L., Schwalm, C., Natali, S., Lombardozzi, D., Keppel-Aleks, G., Watts, J., Lin, X., Zona, D., Oechel, W., Sachs, T., Black, T. A., and Rogers, B. (2020). Addressing Biases in Arctic-Boreal Carbon Cycling in the Community Land Model Version 5. *Geoscientific Model Development Discussions*, 1(December):1–31.

Biskaborn, B. K., Smith, S. L., Noetzli, J., Matthes, H., Vieira, G., Streletskiy, D. A., Schoeneich, P., Romanovsky, V. E., Lewkowicz, A. G., Abramov, A., Allard, M., Boike,

J., Cable, W. L., Christiansen, H. H., Delaloye, R., Diekmann, B., Drozdov, D., Etzelmüller, B., Grosse, G., Guglielmin, M., Ingeman-Nielsen, T., Isaksen, K., Ishikawa, M., Johansson, M., Johannsson, H., Joo, A., Kaverin, D., Kholodov, A., Konstantinov, P., Kröger, T., Lambiel, C., Lanckman, J.-P., Luo, D., Malkova, G., Meiklejohn, I., Moskalenko, N., Oliva, M., Phillips, M., Ramos, M., Sannel, A. B. K., Sergeev, D., Seybold, C., Skryabin, P., Vasiliev, A., Wu, Q., Yoshikawa, K., Zheleznyak, M., and Lantuit, H. (2019). Permafrost is warming at a global scale. *Nat Commun*, 10(1):264.

Blome, T. (2014). *Influence of different permafrost rocesses on the large-scale energy and water cycles over Siberia*, volume 150. Reports on Earth System Science, Hamburg.

Bonan, G. B. (1998). The Land Surface Climatology of the NCAR Land Surface Model Coupled to the NCAR Community Climate Model*. *J. Climate*, 11(6):1307–1326.

Bonan, G. B., Levis, S., Kergoat, L., and Oleson, K. W. (2002a). Landscapes as patches of plant functional types: An integrating concept for climate and ecosystem models. *Global Biogeochemical Cycles*, 16(2).

Bonan, G. B., Oleson, K. W., Vertenstein, M., Levis, S., Zeng, X., Dai, Y., Dickinson, R. E., and Yang, Z.-L. (2002b). The Land Surface Climatology of the Community Land Model Coupled to the NCAR Community Climate Model*. *J. Climate*, 15(22):3123–3149.

Bonfils, C. J., Phillips, T. J., Lawrence, D. M., Cameron-Smith, P., Riley, W. J., and Subin, Z. M. (2012). On the influence of shrub height and expansion on northern high latitude climate. *Environmental Research Letters*, 7(1).

Boone, A., Samuelsson, P., Gollvik, S., Napoly, A., Jarlan, L., Brun, E., and Decharme, B. (2016). The Interactions between Soil-Biosphere-Atmosphere (ISBA) land surface model Multi-Energy Balance (MEB) option in SURFEX – Part 1: Model description. preprint, Climate and Earth System Modeling.

Bowden, W. B. (2010). Climate Change in the Arctic - Permafrost, Thermokarst, and Why They Matter to the Non-Arctic World: Climate change in the Arctic: why it matters. *Geography Compass*, 4(10):1553–1566.

Brooks, R. H. and Corey, A. T. (1966). Properties of Porous Media Affecting Fluid Flow. *J. Irrig. and Drain. Div.*, 92(2):61–88.

Burke, E. J., Dankers, R., Jones, C. D., and Wiltshire, A. J. (2013). A retrospective analysis of pan Arctic permafrost using the JULES land surface model. *Clim Dyn*, 41(3-4):1025–1038.

Burke, E. J., Zhang, Y., and Krinner, G. (2020). Evaluating permafrost physics in the Coupled Model Intercomparison Project 6 (CMIP6) models and their sensitivity to climate change. *Cryosphere*, 14(9):3155–3174.

Cai, L., Lee, H., Westermann, S., and Aas, K. S. (2019). Projecting Circum-Arctic Excess Ground Ice Melt with a sub-grid representation in the Community Land Model. preprint, Frozen ground/Frozen Ground.

Campbell, G. S. (1974). A Simple Method for Determining Unsatarated Conductivity From Moisture Retention Data. *Soil Science*, 117(6):311–314.

Chadburn, S. E., Burke, E. J., Essery, R. L. H., Boike, J., Langer, M., Heikenfeld, M., Cox, P. M., and Friedlingstein, P. (2015). Impact of model developments on present and future simulations of permafrost in a global land-surface model. *The Cryosphere*, 9(4):1505–1521.

Chadburn, S. E., Krinner, G., Porada, P., Bartsch, A., Beer, C., Belelli Marchesini, L., Boike, J., Ekici, A., Elberling, B., Friborg, T., Hugelius, G., Johansson, M., Kuhry, P., Kutzbach, L., Langer, M., Lund, M., Parmentier, F. J. W., Peng, S., Van Huissteden, K., Wang, T., Westermann, S., Zhu, D., and Burke, E. J. (2017). Carbon stocks and fluxes in the high latitudes: Using site-level data to evaluate Earth system models. *Biogeosciences*, 14(22):5143–5169.

Chen, Y., Hu, F. S., and Lara, M. J. (2021). Divergent shrub-cover responses driven by climate, wildfire, and permafrost interactions in Arctic tundra ecosystems. *Global Change Biology*, 27(3):652–663.

Cheng, Y., Musselman, K. N., Swenson, S., Lawrence, D., Hamman, J., Dagon, K., Kennedy, D., and Newman, A. J. (2023). Moving Land Models Toward More Actionable Science: A Novel Application of the Community Terrestrial Systems Model Across Alaska and the Yukon River Basin. *Water Resources Research*, 59(1):e2022WR032204.

Clapp, R. B. and Hornberger, G. M. (1978). Empirical equations for some soil hydraulic properties. *Water Resour. Res.*, 14(4):601–604.

Cox, P. M., Betts, R. A., Bunton, C. B., Essery, R. L. H., Rowntree, P. R., and Smith, J. (1999). The impact of new land surface physics on the GCM simulation of climate and climate sensitivity. *Climate Dynamics*, 15(3):183–203.

Curasi, S. R., Melton, J. R., Humphreys, E. R., Wang, L., Seiler, C., Cannon, A. J., Chan, E., and Qu, B. (2023). Evaluating the Performance of the Canadian Land Surface Scheme Including Biogeochemical Cycles (CLASSIC) Tailored to the Pan-Canadian Domain. *J Adv Model Earth Syst*, 15(4):e2022MS003480.

Côté, J. and Konrad, J.-M. (2005). A generalized thermal conductivity model for soils and construction materials. *Can. Geotech. J.*, 42(2):443–458.

Dai, Y., Wei, N., Yuan, H., Zhang, S., Shangguan, W., Liu, S., Lu, X., and Xin, Y. (2019). Evaluation of Soil Thermal Conductivity Schemes for Use in Land Surface Modeling. *J Adv Model Earth Syst*, 11(11):3454–3473.

Dankers, R., Burke, E. J., and Price, J. (2011). Simulation of permafrost and seasonal thaw depth in the JULES land surface scheme. *The Cryosphere*, 5(3):773–790.

de Vries, D. (1963). Thermal properties of soils. In *W. R. Wijk and A. J. W. Borghorst (EDs.), Physics of the Plant Environment*, pages 210–235. North-Holland, Amsterdam.

Decharme, B., Boone, A., Delire, C., and Noilhan, J. (2011). Local evaluation of the Interaction between Soil Biosphere Atmosphere soil multilayer diffusion scheme using four pedotransfer functions. *J. Geophys. Res.*, 116(D20):D20126.

Decharme, B., Brun, E., Boone, A., Delire, C., LeMoigne, P., and Morin, S. (2016). Impacts of snow and organic soils parameterization on northern Eurasian soil temperature profiles simulated by the ISBA land surface model. *The Cryosphere*, 10(2):853–877.

Dharssi, I., Vidale, P. L., Verhoef, A., Macpherson, B., Jones, C., and Best, M. (2009). New soil physical properties implemented in the Unified Model.

Domine, F., Barrere, M., and Sarrazin, D. (2016). Seasonal evolution of the effective thermal conductivity of the snow and thesoil in high Arctic herb tundra at Bylot Island, Canada. *The Cryosphere*, 10(6):2573–2588.

Domine, F., Belke-Brea, M., Sarrazin, D., Arnaud, L., Barrere, M., and Poirier, M. (2018). Soil moisture, wind speed and depth hoar formation in the Arctic snowpack. *J. Glaciol.*, 64(248):990–1002.

Domine, F., Picard, G., Morin, S., Barrere, M., Madore, J.-B., and Langlois, A. (2019). Major Issues in Simulating Some Arctic Snowpack Properties Using Current Detailed Snow Physics Models: Consequences for the Thermal Regime and Water Budget of Permafrost. *J. Adv. Model. Earth Syst.*, 11(1):34–44.

Dutch, V. R., Rutter, N., Wake, L., Sandells, M., Derksen, C., Walker, B., Hould Gosselin, G., Sonnentag, O., Essery, R., Kelly, R., Marsh, P., King, J., and Boike, J. (2022). Impact of measured and simulated tundra snowpack properties on heat transfer. *The Cryosphere*, 16(10):4201–4222.

Ekici, A., Beer, C., Hagemann, S., Boike, J., Langer, M., and Hauck, C. (2014). Simulating high-latitude permafrost regions by the JSBACH terrestrial ecosystem model. *Geoscientific Model Development*, 7(2):631–647.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.*, 9(5):1937–1958.

Farouki, O. T. (1981). The thermal properties of soils in cold regions. *Cold Regions Science and Technology*, 5(1):67–75.

Fisher, R. A. and Koven, C. D. (2020). Perspectives on the Future of Land Surface Models and the Challenges of Representing Complex Terrestrial Systems. *J Adv Model Earth Syst*, 12(4):e2018MS001453.

Fourteau, K., Brondex, J., Brun, F., and Dumont, M. (2023). A novel numerical implementation for the surface energy budget of melting snowpacks and glaciers. preprint, Cryosphere.

Fourteau, K., Hagenmuller, P., Roulle, J., and Domine, F. (2022). On the use of heated needle probes for measuring snow thermal conductivity. *J. Glaciol.*, 68(270):705–719.

Gibson, C. M., Brinkman, T., Cold, H., Brown, D., and Turetsky, M. (2021). Identifying increasing risks of hazards for northern land-users caused by permafrost thaw: integrating scientific and community-based research approaches. *Environ. Res. Lett.*, 16(6):064047.

Golaz, J., Caldwell, P. M., Van Roekel, L. P., Petersen, M. R., Tang, Q., Wolfe, J. D., Abeshu, G., Anantharaj, V., Asay-Davis, X. S., Bader, D. C., Baldwin, S. A., Bisht, G., Bogenschutz, P. A., Branstetter, M., Brunke, M. A., Brus, S. R., Burrows, S. M., Cameron-Smith, P. J., Donahue, A. S., Deakin, M., Easter, R. C., Evans, K. J., Feng, Y., Flanner, M., Foucar, J. G., Fyke, J. G., Griffin, B. M., Hannay, C., Harrop, B. E., Hoffman, M. J., Hunke, E. C., Jacob, R. L., Jacobsen, D. W., Jeffery, N., Jones, P. W., Keen, N. D., Klein, S. A., Larson, V. E., Leung, L. R., Li, H., Lin, W., Lipscomb, W. H., Ma, P., Mahajan, S., Maltrud, M. E., Mametjanov, A., McClean, J. L., McCoy, R. B., Neale, R. B., Price, S. F., Qian, Y., Rasch, P. J., Reeves Eyre, J. E. J., Riley, W. J., Ringler, T. D., Roberts, A. F., Roesler, E. L., Salinger, A. G., Shaheen, Z., Shi, X., Singh, B., Tang, J., Taylor, M. A., Thornton, P. E., Turner, A. K., Veneziani, M., Wan, H., Wang, H., Wang, S., Williams, D. N., Wolfram, P. J., Worley, P. H., Xie, S., Yang, Y., Yoon, J., Zelinka, M. D., Zender, C. S., Zeng, X., Zhang, C., Zhang, K., Zhang, Y., Zheng, X., Zhou, T., and Zhu, Q. (2019). The DOE E3SM Coupled Model Version 1: Overview and Evaluation at Standard Resolution. *J. Adv. Model. Earth Syst.*, 11(7):2089–2129.

Goodrich, L. E. (1976). *A numerical model for assessing the influence of snow cover on the ground thermal regime*. McGill University, Montreal.

Gouttevin, I., Krinner, G., Ciais, P., Polcher, J., and Legout, C. (2012). Multi-scale validation of a new soil freezing scheme for a land-surface model with physically-based hydrology. *The Cryosphere*, 6(2):407–430.

Gouttevin, I., Langer, M., Löwe, H., Boike, J., Proksch, M., and Schneebeli, M. (2018). Observation and modelling of snow at a polygonal tundra permafrost site: spatial variability and thermal implications. *The Cryosphere*, 12(11):3693–3717.

Gruber, S. (2012). Derivation and analysis of a high-resolution estimate of global permafrost zonation. *The Cryosphere*, 6(1):221–233.

Guimberteau, M., Zhu, D., Maignan, F., Huang, Y., Yue, C., Dantec-Nédélec, S., Ottlé, C., Jornet-Puig, A., Bastos, A., Laurent, P., Goll, D., Bowring, S., Chang, J., Guenet, B., Tifafi, M., Peng, S., Krinner, G., Ducharne, A., Wang, F., Wang, T., Wang, X., Wang, Y., Yin, Z., Lauerwald, R., Joetzjer, E., Qiu, C., Kim, H., and Ciais, P. (2018). ORCHIDEE-MICT (v8.4.1), a land surface model for the high latitudes: model description and validation. *Geosci. Model Dev.*, 11(1):121–163.

Guo, Q., Kino, K., Li, S., Nitta, T., Takeshima, A., Nitta, T., Onuma, Y., Satoh, Y., Suzuki, T., Takata, K., Yoshida, N., and Yoshimura, K. (2021). Description of MATSIRO6.

Hagemann, S. and Stacke, T. (2015). Impact of the soil hydrology scheme on simulated soil moisture memory. *Clim Dyn*, 44(7-8):1731–1750.

He, H., Flerchinger, G. N., Kojima, Y., Dyck, M., and Lv, J. (2021). A review and evaluation of 39 thermal conductivity models for frozen soils. *Geoderma*, 382:114694.

Heijmans, M. M., Magnússon, R., Lara, M. J., Frost, G. V., Myers-Smith, I. H., van Huissteden, J., Jorgenson, M. T., Fedorov, A. N., Epstein, H. E., Lawrence, D. M., and Limpens, J. (2022). Tundra vegetation change and impacts on permafrost. *Nature Reviews Earth and Environment*, 3(1):68–84. Publisher: Springer US ISBN: 0123456789.

Heim, B., Lisovski, S., Wieczorek, M., Pellet, C., Delaloye, R., Bartsch, A., Jakober, D., Pointner, G., Strozzi, T., and Seifert, F. M. (2021). D4.1 Product Validation and Intercomparison Report (PVIR).

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., Rosnay, P., Rozum, I., Vamborg, F.,

Villaume, S., and Thépaut, J. (2020). The ERA5 global reanalysis. *Q.J.R. Meteorol. Soc.*, 146(730):1999–2049.

Hillel, D. (2003). *Introduction to environmental soil physics*. Elsevier.

Hinzman, L. D., Deal, C. J., McGuire, A. D., Mernild, S. H., Polyakov, I. V., and Walsh, J. E. (2013). Trajectory of the Arctic as an integrated system. *Ecological Applications*, 23(8):1837–1868.

Hu, G., Zhao, L., Li, R., Park, H., Wu, X., Su, Y., Guggenberger, G., Wu, T., Zou, D., Zhu, X., Zhang, W., Wu, Y., and Hao, J. (2023). Water and heat coupling processes and its simulation in frozen soils: Current status and future research directions. *CATENA*, 222:106844.

Hu, G., Zhao, L., Wu, X., Li, R., Wu, T., Xie, C., Pang, Q., and Zou, D. (2017). Comparison of the thermal conductivity parameterizations for a freeze-thaw algorithm with a multi-layered soil in permafrost regions. *CATENA*, 156:244–251.

Hudson, B. D. (1994). Soil organic matter and available water capacity. *Journal of Soil and Water Conservation*, 49(2):5.

Hugelius, G., Tarnocai, C., Broll, G., Canadell, J. G., Kuhry, P., and Swanson, D. K. (2013). The northern circumpolar soil carbon database: Spatially distributed datasets of soil coverage and soil carbon storage in the northern permafrost regions. *Earth System Science Data*, 5(1):3–13.

Humphrey, V., Berg, A., Ciais, P., Gentine, P., Jung, M., Reichstein, M., Seneviratne, S. I., and Frankenberg, C. (2021). Soil moisture–atmosphere feedback dominates land carbon uptake variability. *Nature*, 592(7852):65–69.

IPCC (1990). *Climate change: the IPCC scientific assessment*. American Scientist, United States.

IPCC (1996). *The Science of Climate Change: Contribution of Working Group I to the Second Assessment of Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press.

IPCC (2021). *Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, 1 edition.

ITPS, F. a. (2020). *Global Soil Organic Carbon Map (GSOCmap) Version 1.5: Technical Report*. FAO.

Johansen, O. (1975). Thermal Conductivity of Soils:. Technical report, Defense Technical Information Center, Fort Belvoir, VA.

Jones, P. W. (1999). First- and Second-Order Conservative Remapping Schemes for Grids in Spherical Coordinates. *Mon. Wea. Rev.*, 127(9):2204–2210.

Jordan, R. E. (1991). A one-dimensional temperature model for a snow cover: Technical documentation for SNTHERM. 89.

Jorgenson, M. T., Racine, C. H., Walters, J. C., and Osterkamp, T. E. (2001). Permafrost Degradation and Ecological Changes Associated with a WarmingClimate in Central Alaska. *Climatic Change*, 48(4):551–579.

Jury, W. A., Gardner, W. R., and Gardner, W. H. (1991). *Soil Physics*. John Wiley and Sons, Inc.

Koven, C. D., Riley, W. J., and Stern, A. (2013). Analysis of permafrost thermal dynamics and response to climate change in the CMIP5 earth system models. *Journal of Climate*, 26(6):1877–1900.

Kravchenko, A. N. and Guber, A. K. (2017). Soil pores and their contributions to soil carbon processes. *Geoderma*, 287:31–39.

Krinner, G., Derksen, C., Essery, R., Flanner, M., Hagemann, S., Clark, M., Hall, A., Rott, H., Brutel-Vuilmet, C., Kim, H., Ménard, C. B., Mudryk, L., Thackeray, C., Wang, L., Arduini, G., Balsamo, G., Bartlett, P., Boike, J., Boone, A., Chéruy, F., Colin, J., Cuntz, M., Dai, Y., Decharme, B., Derry, J., Ducharne, A., Dutra, E., Fang, X., Fierz, C., Ghattas, J., Gusev, Y., Haverd, V., Kontu, A., Lafaysse, M., Law, R., Lawrence, D., Li, W., Marke, T., Marks, D., Ménégoz, M., Nasonova, O., Nitta, T., Niwano, M., Pomeroy, J., Raleigh, M. S., Schaedler, G., Semenov, V., Smirnova, T. G., Stacke, T., Strasser, U., Svenson, S., Turkov, D., Wang, T., Wever, N., Yuan, H., Zhou, W., and Zhu, D. (2018). ESM-SnowMIP: assessing snow models and quantifying snow-related climate feedbacks. *Geosci. Model Dev.*, 11(12):5027–5049.

Kucharik, C. J., Foley, J. A., Delire, C., Fisher, V. A., Coe, M. T., Lenters, J. D., Young-Molling, C., Ramankutty, N., Norman, J. M., and Gower, S. T. (2000). Testing the performance of a dynamic global ecosystem model: Water balance, carbon balance, and vegetation structure. *Global Biogeochemical Cycles*, 14(3):795–825.

Langer, M., Westermann, S., Heikenfeld, M., Dorn, W., and Boike, J. (2013). Satellite-based modeling of permafrost temperatures in a tundra lowland landscape. *Remote Sensing of Environment*, 135:12–24.

Lawrence, D., Fischer, R., Koven, C., Oleson, K., Swenson, S., and Vertenstein, M. (2018). Technical Description of version 5.0 of the Community Land Model (CLM).

Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., Collier, N., Ghimire, B., van Kampenhout, L., Kennedy, D., Kluzek, E., Lawrence, P. J., Li, F., Li, H., Lombardozzi, D., Riley, W. J., Sacks, W. J., Shi, M., Vertenstein, M., Wieder, W. R., Xu, C., Ali, A. A., Badger, A. M., Bisht, G., van den Broeke, M., Brunke, M. A., Burns, S. P., Buzan, J., Clark, M., Craig, A., Dahlin, K., Drewniak, B., Fisher, J. B., Flanner, M., Fox, A. M., Gentine, P., Hoffman, F., Keppel-Aleks, G., Knox, R., Kumar, S., Lenaerts, J., Leung, L. R., Lipscomb, W. H., Lu, Y., Pandey, A., Pelletier, J. D., Perket, J., Randerson, J. T., Ricciuto, D. M., Sanderson, B. M., Slater, A., Subin, Z. M., Tang, J., Thomas, R. Q., Val Martin, M., and Zeng, X. (2019). The Community Land Model Version 5: Description of New Features, Benchmarking, and Impact of Forcing Uncertainty. *Journal of Advances in Modeling Earth Systems*, 11(12):4245–4287.

Lawrence, D. M. and Slater, A. G. (2005). A projection of severe near-surface permafrost degradation during the 21st century. *Geophys. Res. Lett.*, 32(24):L24401.

Lawrence, D. M. and Slater, A. G. (2008). Incorporating organic soil into a global climate model. *Clim Dyn*, 30(2-3):145–160.

Lawrence, D. M. and Slater, A. G. (2010). The contribution of snow condition trends to future ground climate. *Clim Dyn*, 34(7-8):969–981.

Lawrence, D. M., Slater, A. G., Romanovsky, V. E., and Nicolsky, D. J. (2008). Sensitivity of a model projection of near-surface permafrost degradation to soil column depth and representation of soil organic matter. *J. Geophys. Res.*, 113(F2):F02011.

Lawrence, D. M. and Swenson, S. C. (2011). Permafrost response to increasing Arctic shrub abundance depends on the relative influence of shrubs on local soil cooling versus large-scale climate warming. *Environmental Research Letters*, 6(4).

Le Moigne, P. (2018). SURFEX SCIENTIFIC DOCUMENTATION.

Lee, H., Swenson, S. C., Slater, A. G., and Lawrence, D. M. (2014). Effects of excess ground ice on projections of permafrost in a warming climate. *Environ. Res. Lett.*, 9(12):124006.

Lenaerts, J. T., Medley, B., van den Broeke, M. R., and Wouters, B. (2019). Observing and Modeling Ice Sheet Surface Mass Balance. *Reviews of Geophysics*, page 2018RG000622. Publisher: John Wiley and Sons, Ltd.

Li, G., Zhao, Y., Zhang, W., and Xu, X. (2021). Influence of snow cover on temperature field of frozen ground. *Cold Regions Science and Technology*, 192:103402.

Li, M., Wu, P., and Ma, Z. (2020). A comprehensive evaluation of soil moisture and soil temperature from third-generation atmospheric and land reanalysis data sets. *International Journal of Climatology*, 40(13):5744–5766.

Li, R., Zhao, L., Wu, T., Wang, Q., Ding, Y., Yao, J., Wu, X., Hu, G., Xiao, Y., Du, Y., Zhu, X., Qin, Y., Yang, S., Bai, R., Du, E., Liu, G., Zou, D., Qiao, Y., and Shi, J. (2019). Soil thermal conductivity and its influencing factors at the Tanggula permafrost region on the Qinghai–Tibet Plateau. *Agricultural and Forest Meteorology*, 264:235–246.

Lohmann, D., Nolte-Holube, R., and Raschke, E. (1996). A large-scale horizontal routing model to be coupled to land surface parametrization schemes. *Tellus A*, 48(5):708–721.

Loranty, M. M., Abbott, B. W., Blok, D., Douglas, T. A., Epstein, H. E., Forbes, B. C., Jones, B. M., Kholodov, A. L., Kropp, H., Malhotra, A., Mamet, S. D., Myers-Smith, I. H., Natali, S. M., O'Donnell, J. A., Phoenix, G. K., Rocha, A. V., Sonnentag, O., Tape, K. D., and Walker, D. A. (2018). Reviews and syntheses: Changing ecosystem influences on soil thermal regimes in northern high-latitude permafrost regions. *Biogeosciences*, 15(17):5287–5313.

Loranty, M. M., Goetz, S. J., and Beck, P. S. A. (2011). Tundra vegetation effects on pan-Arctic albedo. *Environmental Research Letters*, 6(2):029601.

Luo, S., Lü, S., and Zhang, Y. (2009). Development and validation of the frozen soil parameterization scheme in Common Land Model. *Cold Regions Science and Technology*, 55(1):130–140.

Matthes, H., Rinke, A., Zhou, X., and Dethloff, K. (2017). Uncertainties in coupled regional Arctic climate simulations associated with the used land surface model. *Journal of Geophysical Research: Atmospheres*, 122(15):7755–7771. Publisher: Wiley-Blackwell.

Mellor, M. (1977). Engineering Properties of Snow. *J. Glaciol.*, 19(81):15–66.

Melton, J. R., Arora, V. K., Wisernig-Cojoc, E., Seiler, C., Fortier, M., Chan, E., and Teckentrup, L. (2020). CLASSIC v1.0: the open-source community successor to the Canadian Land Surface Scheme (CLASS) and the Canadian Terrestrial Ecosystem Model (CTEM) – Part 1: Model framework and site-level performance. *Geosci. Model Dev.*, 13(6):2825–2850.

Melton, J. R., Verseghy, D. L., Sospedra-Alfonso, R., and Gruber, S. (2019). Improving permafrost physics in the coupled Canadian Land Surface Scheme (v.3.6.2) and Canadian Terrestrial Ecosystem Model (v.2.1) (CLASS-CTEM). *Geosci. Model Dev.*, 12(10):4443–4467.

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. (2021). Scientific and Human Errors in a Snow Model Intercomparison. *Bulletin of the American Meteorological Society*, 102(1):E61–E79.

Meredith, M., Sommerkorn, M., Cassota, S., Derksen, C., Ekaykin, A., Hollowed, A., Kofinas, G., Mackintosh, A., Melbourne-Thomas, J., Muelbert, M. M., et al. (2019). Polar regions.

Milly, P. C. D., Malyshev, S. L., Shevliakova, E., Dunne, K. A., Findell, K. L., Gleeson, T., Liang, Z., Phillipps, P., Stouffer, R. J., and Swenson, S. (2014). An Enhanced Model of Land Water and Energy for Global Hydrologic and Earth-System Studies. *Journal of Hydrometeorology*, 15(5):1739–1761.

Minayeva, T., Sirin, A., Kershaw, P., and Bragg, O. (2016). Arctic Peatlands. In Finlayson, C. M., Milton, G. R., Prentice, R. C., and Davidson, N. C., editors, *The Wetland Book*, pages 1–15. Springer Netherlands, Dordrecht.

Miner, K. R., Turetsky, M. R., Malina, E., Bartsch, A., Tamminen, J., McGuire, A. D., Fix, A., Sweeney, C., Elder, C. D., and Miller, C. E. (2022). Permafrost carbon emissions in a changing Arctic. *Nat Rev Earth Environ*, 3(1):55–67.

Mondav, R., McCalley, C. K., Hodgkins, S. B., Frolking, S., Saleska, S. R., Rich, V. I., Chanton, J. P., and Crill, P. M. (2017). Microbial network, phylogenetic diversity and community membership in the active layer across a permafrost thaw gradient. *Environmental Microbiology*, 19(8):3201–3218.

Mooney, P. A., Sobolowski, S., and Lee, H. (2020). Designing and evaluating regional climate simulations for high latitude land use land cover change studies. *Tellus A: Dynamic Meteorology and Oceanography*, 72(1):1853437.

Mudryk, L. R., Derksen, C., Howell, S., Laliberté, F., Thackeray, C., Sospedra-Alfonso, R., Vionnet, V., Kushner, P. J., and Brown, R. (2018). Canadian snow and sea ice: historical trends and projections. *The Cryosphere*, 12(4):1157–1176.

Muller, S. W. (1945). *Permafrost, Or Permanently Frozen Ground: And Related Engineering Problems*. Number 62. Army map service, US Army.

Myers-Smith, I. H., Forbes, B. C., Wilmking, M., Hallinger, M., Lantz, T., Blok, D., Tape, K. D., MacIas-Fauria, M., Sass-Klaassen, U., Lévesque, E., Boudreau, S., Ropars, P., Hermanutz, L., Trant, A., Collier, L. S., Weijers, S., Rozema, J., Rayback, S. A., Schmidt, N. M., Schaepman-Strub, G., Wipf, S., Rixen, C., Ménard, C. B., Venn, S., Goetz, S., Andreu-Hayles, L., Elmendorf, S., Ravolainen, V., Welker, J., Grogan, P., Epstein, H. E., and Hik, D. S. (2011). Shrub expansion in tundra ecosystems: Dynamics, impacts and research priorities. *Environmental Research Letters*, 6(4).

Nicolsky, D. and Shakhova, N. (2010). Modeling sub-sea permafrost in the East Siberian Arctic Shelf: the Dmitry Laptev Strait. *Environ. Res. Lett.*, 5(1):015006.

Noilhan, J. and Planton, S. (1989). A Simple Parameterization of Land Surface Processes for Meteorological Models. *Mon. Wea. Rev.*, 117(3):536–549.

Ochsner, T. E., Horton, R., and Ren, T. (2001). A New Perspective on Soil Thermal Properties. *Soil Sci. Soc. Am. J.*, 65(6):1641–1647.

Oleson, K. W. and Bonan, G. B. (2000). The Effects of Remotely Sensed Plant Functional Type and Leaf Area Index on Simulations of Boreal Forest Surface Fluxes by the NCAR Land Surface Model. *J. Hydrometeor*, 1(5):431–446.

Oleson, K. W., Lawrence, D. M., Bonan, G. B., Drewniak, B., Huang, M., Levis, S., Li, F., Riley, W. J., Swenson, S. C., Thornton, P. E., Bozbiyik, A., Fisher, R., Heald, C. L., Kluzek, E., Lamarque, F., Lawrence, P. J., Leung, L. R., Muszala, S., Ricciuto, D. M., Sacks, W., Sun, Y., Tang, J., and Yang, Z.-L. (2013). Technical Description of version 4.5 of the Community Land Model (CLM).

Oogathoo, S., Houle, D., Duchesne, L., and Kneeshaw, D. (2022). Evaluation of simulated soil moisture and temperature for a Canadian boreal forest. *Agricultural and Forest Meteorology*, 323:109078.

Osterkamp, T. E. and Romanovsky, V. E. (1999). Evidence for warming and thawing of discontinuous permafrost in Alaska. *Permafrost Periglac. Process.*, 10(1):17–37.

Ou, M. and Zhang, S. (2022). Evaluation and Comparison of the Common Land Model and the Community Land Model by Using In Situ Soil Moisture Observations from the Soil Climate Analysis Network. *Land*, 11(1):126.

Overland, J., Dunlea, E., Box, J. E., Corell, R., Forsius, M., Kattsov, V., Olsen, M. S., Pawlak, J., Reiersen, L.-O., and Wang, M. (2019). The urgency of Arctic change. *Polar Science*, 21:6–13.

Palmtag, J., Obu, J., Kuhry, P., Richter, A., Siewert, M. B., Weiss, N., Westermann, S., and Hugelius, G. (2022). A high spatial resolution soil carbon and nitrogen dataset for the northern permafrost region based on circumpolar land cover upscaling. *Earth Syst. Sci. Data*, 14(9):4095–4110.

Paquin, J.-P. and Sushama, L. (2015). On the Arctic near-surface permafrost and climate sensitivities to soil and snow model formulations in climate models. *Clim Dyn*, 44(1-2):203–228.

Park, H., Fedorov, A. N., Zheleznyak, M. N., Konstantinov, P. Y., and Walsh, J. E. (2015). Effect of snow cover on pan-Arctic permafrost thermal regimes. *Clim Dyn*, 44(9-10):2873–2895.

Pearson, R. G., Phillips, S. J., Loranty, M. M., Beck, P. S., Damoulas, T., Knight, S. J., and Goetz, S. J. (2013). Shifts in Arctic vegetation and associated feedbacks under climate change. *Nature Climate Change*, 3(7):673–677. Publisher: Nature Publishing Group.

Peters-Lidard, C. D., Blackburn, E., Liang, X., and Wood, E. F. (1998). The Effect of Soil Thermal Conductivity Parameterization on Surface Energy Fluxes and Temperatures. *J. Atmos. Sci.*, 55(7):1209–1224.

Post, E., Alley, R. B., Christensen, T. R., Macias-Fauria, M., Forbes, B. C., Gooseff, M. N., Iler, A., Kerby, J. T., Laidre, K. L., Mann, M. E., Olofsson, J., Stroeve, J. C., Ulmer, F., Virginia, R. A., and Wang, M. (2019). The polar regions in a 2°C warmer world. *Sci. Adv.*, 5(12):eaaw9883.

Pribyl, D. W. (2010). A critical review of the conventional SOC to SOM conversion factor. *Geoderma*, 156(3-4):75–83.

Ramage, J., Jungsberg, L., Wang, S., Westermann, S., Lantuit, H., and Heleniak, T. (2021). Population living on permafrost in the Arctic. *Popul Environ*, 43(1):22–38.

Rantanen, M., Karpechko, A. Y., Lipponen, A., Nordling, K., Hyvärinen, O., Ruosteenoja, K., Vihma, T., and Laaksonen, A. (2022). The Arctic has warmed nearly four times faster than the globe since 1979. *Communications Earth and Environment*, 3(1):1–10. Publisher: Springer US.

Reick, C. H., Gayler, V., Goll, D., Hagemann, S., Heidkamp, M., Nabel, J. E. M. S., Raddatz, T., Roeckner, E., Schnur, R., and Wilkenskjeld, S. (2021). JSBACH 3 - The land component of the MPI Earth System Model: documentation of version 3.2. page 4990986. Artwork Size: 4990986 Medium: application/pdf Publisher: MPI für Meteorologie Version Number: 1.

Richards, L. A. (1931). Capillary conduction of liquids through porous medias. *Physics*, 1(5):318–333.

Riche, F. and Schneebeli, M. (2010). Microstructural change around a needle probe to measure thermal conductivity of snow. *J. Glaciol.*, 56(199):871–876.

Rinke, A., Kuhry, P., and Dethloff, K. (2008). Importance of a soil organic layer for Arctic climate: A sensitivity study with an Arctic RCM. *Geophys. Res. Lett.*, 35(13):L13709.

Romanovsky, V. E., Smith, S. L., and Christiansen, H. H. (2010). Permafrost thermal state in the polar Northern Hemisphere during the international polar year 2007–2009: a synthesis. *Permafrost and Periglacial*, 21(2):106–116.

Royer, A., Picard, G., Vargel, C., Langlois, A., Gouttevin, I., and Dumont, M. (2021). Improved Simulation of Arctic Circumpolar Land Area Snow Properties and Soil Temperatures. *Front. Earth Sci.*, 9:685140.

Saito, K. (2008). Arctic land hydrothermal sensitivity under warming: Idealized offline evaluation of a physical terrestrial scheme in a global climate model. *J. Geophys. Res.*, 113(D21):D21106.

Sass, J. H., Lachenbruch, A. H., and Munroe, R. J. (1971). Thermal conductivity of rocks from measurements on fragments and its application to heat-flow determinations. *J. Geophys. Res.*, 76(14):3391–3401.

Schaefer, K., Lantuit, H., Romanovsky, V. E., Schuur, E. A. G., and Witt, R. (2014). The impact of the permafrost carbon feedback on global climate. *Environ. Res. Lett.*, 9(8):085003.

Scherler, M., Hauck, C., Hoelzle, M., Stähli, M., and Völksch, I. (2010). Meltwater infiltration into the frozen active layer at an alpine permafrost site. *Permafrost Periglac. Process.*, 21(4):325–334.

Schuur, E. A. and Mack, M. C. (2018). Ecological Response to Permafrost Thaw and Consequences for Local and Global Ecosystem Services. *Annu. Rev. Ecol. Evol. Syst.*, 49(1):279–301.

Schuur, E. A., McGuire, A. D., Schädel, C., Grosse, G., Harden, J. W., Hayes, D. J., Hugelius, G., Koven, C. D., Kuhry, P., Lawrence, D. M., Natali, S. M., Olefeldt, D., Romanovsky, V. E., Schaefer, K., Turetsky, M. R., Treat, C. C., and Vonk, J. E. (2015). Climate change and the permafrost carbon feedback. *Nature*, 520(7546):171–179.

Schuur, E. A. G., Bockheim, J., Canadell, J. G., Euskirchen, E., Field, C. B., Goryachkin, S. V., Hagemann, S., Kuhry, P., Lafleur, P. M., Lee, H., Mazhitova, G., Nelson, F. E., Rinke, A., Romanovsky, V. E., Shiklomanov, N., Tarnocai, C., Venevsky, S., Vogel, J. G., and Zimov, S. A. (2008). Vulnerability of Permafrost Carbon to Climate Change: Implications for the Global Carbon Cycle. *BioScience*, 58(8):701–714.

Schwerdtfeger, P. (1963). Theoretical derivation of the thermal conductivity and diffusivity of snow. *Int. Assoc. Sci. Hydrol.*, 61:75–81.

Shao, M., Zhang, S., Niu, B., Pei, Y., Song, S., Lei, T., and Yun, H. (2022). Soil texture influences soil bacterial biomass in the permafrost-affected alpine desert of the Tibetan plateau. *Front. Microbiol.*, 13:1007194.

Shiklomanov, N. I., Streletskiy, D. A., and Nelson, F. E. (2012). Northern Hemisphere Component of the Global Circumpolar Active Layer Monitoring (CALM) Program. Technical report.

Slater, A. G., Lawrence, D. M., and Koven, C. D. (2017). Process-level model evaluation: a snow and heat transfer metric. *The Cryosphere*, 11(2):989–996.

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

Sturm, M. and Johnson, J. B. (1992). Thermal conductivity measurements of depth hoar. *J. Geophys. Res.*, 97(B2):2129–2139.

Sulman, B. N., Salmon, V. G., Iversen, C. M., Breen, A. L., Yuan, F., and Thornton, P. E. (2021). Integrating Arctic Plant Functional Types in a Land Surface Model Using Above- and Belowground Field Observations. *J Adv Model Earth Syst*, 13(4).

Sun, S., Jin, J., and Xue, Y. (1999). A simple snow-atmosphere-soil transfer model. *J. Geophys. Res.*, 104(D16):19587–19597.

Sushama, L., Laprise, R., Caya, D., Verseghy, D., and Allard, M. (2007). An RCM projection of soil thermal and moisture regimes for North American permafrost zones. *Geophys. Res. Lett.*, 34(20):L20711.

Swenson, S. C. and Lawrence, D. M. (2014). Assessing a dry surface layer-based soil resistance parameterization for the community land model using GRACE and FLUXNET-MTE data. *Journal of Geophysical Research*, 119(17):10,299–10,312. Publisher: Wiley-Blackwell.

Takata, K., Emori, S., and Watanabe, T. (2003). Development of the minimal advanced treatments of surface interaction and runoff. *Global and Planetary Change*, 38(1-2):209–222.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A. (2012). An Overview of CMIP5 and the Experiment Design. *Bulletin of the American Meteorological Society*, 93(4):485–498.

Thackeray, C. W. and Fletcher, C. G. (2016). Snow albedo feedback: Current knowledge, importance, outstanding issues and future directions. *Progress in Physical Geography: Earth and Environment*, 40(3):392–408.

Van Den Hurk, B., Kim, H., Krinner, G., Seneviratne, S. I., Derksen, C., Oki, T., Douville, H., Colin, J., Ducharne, A., Cheruy, F., Viovy, N., Puma, M. J., Wada, Y., Li, W., Jia, B., Alessandri, A., Lawrence, D. M., Weedon, G. P., Ellis, R., Hagemann, S., Mao, J., Flanner, M. G., Zampieri, M., Materia, S., Law, R. M., and Sheffield, J. (2016). LS3MIP (v1.0) contribution to CMIP6: the Land Surface, Snow and Soil-moisture Model Intercomparison Project – aims, setup and expected outcome. *Geosci. Model Dev.*, 9(8):2809–2832.

Van Genuchten, M. T. (1980). A Closed-form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils. *Soil Science Society of America Journal*, 44(5):892–898.

van Kampenhout, L., Lenaerts, J. T., Lipscomb, W. H., Sacks, W. J., Lawrence, D. M., Slater, A. G., and van den Broeke, M. R. (2017). Improving the Representation of Polar Snow and Firn in the Community Earth System Model. *Journal of Advances in Modeling Earth Systems*, 9(7):2583–2600. Publisher: Blackwell Publishing Ltd.

Verseghy, D. L. (1991). Class-A Canadian land surface scheme for GCMS. I. Soil model. *Int. J. Climatol.*, 11(2):111–133.

Verseghy, D. L. (2012). CLASS - The Canadian Land Surface Scheme (Version 3.6) - Technical Documnetation.

Viterbo, P. and Beljaars, A. C. M. (1995). An Improved Land Surface Parameterization Scheme in the ECMWF Model and Its Validation. *J. Climate*, 8(11):2716–2748.

Walker, M. D., Wahren, C. H., Hollister, R. D., Henry, G. H. R., Ahlquist, L. E., Alatalo, J. M., Bret-Harte, M. S., Calef, M. P., Callaghan, T. V., Carroll, A. B., Epstein, H. E., Jónsdóttir, I. S., Klein, J. A., Magnússon, B., Molau, U., Oberbauer, S. F., Rewa, S. P., Robinson, C. H., Shaver, G. R., Suding, K. N., Thompson, C. C., Tolvanen, A., Totland, O., Turner, P. L., Tweedie, C. E., Webber, P. J., and Wookey, P. A. (2006). Plant community responses to experimental warming across the tundra biome. *Proc. Natl. Acad. Sci. U.S.A.*, 103(5):1342–1346.

Wang, F., Cheruy, F., and Dufresne, J.-L. (2016a). The improvement of soil thermodynamics and its effects on land surface meteorology in the IPSL climate model. *Geosci. Model Dev.*, 9(1):363–381.

Wang, L., Zhou, J., Qi, J., Sun, L., Yang, K., Tian, L., Lin, Y., Liu, W., Shrestha, M., Xue, Y., Koike, T., Ma, Y., Li, X., Chen, Y., Chen, D., Piao, S., and Lu, H. (2017). Development of a land surface model with coupled snow and frozen soil physics. *Water Resour. Res.*, 53(6):5085–5103.

Wang, T., Ottlé, C., Boone, A., Ciais, P., Brun, E., Morin, S., Krinner, G., Piao, S., and Peng, S. (2013). Evaluation of an improved intermediate complexity snow scheme in the ORCHIDEE land surface model: ORCHIDEE SNOW MODEL EVALUATION. *J. Geophys. Res. Atmos.*, 118(12):6064–6079.

Wang, W., Rinke, A., Moore, J. C., Ji, D., Cui, X., Peng, S., Lawrence, D. M., McGuire, A. D., Burke, E. J., Chen, X., Decharme, B., Koven, C., MacDougall, A., Saito, K., Zhang, W., Alkama, R., Bohn, T. J., Ciais, P., Delire, C., Gouttevin, I., Hajima, T., Krinner, G., Lettenmaier, D. P., Miller, P. A., Smith, B., Sueyoshi, T., and Sherstiukov, A. B. (2016b). Evaluation of air-soil temperature relationships simulated by land surface models during winter across the permafrost region. *Cryosphere*, 10(4):1721–1737.

Warrick, A. W., editor (2002). *Soil physics companion*. CRC Press, Boca Raton, Fla.

Westermann, S., Bartsch, A., and Strozzi, T. (2020). D2.2 Algorithm Theoritical Basic Document (ATDB).

Winstral, A., Magnusson, J., Schirmer, M., and Jonas, T. (2019). The Bias-Detecting Ensemble: A New and Efficient Technique for Dynamically Incorporating Observations Into Physics-Based, Multilayer Snow Models. *Water Resources Research*, 55(1):613–631.

Woodside, W. (1958). Calculation of the thermal conductivity of porous media. *Can. J. Phys.*, 36(7):815–823.

Xiao, Y., Zhao, L., Dai, Y., Li, R., Pang, Q., and Yao, J. (2013). Representing permafrost properties in CoLM for the Qinghai–Xizang (Tibetan) Plateau. *Cold Regions Science and Technology*, 87:68–77.

Yang, S., Li, R., Zhao, L., Wu, T., Wu, X., Zhang, Y., Shi, J., and Qiao, Y. (2022). Evaluation of the Performance of CLM5.0 in Soil Hydrothermal Dynamics in Permafrost Regions on the Qinghai–Tibet Plateau. *Remote Sensing*, 14(24):6228.

Yang, Z.-p., Ou, Y. H., Xu, X.-l., Zhao, L., Song, M.-h., and Zhou, C.-p. (2010). Effects of permafrost degradation on ecosystems. *Acta Ecologica Sinica*, 30(1):33–39.

Yen, Y.-C. (1962). Effective thermal conductivity of ventilated snow. *J. Geophys. Res.*, 67(3):1091–1098.

Yen, Y.-C. (1981). *Review of thermal properties of snow, ice, and sea ice*, volume 81. US Army, Corps of Engineers, Cold Regions Research and Engineering Laboratory.

Zhang, G., Chen, Y., and Li, J. (2021). Effects of organic soil in the Noah-MP land-surface model on simulated skin and soil temperature profiles and surface energy exchanges for China. *Atmospheric Research*, 249:105284.

Zhao, M., Golaz, J.-C., Held, I. M., Guo, H., Balaji, V., Benson, R., Chen, J.-H., Chen, X., Donner, L. J., Dunne, J. P., Dunne, K., Durachta, J., Fan, S.-M., Freidenreich, S. M., Garner, S. T., Ginoux, P., Harris, L. M., Horowitz, L. W., Krasting, J. P., Langenhorst, A. R., Liang, Z., Lin, P., Lin, S.-J., Malyshev, S. L., Mason, E., Milly, P. C. D., Ming, Y., Naik, V., Paulot, F., Paynter, D., Phillipps, P., Radhakrishnan, A., Ramaswamy, V., Robinson, T., Schwarzkopf, D., Seman, C. J., Shevliakova, E., Shen, Z., Shin, H., Silvers, L. G., Wilson, J. R., Winton, M., Wittenberg, A. T., Wyman, B., and Xiang, B. (2018). The GFDL Global Atmosphere and Land Model AM4.0/LM4.0: 2. Model Description, Sensitivity Studies, and Tuning Strategies. *J. Adv. Model. Earth Syst.*, 10(3):735–769.

Zhao, W., Mu, C., Han, L., Sun, W., Sun, Y., and Zhang, T. (2023). Spatial and temporal variability in snow density across the Northern Hemisphere. *CATENA*, 232:107445.

Zhu, D., Ciais, P., Krinner, G., Maignan, F., Jornet Puig, A., and Hugelius, G. (2019). Controls of soil organic matter on soil thermal dynamics in the northern high latitudes. *Nat Commun*, 10(1):3172.

Declaration of independence

I hereby assure that I have independently composed the present dissertation, using no sources or aids other than those indicated, and I have appropriately marked all direct and paraphrased quotations as such. This scholarly work has not been presented to any examination authority in the same or similar form and has not been published.

_____

Adrien Damseaux