Vegetation Distribution and Biophysics, Permafrost, and Feedback Mechanisms to Arctic Change

by

Wenbo Zhou

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Civil Engineering)
in The University of Michigan
2023

Doctoral Committee:

Professor Valeriy Y. Ivanov, Chair
Associate Professor Brian R. Ellis
Professor George W. Kling
Associate Professor Aleksey Y. Sheshukov, Kansas State University

Wenbo Zhou

zhouwb@umich.edu

ORCID iD: 0000-0003-4010-1466

This dissertation is dedicated to my family.

ACKNOWLEDGEMENTS

Many people have provided significant support to accomplish this dissertation and have accompanied me for this part of my academic journey. I would like to take this chance to express my gratitude to them.

I will always be grateful to my advisor and mentor, Professor Valeriy Yuryevich Ivanov, for his invaluable guidance, advice, and support starting from the very first day of this program. His patience help me pass through those bottlenecks I encountered. He has also taught me a lot more than just knowledge. I remember his words when the pathway to start my doctoral program wasn't clear as we were waiting for the grants: "It's like playing piano, it won't always be the black keys, you will hit the white keys at some point." I also remember when he tried to let me do impromptu Salsa in Puerto Rico: "There isn't always a chance to practice." These little moments of his give me a good reference to view stuff from a different angle.

I am also grateful to my committee members: Professor George W. King, and Professor Brian R. Ellis, for their great support on this dissertation and their advice on how to view the study in the big picture and how to tell the story of my research in a better way.

I want to give special thanks to my committee member Professor Aleksey Y. Sheshukov at Kansas State University, along with Professor Jingfeng Wang at the Georgia Institute of Technology, and Dr. Khachik Sargsyan at the Sandia National Laboratories, for sharing their expertise unselfishly in permafrost freeze-thaw modeling, ground heat flux modeling, and uncertainty quantification methods, respectively.

Their knowledge and support provide cornerstones for this dissertation.

I would also like to thank the professors in the CEE program: Prof. Nikolaos Katopodes, Prof. Steven Wright, Prof. Aline Cotel, Prof. Avery Demond, and Prof. Jeremy Bricker, who provided great classes and mentorship, and filled my knowledge reservoir to conduct further research.

The biggest missing parts of my Ph.D. experience are traveling to Yamal and conducting real fieldwork due to the unprecedented situation caused by the pandemic. This makes the help from our Russian colleagues throughout the research even more valuable. I would like to thank Dr. Valeriy Mazepa, Dr. Stepan Shiyatov, Dr. Yulia V Shalaumova, Dr. Aleksandr Sokolov, Dr. Victor Valdayskikh, and the people in the Yamal research group for their support in field data collection.

I am glad to have had my HYDROWIT labmates accompany me through my time in EWRE: Chase Dwelle, Elizabeth Agee, Donghui Xu, Weicheng Huang, Kevin Murphy, Liujing Zhang, Tyeen Taylor, Vinh Tran, Boqi Tian, and Xin Shen. Thank you for your friendship and support which make the lonely trip not so lonely. There are also many other friends who I met here in the class and on the soccer field that I cannot fully list here, thank you all for making Michigan a warm place.

Finally, I have to express my deepest gratitude to my parents, Meihua Hu and Ruoyu Zhou. Thank you for letting me chase freely the target I select and for the support since I was a child, especially during the recent three years that I cannot come back home. I will never be where I am now without your unconditional love and sacrifice. I have nothing to pay in turn but keep moving forward. The pandemic has taken many lives away, including my dearest grandmother, I hope you and my grandfather can see these words from above with joy.

TABLE OF CONTENTS

CHAPTER

LIST OF FIGURES

Figure

LIST OF TABLES

<u>Table</u>

LIST OF APPENDICES

Appendix

ABSTRACT

Arctic terrestrial ecosystems have been changing in response to global warming over the past decades. Studies have reported shrub and tree expansion into tundra ecosystems. This process will have feedback on climate change through changes in albedo, evapotranspiration, and the carbon cycle. However, such vegetation shift is not happening uniformly towards the tundra area. The factors that control this encroachment need to be identified and the underlying mechanism of their interaction with vegetation distribution patterns needs to be explained at a regional and local scale. Meanwhile, historical field observations of ground soil temperatures have shown a general warming trend in permafrost soils across the Arctic. The ongoing warming process of the permafrost strongly affects the regional ecosystems, infrastructure, and indigenous communities in the Arctic. However, the knowledge about the surface ground heat flux, a key factor that controls the subsurface thermal dynamics, is still limited in the cold region. Projections of future climate and the subsurface thermal state are conducted based on model simulations with significant uncertainties yet to be addressed before such results can be used to make any conclusion. It is important to solve these issues first to understand the potential Arctic change in the future and infer its feedback to the global climate system.

In this dissertation, these processes have been investigated using field measurements and model results, with a focus on study sites located in Northwestern Siberia. The relationships between spatiotemporal patterns of the encroaching tall vegetation to the microtopography and snow cover are demonstrated. Preference for tree growth locations has been shown for surfaces with divergent characteristics (well-drained).

The change of distribution patterns of tree locations is related to the snow distribution shifts by applying clustering analysis.

Data sets for surface ground heat flux G_S are sparse in the Arctic region from observation due to challenges in field measurements, and not available from the Global Climate Model (GCM) direct outputs. An uncertainty-informed framework is developed by combining the analytical model and the state-of-art uncertainty quantification (UQ) machinery to reconstruct G_S from field measurements and GCM outputs of shallow soil temperatures. The reconstructed G_S and its probability distribution are valuable to be used as the boundary forcing for subsurface thermal regime simulations.

Climate projections from GCMs are retrieved to reconstruct the G_S for the future in the study sites. A Bayesian Weighted Averaging (BWA) stochastic downscaling method is applied to reduce the bias and uncertainty carried in the GCMs with historical borehole observations. The long-term ground soil temperature is simulated by a physical numerical model calibrated by historical observations until the end of this century using the downscaled annual mean $\bar{G}_S$ as the top boundary control. The projected degree of permafrost degradation in terms of surface soil temperature is higher than the global average at two of the borehole sites for the selected future climate scenarios. Almost one-seventh of the current permafrost will disappear at one of the study sites under the worst carbon emission scenario, with other cases also showing some level of permafrost loss. These will in turn have a huge impact on global climate through carbon release, destabilize the infrastructure foundation, and change the surface and subsurface hydrology and thermal regime. The results stress the importance of the ongoing permafrost monitoring work and the urgent need to control the global warming trend. The study of the above-ground vegetation dynamics and below-ground permafrost cryospheric processes all together provides insights to better understand the Arctic change in the future.

CHAPTER I

Introduction

1.1 Impacts of climate change in the Arctic

The global surface temperature is increasing at an unprecedented rate for the last few decades according to the latest Intergovernmental Panel on Climate Change (IPCC) Sixth Assessment Report (AR6) (*Masson-Delmotte et al.*, 2021). Specifically, the global land surface temperature is 1.59 °C higher in the last decade compared to the pre-industrial condition. Observations have shown an air temperature increase in the Arctic higher than twice the global average. As a consequence, the Arctic terrestrial system is changing rapidly with overall greening and regional browning of the tundra, shorter snow cover period, and increased permafrost thaw (*Bieniek et al.*, 2018; *Constable et al.*, 2022). A warming Arctic will have feedback to the climate change through increased permafrost carbon releases (*Schaefer et al.*, 2014; *Schuur et al.*, 2015), changes in surface albedo and energy budget caused by vegetation shifts (*Pearson et al.*, 2013; *Myers-Smith et al.*, 2020), influence the infrastructure systems in cold regions (*Streletskiy et al.*, 2012; *Hjort et al.*, 2018, 2022), and force adaptation in the social systems (*Tyler et al.*, 2007; *Ford et al.*, 2014, 2015). Considering the ongoing rapid warming trend in the Arctic and its important role in global climate circulation, an emerging research question is: how is the Arctic going to change in the future?

To view potential shifts in the Arctic in the future, the historical and ongoing trend of changes need to be understood first. Since the Arctic is a complex system with many sub-components, a logical and feasible strategy is to study the impacts of climate change on individual components in the Arctic systems first, then combine them together considering the inter-component interactions to yield a holistic view of the possible Arctic change. Two of the sub-components in the Arctic with the most significant changes are considered in this study: the above-ground vegetation change and the below-ground permafrost degradation. The following chapters in this dissertation aim to address the big research question by focusing on these two sub-components in a detailed view.

1.2 Research scope

Chapter II reveals the relationship between the spatial patterns of tree stands encroachment and two environmental factors: topography and snow cover. Observations for tree stem diameters, locations of tree growth, and point snow depth are recorded for a study plot in three campaigns starting in the 1960s. Snow distribution maps are extrapolated from the point snow depth measurements. Topographical characteristics such as the curvature indices are derived from surface elevation information at different resolutions (provided by the Arctic Digital Elevation Model (ArcticDEM) products) to capture the surface features at the most significant scale. Statistic analyses are conducted to identify tree growth preference over the curvature indices which are further related to hydrologic characteristics. Clustering analysis for tree stands locations are made to recognize the change in their spatial distribution patterns and their correlation to snow distribution shifts.

Chapter III outlines the general framework that can be applied to reconstruct the surface ground heat flux G_S from different shallow soil temperature data sources. The three fundamental components of this framework are described: (1) An analytical model relates surface soil temperature T_S and ground heat flux G_S through a half-order derivative (HOD) method that is used to retrieve G_S from T_S. (2) A physical-based numerical model that solves the freeze-thaw process in the soil is used to simulate subsurface temperature and water phase series. (3) The uncertainty quantification (UQ) machinery and probabilistic learning techniques are used to derive soil properties used in the HOD model in view of probabilistic description and improved computational efficiency. Applications for field site data measured in detail, GCM outputs, and borehole temperature observations are presented. For all three cases, the commonly required variable is the soil temperature at a shallow depth.

In Chapter IV, projection of the ground soil temperature is conducted using the freeze-thaw numerical model with projected G_S derived from the Bayesian Weighted Averaged (BWA) stochastic downscaling process as the top boundary input. First, GCM outputs for historical and selected future climate scenarios and historical borehole temperature observations are used in the framework developed in Chapter III to reconstruct their corresponding G_S. Second, the G_S reconstructed from GCMs are downscaled by applying BWA with G_S from borehole observations to reduce model bias and uncertainty related to model resolution. Third, the projected G_S at three borehole sites after downscaling is applied as the top boundary condition in the numerical model to conduct the projection for ground temperature. The uncertainty associated with the parameter inference and the downscaling process is addressed at the end.

Chapter V summarizes this dissertation. The critical assumptions and limitations are listed for the research work in each chapter. Finally, the potential future work is discussed as including more comprehensive dynamics through model coupling in the current framework.

CHAPTER II

Spatiotemporal Dynamics of Encroaching Tall Vegetation, a Case Study

The content of this chapter is published in (*Zhou et al.*, 2021). Zhou. W. and Ivanov. V. conceived of the presented idea. Zhou. W. performed the analytical methods, processed the data, analyzed the results, and prepared the draft manuscript. Ivanov. V. supervised the findings of this work. Mazepa. V., Shiyatov. S. and Shalaumova. Y. provided field observations of vegetation and snow characteristics. Zhang. T. and Liu. D. analyzed and provided the digital elevation map. All authors discussed the results and contributed to the final version of the manuscript.

2.1 Introduction

The Arctic terrestrial ecosystems have been changing in response to global warming (*Hinzman et al.*, 2005) over the past decades. Numerous studies have reported shrubification and tree expansion into tundra ecosystems toward higher latitudes and areas of higher elevation (*Sturm et al.*, 2001; *Tape et al.*, 2006; *Forbes et al.*, 2010; *Myers-Smith et al.*, 2011; *Elmendorf et al.*, 2012; *Frost and Epstein*, 2014). The process has been proliferating and is well documented for many places in the Arctic from both direct *in situ* observations, dendrochronological analyses (*Lloyd and*

Fastie, 2002; *Shiyatov et al.*, 2007; *Devi et al.*, 2008; *Van Bogaert et al.*, 2011; *Martin et al.*, 2017) and satellite-based observations (*Lloyd et al.*, 2011; *Frost et al.*, 2014; *Mathisen et al.*, 2014). With a warming trend projected for the coming decades, the change of tundra plants at the circumpolar scale is expected to continue in the future (*Bjorkman et al.*, 2018), including both structural shifts from low to high shrubs or trees (*Forbes et al.*, 2010; *Macias-Fauria et al.*, 2012) and migration-based treeline advances (*MacDonald et al.*, 2008; *Vavrus et al.*, 2012; *Koenigk et al.*, 2013).

The observed replacement of short tundra vegetation (graminoids, forbs, and non-vascular vegetation) with tall erect woody plants (shrubs and trees) is thought to be induced mainly by increasing temperature (*Walker et al.*, 2006; *Elmendorf et al.*, 2012), along with the impacts of other important drivers like precipitation, soil moisture, snow dynamics, and herbivory (*Frost and Epstein*, 2014; *Martin et al.*, 2017). This process will likely have positive feedback on climate and enhance warming through changes in albedo, evapotranspiration, and carbon cycle (*Lafleur et al.*, 2001; *Foley et al.*, 2003; *Chapin III et al.*, 2005; *Pearson et al.*, 2013; *Zhang et al.*, 2013; *Lafleur and Humphreys*, 2018). Furthermore, this process can potentially alter wildlife habitat (*Tape et al.*, 2016). Forest-tundra boundaries have been found to be temperature-sensitive transitional zones (*Esper and Schweingruber*, 2004; *Harsch et al.*, 2009). While treeline location is mainly controlled by summer temperatures and growing season length (*MacDonald et al.*, 2008; *Hoch and Körner*, 2009), treeline advances have also been found to be associated with winter warming (*Kullman*, 2007; *Hagedorn et al.*, 2014). Although tree growth is considered to be mainly limited by low temperatures (*Körner and Paulsen*, 2004), other environmental conditions also affect tree distribution patterns (*Holtmeier and Broll*, 2005) leading to spatial heterogeneity in treeline advances (*Lloyd et al.*, 2002; *Wilmking et al.*, 2004).

The northern Siberian region encompasses the largest forest-tundra ecotone in the world (*Frost and Epstein*, 2014) and exhibits large gradients in summer conditions and regional differences in vegetation distribution (*MacDonald et al.*, 2008). Previous studies have shown various rates of vegetation expansion across northern Siberia strongly correlated with increased winter precipitation (*Frost and Epstein*, 2014; *Devi et al.*, 2020). This can be explained by the impact of winter precipitation on soil hydrological and thermal mechanisms. On the one hand, the increase in winter precipitation may provide extra water supply for seedlings and saplings during the growing season (*Grigorieva and Moiseev*, 2018; *Devi et al.*, 2020); on the other hand, snow cover is essential for new trees to survive along the treeline as snowpack increases soil temperature and protects against frost and wind abrasion stresses (*Holtmeier and Broll*, 2007; *Kharuk et al.*, 2010a; *Hagedorn et al.*, 2014). Landscape features also influence the spatial distribution of forests, thereby mediating their response to ambient warming (*Kharuk et al.*, 2010b; *Dearborn and Danby*, 2018). In the areas of complex topography, including mountainous regions, tall vegetation would have higher or lower favorability over different microtopographic niches. This preference of vegetation depends on environmental factors such as slope type and snow redistribution, which can lead to spatially varying treeline sensitivity to climate (*Holtmeier and Broll*, 2005). Topography-mediated hydrological conditions such as drought and flooding stresses can also affect the boreal forest response to climate at regional scales (*Sato and Kobayashi*, 2018). Soil nutrient is also a key player in biomass dynamics if the low-temperature limitation of tree growth at the treeline starts weakening in the future (*Hagedorn et al.*, 2020). An emerging question is: as the climate is becoming less limiting for the encroachment of tall vegetation, are there other appreciable factors controlling the distribution of trees over the forest-tundra ecotone?

This study aims to analyze spatiotemporal patterns of the larch forest encroachment into tundra and relevant environmental factors with a focus on two features: microtopography and snow cover. We analyze the timberline ecotone in the Polar Urals region, where a unique set of records on tree growth dynamics along with environmental features data have been available since the 1960s. The 'timberline ecotone' refers to a transitional belt of mountain vegetation between the upper limit of closed forests and the upper limit of single tree growth in the tundra (*Körner*, 2003; *Shiyatov et al.*, 2005). We hypothesize that the spatiotemporal pattern of tree distribution of this ecotone is associated with niches of favorability caused by microtopography and snow cover.

2.2 Materials and methods

2.2.1 Study site and vegetation

The study region of the Rai-Iz massif is located on the eastern slope of the Polar Urals range (200– 350 m asl.), in the Sob' River watershed (ca. 66°46' N, 65°49' E) underlain by the continuous permafrost (figures 2.1(a) and (c)). The bedrock is mainly comprised of Paleozoic amphibolite and crystal granodiorite (*Mazepa et al.*, 2011). The most common soil in the Polar Urals is mountain-tundra gleysol, and rare inclusion of mountain-tundra turf soils. On the ultramafic massif of Rai-Iz, soils are characterized by neutral and near-neutral pH (*Kataeva et al.*, 2004). According to data collected at Salekhard meteorological station (55 km southeast of the study area, 35 m asl.) between 1883 and 2005., the mean annual, mean monthly minimum (January) and maximum (July) air temperatures are -6.7°, -24.4°, and 13.8°C, respectively. The mean annual precipitation is 600 mm with 50% as snow and sleet. The growing season lasts from mid-June to early August with a mean frost-free period of 94 d. Westerly wind prevails, while the northeasterly wind is also present in early summer, with an average wind speed of 1–3 m s^{-1} and maximum speed observed during the winter season (*Mazepa*, 2005). The wind condition at the study site is shown as wind rose maps in figure 2.2. Data of wind speed and direction are measured from October 2014 to August 2016 at a location around 50 meters to the studied transect.

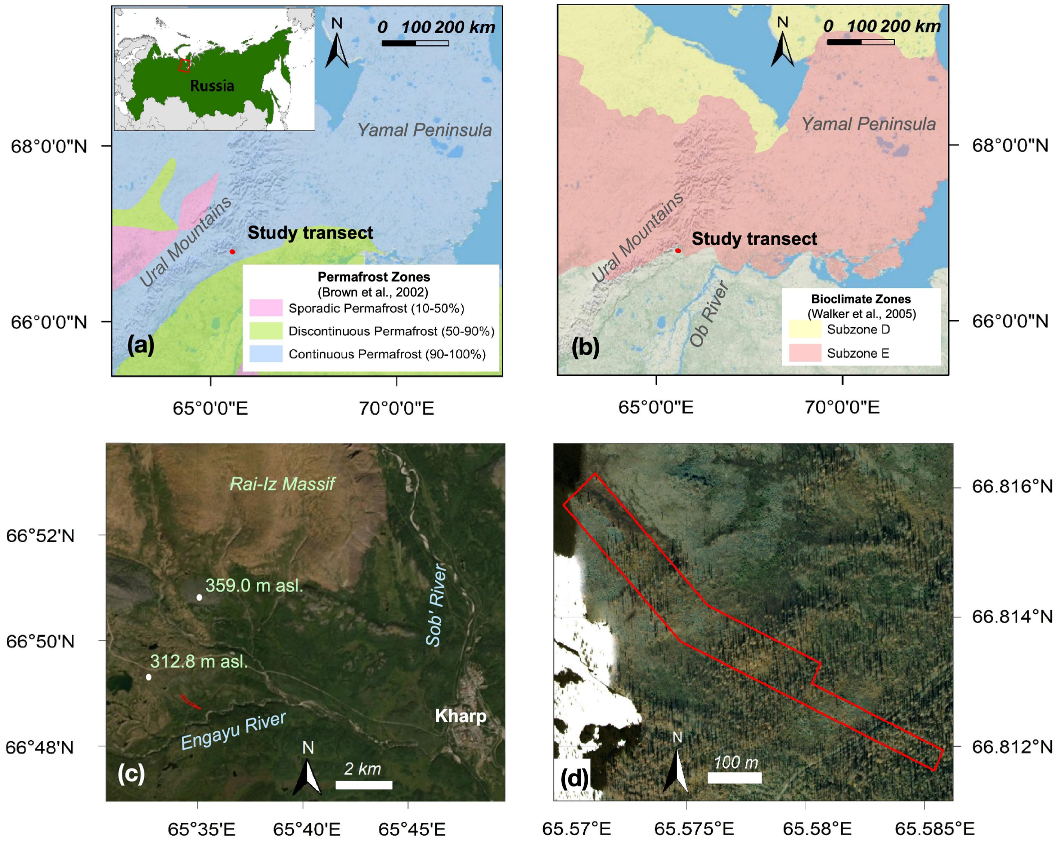

Figure 2.1: Maps of the study area: (a) the permafrost zones (*Brown et al.*, 2002) and (b) the bioclimate subzones (*Walker et al.*, 2005) in the Polar Urals region. (c) The vicinity of the Rai-Iz massif and the Sob' river watershed, where the studied transect (red polygon) is located. Elevations of hilltops near the transect have been displayed. (d) Outline of the studied transect. The permafrost zones in (a) are classified based on the percentage (as in the legend) of landscape underlain by permafrost. In (d), elevation ranges from 194 (bottom-right) to 257 (top-left) m asl.

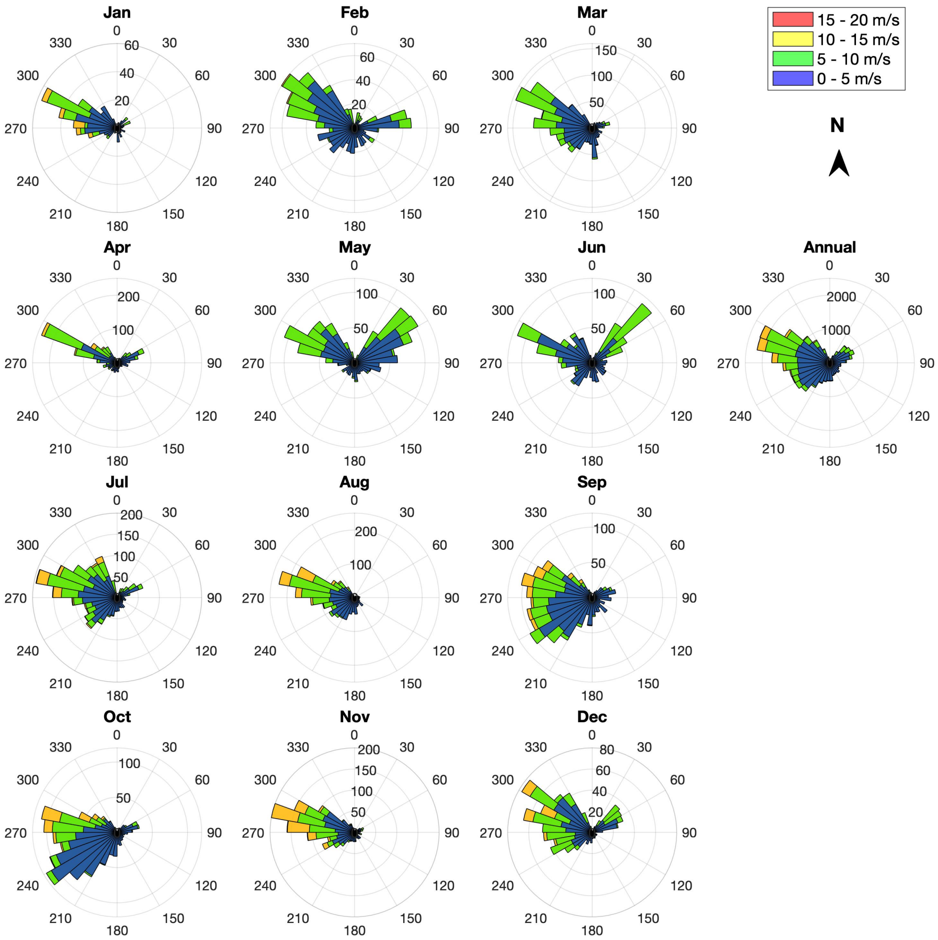

Figure 2.2: The monthly and annual wind rose maps. The azimuth represents the angle to true north in degree, and the radius represents the number of events.

Starting in the early 1960s, six permanent altitudinal transects of 300–1100 m long and 20–80 m wide were established in the eastern foothills of the mountain range for long-term monitoring of spatiotemporal dynamics of alpine forest-tundra and forest-meadow communities. Transects typically start at the upper timberline boundary, where live trees are not yet present, and extend down into the upper limit of closed forest (*Mazepa et al.*, 2011). This study focuses on the earliest ecotone transect within which complete data on tree characteristics, terrain elevation, and snow distribution have been collected for multiple years (figure 2.1(d)). This transect also exhibits varying topography as compared to the other transects. The transect is 860 m long and 40–80 m wide (the total area is 5.6 ha, with the upper left corner at 66°48'57" N, 65°34'09" E) and oriented along the direction of the predominant westerly wind (*Mazepa*, 2005; *Shiyatov and Mazepa*, 2011).

This transect is located close to the southern boundary of the bioclimate subzone E (*Walker et al.*, 2005) as shown in figure 2.1(b). Tree stand composition within the timberline ecotone is relatively simple: larch (*Larix sibirica* Ledeb.) forest-tundra communities dominate in the upper part, while open larch forest with Siberian spruce (*Picea obovata* Ledeb.) and downy birch (*Betula tortuosa* Ledeb.) occur in the lower part of the area (*Shiyatov et al.*, 2005). Dwarf birch (*Betula nana* L.) is abundant in the lower and middle parts of the transect. Names of plant species correspond to the Integrated Taxonomic Information System (*ITIS*, 2021). Historically, boreal trees grow at the fringe of the forest-tundra ecotones in various growth forms (creeping, prostrate, single-stemmed, and multi-stemmed), which reflect adaptations to environmental changes (*Mazepa and Devi*, 2007; *Devi et al.*, 2008). Over 90% of the young trees appearing in the study transect after 1950 are single-stemmed (*Mazepa et al.*, 2011). The studied forest stands and recent vegetation growth have emerged without significant impacts from reindeer grazing, fire, and human activities (such as logging) over at least the last millennium (*Mazepa*, 2005; *Devi et al.*, 2020).

2.2.2 Tree structural and allometric characterization

The records of three census campaigns in 1960, 1999, and 2011 are used in this study to quantitatively assess the changes in the composition, structure, and spatial distribution of the forest-tundra communities. Specifically, detailed mappings of all alive and dead tree locations, and measurements of their allometric and geometric characteristics such as height, crown size, branch position, and diameter at multiple heights were produced during these campaigns. All three censuses predominantly recorded larch with the occasional presence of spruce along the study transect. The numbers of trees for the 1960, 1999, and 2011 censuses are 1853, 1700, and 1398, respectively, after pre-screening.

The mapping criteria between censuses were different. Specifically, many saplings were recorded in the first census but not in the two later censuses. For consistency, we pre-screen the records and consider only trees with a diameter larger than 1 cm for all three censuses after reconstruction. The missing tree diameter data were reconstructed based on an empirical regression model between trunk height and diameter. A power-law model was fitted using a nonlinear least-squares method and is shown in figure 2.3.

Tree diameters were measured at tree base, 10 cm, and 130 cm height (if the stem was sufficiently long). The accuracy of diameter measured at the tree base was affected by ground conditions and stem irregularities, while many measurements at 130 cm were missing since most trees were shorter than this height in the earliest census of 1960. Thus, tree diameters measured at 10 cm height are referred to as the "diameter D" in this analysis.

Sapwood is used in this study as an indicator of tree productivity and total biomass (*Kaufmann and Troendle*, 1981). Sapwood of growing trees conducts water and stores living (parenchyma) cells that can be used for a number of physiological processes including xylem tracheid refilling after the seasonal freeze-thaw cycle (*Mayr et al.*,

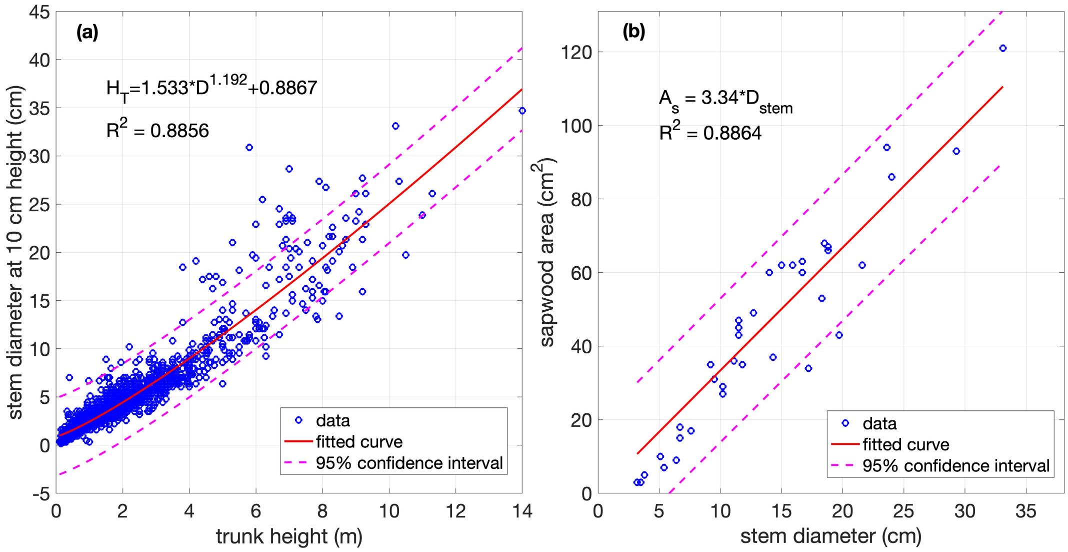

Figure 2.3: (a) A nonlinear fitting model between trunk height H_T and tree diameter D at 10 cm. Measurements from the census of 1960 (most complete) are used to derive the relationship. (b) Linear regression between stem diameter D_{stem} and sapwood area A_S based on 37 samples of larch cross-sectional cuts from sawed trees. The cuts were located at different heights of the tree.

2014). While measuring diameters and heights of trees is relatively straightforward, direct measurements of sapwood area are cumbersome as they can be done accurately only by cutting trees down. However, previous studies show significant correlation between sapwood area and stem diameters for several boreal tree species (*Bond-Lamberty et al.*, 2002). Several larch trees sawn in the 2000s were used to establish a linear regression relationship between sapwood area and stem diameter along the tree trunk (figure 2.3(b)). This linear relationship and the assumption that sapwood changes reflect biomass production make it possible to use the changes in stem diameter as a direct proxy for tree biomass changes.

The allometric relationship between sapwood area and tree diameter is usually exponential (*Vertessy et al.*, 1995; *Pazdrowski et al.*, 2007). The highly linear relationship presented in figure S.1b may be dominated by the young ages and small sizes of the studied tree samples and is site-specific. While sapwood area is a better indicator for biomass for many conifer species, research from *Kajimoto et al.* (2006)

has shown that diameters measured at different height are also good predictors for two larch species and the D-based site-specific allometry can be used as an easy and reliable model to estimate biomass. These studies combining with our site-specific highly linear allometric relationship support our statement to use tree diameter as an indicator for biomass.

After reconstructing the missing tree diameter data using an allometric function, we compute the change of tree diameters $\mathit{\Delta} D$ between censuses as the proxies for the increase of biomass and tree productivity, and then relate them to environmental features. Specifically, to compare changes of tree diameters $\mathit{\Delta} D$ for different periods (i.e. from 1960 to 1999, and from 1999 to 2011), the relative diameter change $\mathit{\Delta} D_{rel}=\mathit{\Delta} D/D_0$, where D_0 is the tree diameter at the first year of the corresponding period, is computed first. $\mathit{\Delta} D_{rel}$ is then normalized by the period of time T (years) to get the annual relative change rate of stem diameter $\dot{D}_{rel} = \sqrt[T]{1+\Delta D_{rel}} - 1$ assuming a constant annual change rate for a given period. This rate was computed only for trees recorded both at the start and end year of the time interval between the two censuses. The changes in tree diameter may also be used as a proxy of tree biomass change through its relationship with sapwood as shown in figure 2.3(b).

2.2.3 Snow distribution

Snow data were collected along the study transect in 1961, 2006, and 2013 in the spring, when snow depth (Z_S) was at its peak by direct measurement, tree marking, and visual assessment. Direct measurements with a snow scale were used in the upper part of the transect, where shallow snow depths and low density of tree stands prevailed. In the middle and bottom parts of the transect, different approaches were used. Whenever possible, the snow level was marked with paint on tree trunks, and the mark height was measured during the summer period of the year. Visual assessment of the snow depth relative to trees located within sight was applied on sections of the transect where trees did not grow and snow was too deep to measure using a snow scale. Spatially irregular measurements were subsequently interpolated using the method of inverse distance weighting to build spatial distributions of snow depth. The relative snow depth, defined as $\hat{Z}_S = (Z_S - Z_{Smin})/(Z_{Smax} - Z_{Smin})$ where Z_{Smax} and Z_{Smin} are the maximum and minimum snow depths within the transect for a given survey year from the three surveys, is shown in figure 2.4.

2.2.4 Topographic analysis

Digital terrain models (DTMs) are needed to capture topography at fine scales, but there are limited areas in the Arctic for which high-resolution DTMs are available. There is a range of digital elevation model (DEM) products that can provide high-resolution elevation maps. However, these DEM products are affected by the presence of surface biomass masking the location of actual topography. To reconstruct the DTM for the transect, we combined a resampled 2 × 2 m resolution grid from a 20 × 20 m regularly spaced field terrain measurements of the transect, and a 2 m resolution DEM product called ArcticDEM (*Porter et al.*, 2018). After co-registering two ground sources (i.e. field measurement and ArcticDEM), we manually selected bare ground grids from ArcticDEM together with the field measurement to extract a high-resolution DTM for the study transect by natural neighbor interpolation (*Tily and Brace*, 2006).

Terrain curvature is one of the fundamental topographic properties and an important driver of hydrologic processes (*Moore et al.*, 1991). Two types of curvature indices, i.e. profile curvature and platform curvature defined by the Environmental Systems Research Institute (*ESRI*, 2021), are used to represent micro-topographic features in this study. Specifically, profile curvature is the curvature of the surface in the direction of the steepest slope, while planform curvature is the curvature in the plane perpendicular to the direction of the steepest slope. For profile curva-

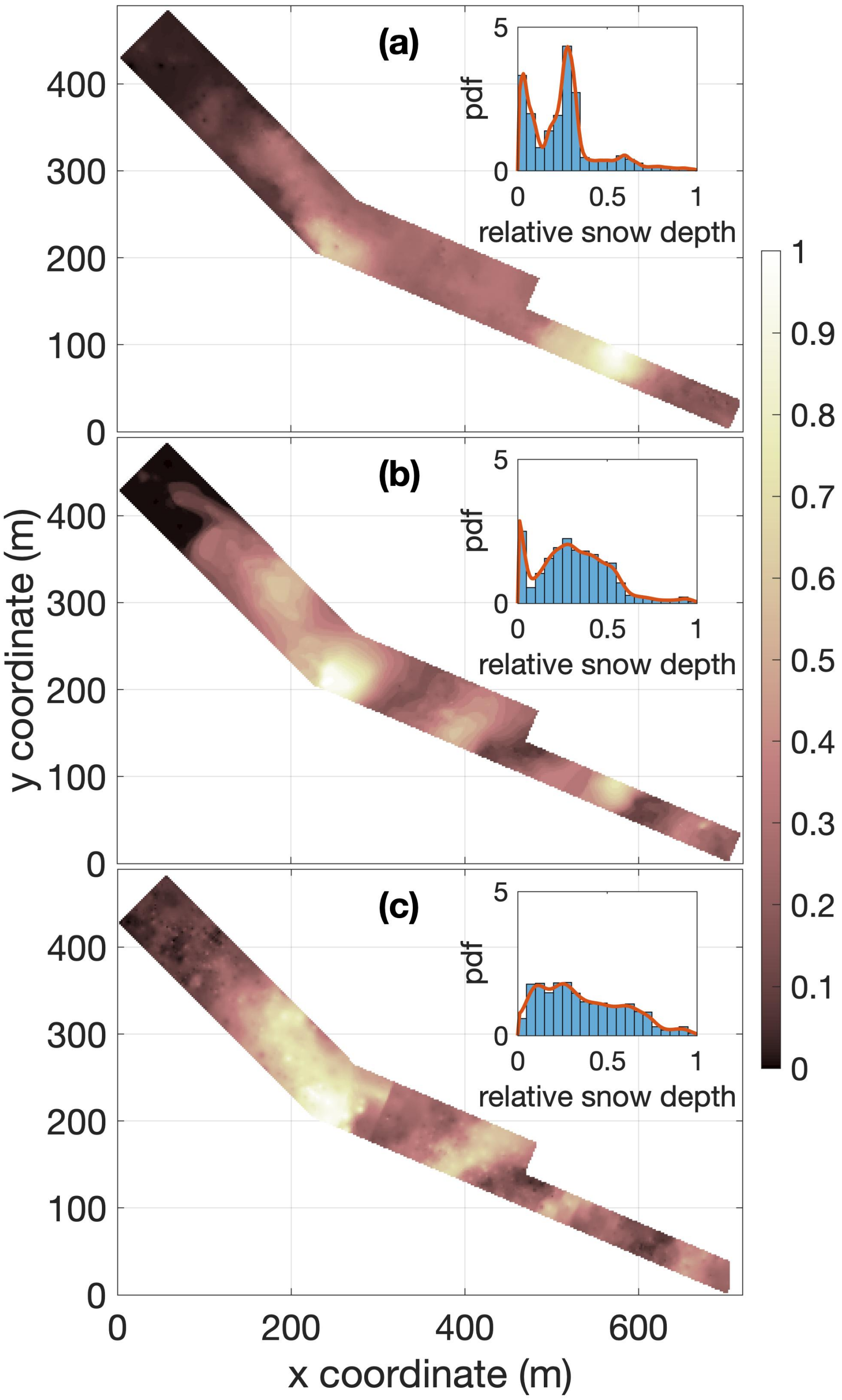

Figure 2.4: The spatial distribution of the relative snow depth in the study transect in 1961 (a), 2006 (b), and 2013 (c). Plot insets show the relative density (bars) and the probability density function of the relative snow depth estimated with kernel density (red curves in the insets) using a Gaussian kernel. The transect was mapped to a relative coordinate frame, where the x and y axes align with the lowermost and leftmost point of the transect, respectively.

ture index, a negative value indicates an upwardly convex or dome-shaped surface, while a positive value represents a concave surface or depression; for planform curvature index, negative and positive values indicate a laterally concave and convex surface, respectively. Note that the opposite sign is used in the definition of planform curvature concavity/ convexity to that of profile curvature. In conventional terrain analysis, the curvature indices are calculated from DTM as the second-order derivative of the elevation map using various geomorphometric methods (e.g. *Zevenbergen and Thorne*, 1987). The curvature indices calculated from the resampled DTM with various resolutions are used to capture topographical features at different scales.

Common algorithms of terrain curvature of a grid use a 3 × 3 cell size moving window and the curvature is computed by using elevations of eight neighboring cells (*Krebs et al.*, 2015; *Tarolli et al.*, 2012). This index therefore only represents the scale at which it is derived. Since the computed curvature indices are scale-dependent (e.g., curvature computed at high-resolution and thus small-scale cannot recognize larger-scale features), it is necessary to consider the effect of the DTM resolution on the topographic features at the scale of interest (*Shary et al.*, 2005; *Shi et al.*, 2007; *Krebs et al.*, 2015). To better capture terrain curvature variations across scales, we first resampled the DTM of the base resolution at 2 m × 2 m to progressively coarser resolutions, up to 20 m × 20 m. We subsequently extracted curvature indices using the 3 × 3 grid cell window for each grid of the resampled DTMs. All manipulations were conducted using ArcMap of ESRI's ArcGIS Desktop software package (*Environmental Systems Research Institute*, 2011).

To attribute topographic locations to niches of higher or lower favorability for tree encroachment, we computed the weighted average curvature index using stem diameter at locations where trees are present as a weight:

$$\bar{\kappa}_p = \frac{\sum D_i \kappa_{i,p}}{\sum D_i}, \tag{2.1}$$

where D_i is the diameter of tree i, $\kappa_{i,p}$ is the curvature index of a terrain grid cell where the tree is located, and the summation is for all trees along the transect. The index p represents either planform or profile curvature index. Noted that $\kappa_{i,p}$ as well as the number of trees within a given grid cell change with resolutions, so the summation is not carried out over the same set of trees for different grid resolutions. The above formulation emphasizes the contributions of curvature indices of terrain locations with higher D. The rationale is that stem diameter is argued to be an indicator of biomass and long-term productivity (section 2.2.2). If there is an enhanced tree growth at topographic locations of specific curvature, it can be elucidated by comparing $\bar{\kappa}_p$ in (2.1) with the average curvature of transect topography as the arithmetic mean of curvature over all grid cells along the transect at different grid resolutions.

2.2.5 Stem density clustering analysis

Spatial point pattern analysis is commonly used to characterize the patterns of the spatial distribution of tree species (e.g. *Condit et al.*, 2000). Most of these analyses use a family of statistics derived from the Ripley's K-function (*Ripley*, 1976) to quantify the species clumping characteristics by comparing them with complete spatial randomness (CSR). In this study, following the approach proposed by *Plotkin et al.* (2002), we apply a density-based spatial clustering method with the presence of noise (i.e. points located too far from other points) (*Ester et al.*, 1996) to capture spatial pattern of tree distribution. The advantage of this method is that it does not require *a priori* number of clusters, such as in k-means analysis (*MacQueen*, 1967), and can identify clusters of arbitrary shape. Clusters of tree locations (referred to as 'points') are derived at different spatial resolutions using the records of the three tree censuses (section 2.2.2). The algorithm requires two parameters: neighborhood search radius (d) and the minimum number (q) of points required to form a cluster located within the search radius. By setting the parameter q to 1 in all analyses, we make the clustering process depend only on the parameter d, which thus represents the maximum distance between any pair of connected trees belonging to the same cluster. Therefore, for any given value of d all stem locations are partitioned into a unique set of clusters.

Tree clusters were formed at different spatial scales by changing the parameter d from 0 to the maximum distance between locations of any two connected trees in the transect (65 m). For a given d, the total number of obtained clusters is denoted as m and the size of an ith cluster as c_i. The sample size, i.e. the total number of trees in the transect, is denoted as $n = \sum_{i=1}^{m} c_i$. The unique arrangement of clusters for a specific value of d can be summarized with the normalized mean cluster size:

$$\hat{c} = \frac{1}{n^2} \sum_{i=1}^{m} c_i^2. \tag{2.2}$$

This variable represents the probability that two randomly chosen tree locations belong to the same cluster (*Plotkin et al.*, 2002). For example, $\hat{c} = 1/n$ when $d = 0$, corresponds to a clustering pattern that each point forms its own cluster; $\hat{c} = 1$ when d equals the maximum pairwise distance between any two connected tree locations, i.e. all trees form a single cluster. The clustering pattern varies for intermediate values of d and can be obtained by recording the concomitant change of $\hat{c}$, known as the mean cluster size curve $\hat{c}(d)$. To make the results for censuses with different sample sizes comparable in the mean cluster size curves, we use the normalized distance $\hat{d}$ defined as

$$\hat{d} = d\sqrt{n/A}, \tag{2.3}$$

where A is the transect area. Spatial patterns of clustering can then be compared with CSR. Specifically, to represent CSR we generated an ensemble set (1000 total) of random tree locations uniformly distributed in the transect. The number of random locations was set to the total number of trees. The difference between $\hat{c}(d)$ curves derived using actual tree locations and CSR locations represents the degree of non-randomness in the tree spatial distribution.

2.3 Results

2.3.1 Stem diameter change

The sampling distribution of the annual relative change rate of stem diameter $\dot{D}_{rel}$ for the two periods between the three censuses (1960–1999 and 1999–2011) is shown in figure 2.5(a). The mode of the distribution increased from 2%/year to 3%/year, indicating accelerated growth in the later period. The right-skewed distribution of the second period implies that the proportion of trees with higher growth rates is larger in the recent period. Figure 2.5(b) shows the relationship between $\dot{D}_{rel}$ and D_0 for the two periods between the three field surveys. The maxima of $\dot{D}_{rel}$ form a clear upper boundary with respect to D_0 for both periods, indicating the upper limit of the annual growth rate, and the results illustrate an increase of these growth rates from the earlier to the later period (red and blue lines). The highest growth rate decreasing with D suggests that larger, older trees grow slower than smaller, younger trees. Growth rates can vary significantly, between just above zero and to a few percent of their stem diameter per year.

2.3.2 Tree locations and terrain curvature indices

Figure 2.6 compares the average, D-weighted terrain curvature index for grid cells with live trees (Eq. (2.1)) and the transect average index (i.e. uses all grid cells), as a function of DTM grid cell resolution. The results indicate scale-dependent differences with respect to the transect average for both planform and profile curvatures, and for all censuses. The difference is significant for DTM grid sizes ranging from 2 to

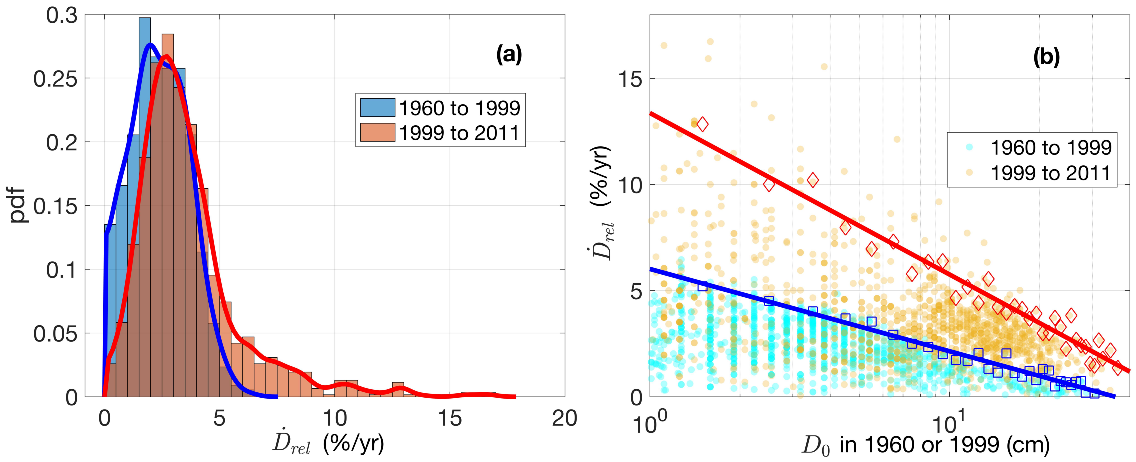

Figure 2.5: (a) The sampling probability density function (pdf) of the annual relative rate of stem diameter change, $\dot{D}_{rel}$ (a proxy for stem growing rate), for the two periods between field surveys: from 1960 to 1999, and 1999–2011. The blue and red curves are the probability density functions estimated using a Gaussian kernel for 1960–1999 and 1999–2011, respectively. (b) A relationship between the $\dot{D}_{rel}$ from 1960 to 1999 (cyan dots) and 1999–2011 (orange dots) and the stem diameter at the start year (D_0) for each period. The logarithmic scale is used for the horizontal axis. Blue (1960–1999) and red (1999–2011) lines are the upper boundaries of the change rate. To construct the boundaries, we calculated 95% percentiles of $\dot{D}_{rel}$ for bins of D_0 (plotted as blue squares and red diamonds) and a linear regression with respect to $\log(D_0)$ at the center of each bin was fitted.

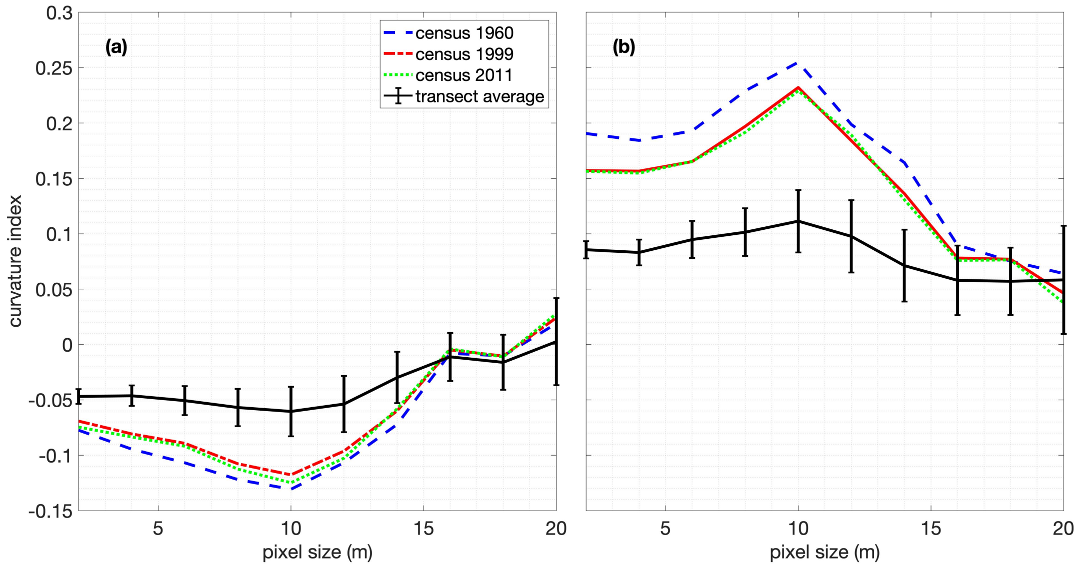

Figure 2.6: Transect curvature indices as a function of the DTM resolution: (a) average profile curvature index and (b) average planform curvature index computed over all locations with live trees, all censuses (colored curves), for different grid cell sizes. Averaged curvature indices estimated using all grids in the transect are shown with black lines; whiskers represent the 95% confidence interval.

14 m, with the largest difference occurring at the grid size of 10 m. This implies the preference for tree growth in terrain curvature is characterized most significantly at this scale (i.e. 25–30 m length scale of landforms, which is the size of the moving window used to compute the curvatures, see section 2.2.4). At locations with live trees, the profile curvature index is more negative, and the planform curvature index is more positive than the overall transect averages, implying that trees tend to cluster at sites with convex topography. The curvatures of the locations with trees and the transect average become indistinguishable for grid cells of sizes greater than 14 m. Since the number of grids along the transect decreases with grid size, this leads to the convergence of the curvature indices to the same value.

2.3.3 Tree cluster dynamics

The resultant cluster size curves for the three censuses are shown in figure 2.7. The sharp transitions in the three curves at various distances $\hat{d}$ (Eq.(2.3)) correspond to the 'continuum percolation' phenomenon (*Meester and Roy*, 1996; *Plotkin et al.*, 2002), i.e. the aggregation of isolated smaller clusters into a larger cluster. The corresponding distance is known as the 'critical distance' or 'percolation threshold'. When $\hat{d}$ is below the threshold, elements in the system of interest are disconnected and form many smaller-sized clusters; above the percolation threshold, the elements are aggregated to form larger clusters. In theory, for true CSR (i.e. if the sample size is infinitely large), the cluster size curve will show a discontinuous transition corresponding to the step function at a critical distance of $\hat{d} \approx 1.2$ (an abrupt transition for the developed CSR curve in figure 2.7 is not discontinuous due to the finite sample size and transect edge effects on the computation of $\hat{d}$).

The cluster size curves constructed using the actual tree locations from all censuses have three transitions at varying critical distance $\hat{d}$. Their differences with respect to the CSR curve imply a non-random spatial tree distribution at different spatial scales. The plateaus of the curves between the abrupt transition in cluster sizes represent the ranges of d within which the change of cluster pattern is insignificant. The number of plateaus (two for this case) indicates the number of scales of nonrandom aggregation (*Plotkin et al.*, 2002). Therefore, one may perform a clustering analysis using d corresponding to a plateau region to represent such a stable distribution at a selected scale. Here we choose the values of normalized distance corresponding to the start of plateaus immediately after the occurrence of 'continuum percolation'. The following clustering analysis uses $\hat{d}$ corresponding to the first plateaus ($\hat{d}_1$) in figure 2.7: 1.24, 1.4, and 1.51 (i.e. the actual distances are $d_1 = 6.7$, 7.9, and 9.4 m) for the 1960, 1999, and 2011 censuses, respectively. For the second plateau ($\hat{d}_2$) in figure 2.7, the normalized distances are $\hat{d}_2$=2.37, 3.63, and 5.58 ($d_2 = 12.8$, 20.5, and 34.7 m). When a cluster pattern corresponding to $\hat{d}_1$ is defined as the 'primary cluster set', then the value of $\hat{d}_1$ is a representative proxy of the distance among trees within any cluster of this set, while $\hat{d}_2$ represents a proxy of the distance between borders of clusters formed by the primary set. As figure 2.7 shows, over time, relatively smaller changes of the distribution pattern occurred within the primary cluster set, as reflected by the comparable magnitudes of $\hat{d}_1$ across the three censuses. However, the distances between the cluster borders become larger, as illustrated by the increase of $\hat{d}_2$, particularly for the census of 2011 (mainly due to the mortality of trees at cluster borders).

Figure 2.8 shows the spatial distributions of tree clustering for the two critical distances $\hat{d}$. Transitions of cluster patterns can be detected by comparing plots in the same row. For example, when the distance changes from d_1 =6.7 m to d_2 =12.8 m (the census of 1960), the number of clusters decreases significantly from 102 to 21 due to cluster aggregation. The same 'continuum percolation' phenomenon is observed for the other census data of 1999 and 2011. Comparing plots in the same column reveals that the number of primary clusters decreases over time. This phenomenon can be explained by the disappearance of cluster 'noise', i.e. those trees that form their own cluster. Notice that for the primary cluster set (the left panel of figure 2.8), the sum of cluster sizes of the three largest clusters is over 80% of the total number of surveyed trees (not shown).

Figure 2.9 illustrates the temporal change of the distribution of the largest cluster from 1960 to 2011. In 1960, there were many small trees scattered in this cluster, and larger trees were mainly located in the middle part. A large fraction of small trees previously located in the lower part disappeared in later years, while those in the upper part grew bigger and likely seeded new trees that appeared in the transect in 2011.

Concurrently with the change of the tree distribution pattern, the area with high snow depth had shifted from the lower part to the middle part of this cluster. Similar transitions of areas with high snow depth are also apparent in the regions outside this cluster (not shown). The area of high snow accumulation shifted from the lower part

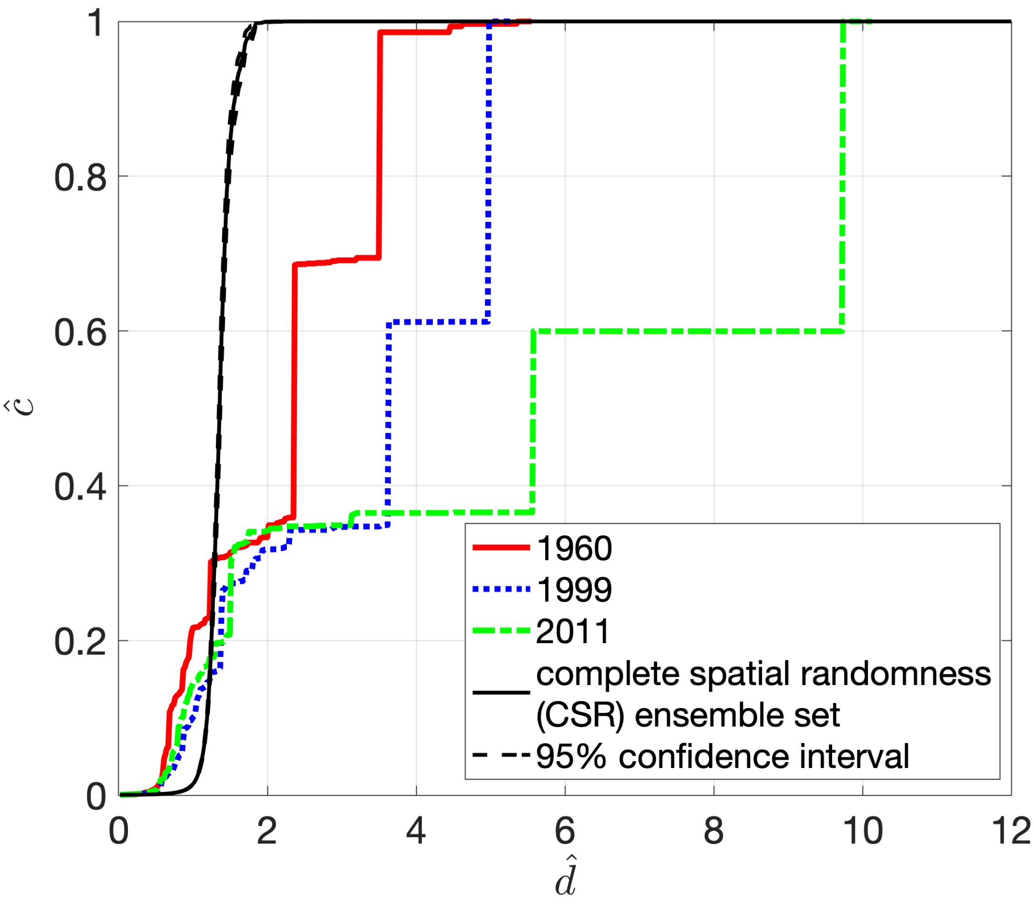

Figure 2.7: Normalized mean cluster size curves for the three censuses (colored lines) and the complete spatial randomness (CSR) ensemble set. The CSR curve (black solid line) shows the curve averaged over 1000 members of the CSR ensemble set; the black dashed curves represent the 95% confidence interval for this set. The normalized mean cluster size, $\hat{c}$, represents the probability that two randomly chosen tree locations fall in the same cluster. While $\hat{d}$ is the normalized maximum pairwise distance between two connected tree locations (see text for details). The plateaus of the curves between the abrupt transition in cluster sizes represent the ranges of $\hat{d}$ within which the change of cluster pattern is insignificant. The $\hat{d}$ corresponding to the start of plateaus is the critical distance.

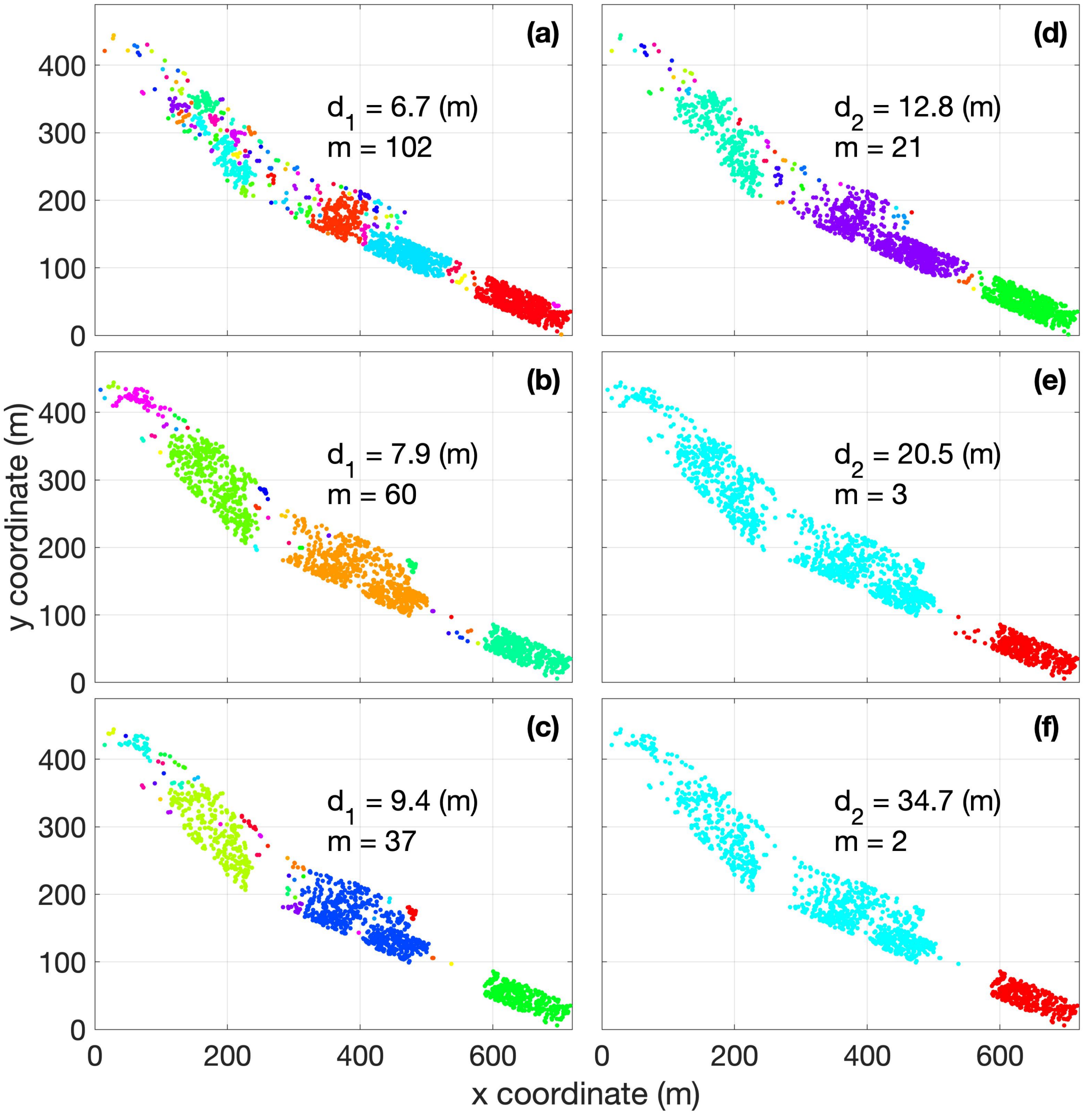

Figure 2.8: Spatial representation of tree clusters corresponding to the two critical distances $\hat{d}_1$ (left panel, the first critical distance in figure 2.7, the actual distance is d_1) and $\hat{d}_2$ (right panel, the second critical distance in figure 2.7, the actual distance is d_2); m is the corresponding number of clusters. The first, second, and third row corresponds to the censuses in 1960, 1999, and 2011, respectively. Circles represent tree locations, and each cluster is plotted with a unique color.

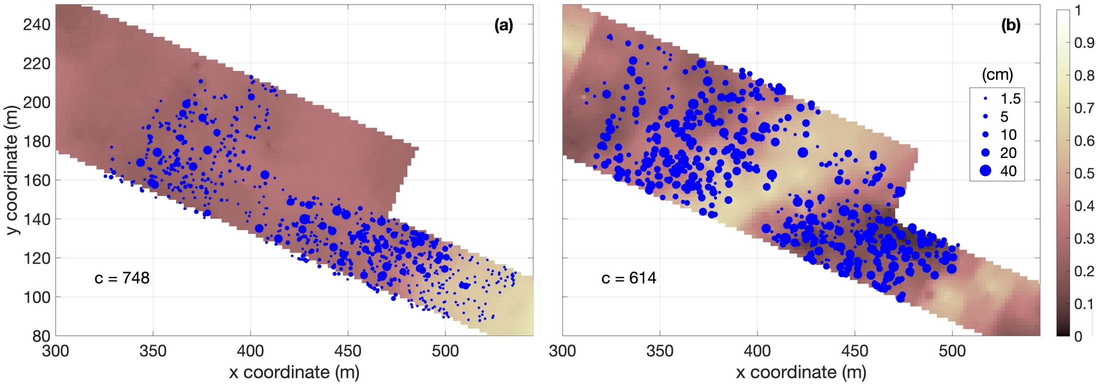

Figure 2.9: Change of spatial pattern of the largest tree cluster (blue circles) identified for the first 'percolation threshold' in figure 2.8 (left panel) in (a) 1960 and (b) 2011. Circle sizes are scaled by the tree diameters. The color map indicates the distribution of relative snow depth $\hat{Z}_S$ in (a) 1961 and (b) 2013. Letter '*c*') in the panels denotes the cluster size, i.e. the number of trees within this cluster.

to the middle-upper part of the transect, coinciding with the growth and densification of trees in clusters at the top.

2.4 Discussion and conclusion

Forest expansion in the Polar Urals region is thought to be induced by the increased summer temperature and winter precipitation (*Shiyatov et al.*, 2005; *Devi et al.*, 2008, 2020). Analyses of treeline dynamics are usually based on repeated landscape photographs and remotely sensed imagery (*Kharuk et al.*, 2006; *Beck and Goetz*, 2011; *Frost and Epstein*, 2014), dendrometric survey on tree remnants (*Briffa et al.*, 2008; *Shiyatov and Mazepa*, 2011, 2015), and treeline models (*Kaplan and New*, 2006; *Paulsen and Körner*, 2014). However, long-term records of allometric characteristics of living trees, as well as snow distribution back to the beginning of the satellite era are rare at tree stand level. This study analyzed spatiotemporal characteristics of the encroachment process of Siberian larch into timberline ecotone of Polar Urals between 1960 and 2011. The unique data sets spanning over 50 years allowed us to carry out tree-level characterizations of landscape favorability for the encroachment process and other features of tree expansion.

Various factors influence the spatial distribution of trees in the timberline ecotone. Considering the small scale of this study that spans a ∼1 km-long gentle mountain hillslope, climatic factors such as temperature and precipitation as well as topographic features such as elevation and slope aspect are relatively uniform along the transect. In spite of the spatially homogeneous climate, our analysis reveals a consistent signature of terrain favorability with respect to the curvature indices derived from the data for all censuses. Trees tend to grow on convex or 'bulged' surfaces along the study transect, indicating the preference of trees to be located at sites with divergent surface characteristics. Considering topographic features as proxies for hydrologic dynamics, these locations correspond to well-drained sites. Convex topography may also have negative effects on tree growth, such as extreme windswept and the likelihood of summer drying, impacting seedling establishment (*Holtmeier and Broll*, 2005). The preference of terrain locations with convex curvature is an overall outcome of local-scale hydrologic dynamics combined with more favorable climate conditions (e.g. increasing winter precipitation), dominating over concomitant disadvantages. The subsurface conditions are also important factors affecting tree distribution. *Grigorieva and Moiseev* (2018) have pointed out the importance of soil moisture on the survival and growth of larch seedlings. *Shiryaev et al.* (2019) presented a strong correlation between larch forest advance and increase of soil thaw depth and soil microbiota activities. However, such data are currently unavailable and extremely laborious to collect in the field. Future work augmenting terrain data with soil textural analysis and characterization of the subsurface thermal state is needed.

An increase in the maximum stem growth rate during the most recent period has been observed in this study. Such an increase is expected as the climate warms and the environment becomes more favorable for tree growth, as had been inferred from long-term dendrochronology studies (*Briffa et al.*, 2008, 2013), field measurements (*Devi et al.*, 2008), as well as remotely sensed Normalized Difference Vegetation Index analysis (*Zeng et al.*, 2013; *Berner et al.*, 2020). The variation in growth rates can be explained by the influence of multiple environmental factors such as soil moisture (*Myers-Smith et al.*, 2015, 2020), soil nutrient availability and non-homogeneity of subsurface freeze-thaw processes (e.g. *Sullivan et al.*, 2015) for trees located at different locations. Quantitative assessment of their contributions is the topic of future research.

This study shows a clear interaction between changes in spatial patterns of tree locations and the snow distribution shifts. Many small trees and saplings of the largest cluster died between 1960 and 1999 (section 2.3.2) likely caused by high snow accumulation due to the wind drift—a phenomenon well pronounced at the field transect. Specifically, when snow is redistributed by wind, high roughness obstacles such as trees reduce wind speed at their lee sides leading to leeward snow accumulation (*Hiemstra et al.*, 2002). In 1960, this accumulation was observed at the bottom of the largest cluster (figure 2.9(a)). Although snow cover corresponds to a favorable microclimate (*Hagedorn et al.*, 2014), excessive snow accumulation results in a longer snowmelt period and a shorter growing season reducing total photosynthetic carbon gain and survival chances for small trees. As trees grew and became larger at locations with snow depths that did not limit the duration of the growing season (e.g. the upper part of the largest cluster), the leeward snow accumulation moved towards upslope (figure 2.9(b)). The increased tree presence in the upper areas in 2011 resulted in a new area of high snow accumulation in the middle part of the cluster. Such a change in surface roughness condition likely made the lower part of the transect even more favorable for trees to grow as indicated by a large number of new stems growing in the lower part. Such growth patterns at different locations with a phase shift signal the effect of spatial teleconnection. This effect has been displayed within and among clusters through interactions between new tree growth and snow distribution changes, which partially explain the heterogeneity of trees encroaching upslope areas along this mountain transect. Since snow distribution can change drastically between consecutive years, one obvious limitation is the incompleteness of the snow record. Improved understanding of the spatial interactions of snow with the tree encroachment process requires a more continuous record of snow distribution at fine-scale resolutions (i.e. at 1 m).

Treeline is considered a critical indicator of climate warming (*Harsch and Bader*, 2011). Forest extent is predicted to increase 55% in the Arctic with a 42% decrease in tundra area under projected warming in the next few decades (*Kaplan and New*, 2006). The spatial variability of tree expansion, however, indicates the importance of other drivers of vegetation shifts besides summer warming (*Elmendorf et al.*, 2012). The results presented in this study imply the necessity to take microtopography and snow conditions into account when predicting future timberline ecotone dynamics. Specifically, two thresholds of snow depth need to be determined as a prior: the lower limit, above which trees can survive winter stresses such as wind abrasion and frost, and the upper limit, under which the growing season is long enough for trees to grow. Admittedly, however, these challenges are difficult to overcome as snowpack modeling can be a major source of error in treeline dynamics models due to the lack of precipitation data in mountain or remote regions (*Paulsen and Körner*, 2014). The complexities displayed in the interaction between trees and snow distribution implies that until this dynamic process approaches a steady state, prediction of treeline dynamics at regional scales will remain difficult and subject to uncertainties.

In conclusion, the study in this chapter reveals the relationship between the spatial pattern of encroachment of tree stands and environmental factors including topography and snow cover. Structural and allometric characteristics of trees, along with terrain elevation and snow depth collected along a transect 860 m long and 80 m wide were used for the analysis. Results of the topographic analysis suggest that trees tend to cluster in areas with convex surfaces. The clustering analysis also indicates that the patterns of tree locations are linked to snow distribution. Records from the earliest campaign in 1960 show that trees lived mainly at the middle and bottom of the transect across the areas of high snow depth. As trees expanded uphill following a warming climate trend in recent decades, the high snow depth areas also shifted upward creating favorable conditions for recent tree growth at locations that were previously covered with heavy snow. The identified landscape signatures of increasing tall vegetation, and the effects of microtopography and snow may facilitate the understanding of treeline dynamics at larger scales.

CHAPTER III

Uncertainty Informed Surface Ground Heat Flux Reconstruction

The content of this chapter is prepared to be submitted to the *Journal of Geophysical Research: Earth Surface* by the time this dissertation is submitted. Zhou. W., Ivanov. V. and Sheshukov. A. conceived of the presented idea. Zhou. W. processed the data, performed the computations, analyzed the results, and prepared the draft manuscript. Ivanov. V. supervised the findings of this work. Wang. J., Zhu. M., Sheshukov. A., and Sargsyan. K. developed the methodologies. Xu. D. suggested proper downscaling methods. Zhang. L. pre-processed the data. Ivanov. V. and Sheshukov. A. collected the field data with support from Sokolov. A., Valdayskikh. V., and Mazepa. V. All authors discussed the results and contributed to the final manuscript.

3.1 Introduction

Surface ground heat flux (G_S) is defined as the diffusive heat flow across the boundary that separates the bottom of the atmosphere and below ground. It is a crucial component of the land surface energy balance, especially in the high-latitude regions as the energy supply of phase change of water (*Lunardini*, 1981; *Halliwell* *and Rouse*, 1987; *Ochsner and Baker*, 2008; *El Sharif et al.*, 2019). G_S controls the dynamics of freeze and thaw and permafrost degradation in Arctic regions (*Biskaborn et al.*, 2019). However, the quantification of G_S has received relatively less attention in the past (*Heusinkveld et al.*, 2004; *Wu et al.*, 2020) and the impact of climate change on G_S is not well understood (e.g., *Nicolsky and Romanovsky*, 2018).

Direct measurement of G_S can be made using heat flux plates (e.g., HuksefluxUSA Inc. HFP01 Heat Flux Sensor) buried at shallow depths in the soil to ensure tight contact between the plate surface and soil substrate (surficial deployment is impossible due to radiative energy and precipitation impact) and combined with soil temperature measurements in the layer above the plate (*El Sharif et al.*, 2019). However, such measurements are sparse in remote Arctic regions because they are costly and logistically challenging. Large-scale data products from global climate models (GCMs) are often used to project permafrost thermal states in the future under different climate scenarios. However, they usually do not provide direct outputs of G_S for either historical or future periods. Instead, G_S needs to be derived as the residual term of the surface energy balance equation, using net radiation (R_n), sensible (H), and latent (λE) heat flux. However, this is only valid for snow-free surfaces. For snow-covered surfaces, the change in the internal energy of the snowpack (heat associated with heating/cooling and water phase change) is required to derive G_S. Yet, the internal energy change of snowpack is usually not an explicit output from most GCMs.

Several analytical methods have been developed to derive G_S. Traditional methods for modeling G_S require multi-depth data of soil temperature with soil properties such as bulk density, specific heat, and thermal conductivity (*Fuchs and Tanner*, 1968a; *Kimball et al.*, 1976; *Horton and Wierenga*, 1983; *Cobos and Baker*, 2003; *Ochsner et al.*, 2006; *Ochsner and Baker*, 2008). For example, in the case of homogeneous soils without water phase change, an analytical solution of soil heat flux and temperature at any depth can be derived from a harmonic analysis using measured soil temperature (*Heusinkveld et al.*, 2004). However, its accuracy decreases as measurement depth increases. Under the same assumptions of homogeneity and water state condition for a semi-infinite soil column, an analytical solution of G_S can be derived assuming sinusoidal surface temperature boundary condition (*Gao et al.*, 2010). Analogously, G_S can also be derived from the Green's function solution (*Wang and Bou-Zeid*, 2012) or from the soil heat flux at a deeper depth through 'damping depth' (*Leuning et al.*, 2012). *Wang and Bras* (1999) developed an analytical relationship between soil temperature and ground heat flux at the same depth through a half-order derivative (HOD) method (see section 3.2 for details). Using this analytical solution (referred to as the HOD model hereafter), ground heat flux can then be derived from surface temperature time series given the thermal inertia of the bulk soil. In general, all analytical solutions require surface soil temperature as input.

Representative surface temperature measurements using ground-based instruments are difficult to obtain and not commonly available in the Arctic region. For example, measurements using thermometers are subject to substantial sampling errors caused by poor contact between the sensor and the surface and the disturbances from incident direct sun and diffuse sky radiation (*Fuchs and Tanner*, 1968b). The measurements of contactless infrared temperature sensors are affected by calibration errors and other technical problems such as optical path blocking by snowflakes or vegetation elements (Apogee Instruments, Inc., SI-400 infrared radiometer manual). Instead, to monitor the subsurface conditions, temperature sensors are usually placed at some depth below ground, in narrow and deep boreholes. As a result of such a monitoring approach, long-term time series of multi-level soil temperatures have been obtained from boreholes in the Arctic regions. However, existing analytical solutions cannot be used to *directly* derive surface ground heat flux from belowground borehole temperature measurements.

Besides temperature, soil thermal properties such as thermal inertia and thermal diffusivity also need to be specified in model formulation(e.g., the HOD model). These model parameters can be *a priori* determined given the soil composition, i.e. soil texture, soil organic material content, and soil water content (*Johansen*, 1977; *Farouki*, 1981; *Lawrence and Slater*, 2008). However, measurements of soil texture needed to derive such thermal properties are still lacking in the Arctic, and due to the intrinsic environmental heterogeneity and their small-scale nature of sampling (e.g., 'fist-size' auger soil samples), they can be extremely noisy. The emergent questions are then: can effective magnitudes of these soil thermal properties be objectively evaluated using only soil temperature measurements? Corollary, if the effective soil properties can be inversely estimated from observed physical states (such as soil temperature), can surface ground heat flux be reconstructed and its uncertainty be quantified?

The goal of this chapter is the development and demonstration of efficacy and uncertainty-informed accuracy of a novel methodological approach for estimating the surface ground heat flux over different seasons from soil temperature and soil moisture series (if available) measured at a range of depths below the ground surface. Three methodological "blocks" are applied to achieve this goal. The analytical HOD model is used to derive surface ground heat flux from soil temperature (and moisture if available). A physically based numerical model (*Sheshukov and Egorov*, 2002) is used to simulate non-isothermal heat-moisture dynamics and the resultant temperature and water phase series from the surface flux G_S. Finally, modern advancements in probabilistic learning and uncertainty quantification (UQ) machinery (*Sivia and Skilling*, 2006; *Sraj et al.*, 2014; *Dwelle et al.*, 2019) are used to derive the probabilistic description of soil thermal properties used in the analytical model. Accurate G_S derived from field measurements or obtained as GCM outputs during snow-free seasons are used as a reference for the parameter inference procedure. In the more typical case when G_S data are not available, temperature series at different depths of the soil column (such as in a borehole) are used to estimate flux G_S, which is subsequently verified by comparing the simulations of soil temperature from a physically based numerical model with temperature measurements.

3.2 HOD solution of ground heat flux and soil temperature

3.2.1 Analytical model formulation

The HOD model formulation is briefly described here with more details in *Wang and Bras* (1999). For a one-dimensional semi-infinite soil column, the heat transfer process follows the classical thermal diffusion equation with no phase change assuming constant thermal diffusivity and a uniform initial temperature profile. Ground heat flux at depth z_2 ($G(z_2, t)$) can be derived from soil temperature at depth z_1 ($T_S(z_1, t)$),

$$G(z_2, t) = \frac{I}{\sqrt{\pi}} \int_0^t \exp\left[-\frac{(z_2 - z_1)^2}{4\alpha(t - \tau)}\right] \frac{dT_S(z_1, \tau)}{\sqrt{t - \tau}}, \tag{3.1}$$

where I is the soil thermal inertia (W^2 s m^{-4} K^{-2}), α is the soil thermal diffusivity (m^2 s^{-1}), t is time (s), and τ is the the integration dummy variable (s). With a positive direction for depth defined downward from the surface, this solution requires $0 \leq z_1 \leq z_2$. Analytical solutions for soil temperatures are

$$T_S(z_2, t) = T_S(z_2, 0) + \int_0^t \operatorname{erfc}\left[\frac{z_2 - z_1}{2\sqrt{\alpha(t - \tau)}}\right] dT_S(z_1, \tau), \tag{3.2}$$

and temperature at z_2 is related to ground heat flux at z_1 by

$$T_S(z_2, t) = T_S(z_2, 0) + \frac{1}{I\sqrt{\pi}} \int_0^t \exp\left[-\frac{(z_2 - z_1)^2}{4\alpha(t - \tau)}\right] \frac{G(z_1, \tau)d\tau}{\sqrt{t - \tau}}, \tag{3.3}$$

$\tau = 0$ in Eqs. (3.1)-(3.3) is set at times when G_S is as close to zero as possible. In the equations above, soil thermal inertia, I, is defined as

$$I = \sqrt{kc_v}, \tag{3.4}$$

and soil thermal diffusivity, α, is defined as

$$\alpha = \frac{k}{c_v}, \tag{3.5}$$

where k is the soil thermal conductivity (W m^{-1} K^{-1}), and c_v is the soil volumetric heat capacity (J m^{-3} K^{-1}).

3.2.2 Parameter separation

Thermal properties of the soil are affected by its water content, in liquid or solid phase. In the original formulation of the HOD model, thermal inertia is assumed to be time-invariant. Water content and phase-dependent thermal inertia vary at seasonal scale in the Arctic. The top soil layer in the Arctic, known as the active layer, also undergoes seasonal thawing and freezing cycles. The mineral soil in the active layer is usually near saturation for much of the year and may become unsaturated in summer, while the shallow layer of well-drained and organic-rich soil at the top of the active layer may be desiccated even in winter (*Kane et al.*, 1989; *Hinzman et al.*, 1991; *Hinkel et al.*, 2001). The effects of soil water content and phase are accounted for by separating the parameters previously considered as time-invariant into time-dependent and time-independent parts.

Following a linear mixing model, soil thermal conductivity k and volumetric heat capacity c_v are computed using the set of equations as in *Lawrence and Slater* (2008).

For soil thermal conductivity:

$$k = K_e k_{sat} + (1 - K_e) k_{dry}, \tag{3.6}$$

$$K_e = \begin{cases} \log S_r + 1, & T_S \geq T_f \\ S_r, & T_S < T_f \end{cases}, \tag{3.7}$$

$$k_{sat} = k_s^{1-\theta_{sat}} k_{liq}^{\frac{\theta_{liq}}{\theta_{liq}+\theta_{ice}}\theta_{sat}} k_{ice}^{(1-\frac{\theta_{liq}}{\theta_{liq}+\theta_{ice}})\theta_{sat}}, \tag{3.8}$$

$$S_r = \frac{\theta_{liq} + \theta_{ice}}{\phi} \approx \frac{\theta_{liq} + \theta_{ice}}{\theta_{sat}}, \tag{3.9}$$

For soil volumetric heat capacity:

$$c_v = c_h + \theta_{liq} c_{liq} + \theta_{ice} c_{ice}, \tag{3.10}$$

Names and units of the parameters used in Eqs. (3.6)-(3.10) are listed in table 3.1. The parameters k_{liq}, k_{ice}, c_{liq}, and c_{ice} are certain and treated as constants. The parameter θ_{sat} is determined by finding the maximum value in the observed series of $\theta_{liq} + \theta_{ice}$ and is considered to be equal to porosity, i.e., $\phi \approx \theta_{sat} = \theta_{liq} + \theta_{ice}$. Three time-independent parameters can be identified from these equations: c_h, k_{dry}, and k_s. Other parameters are time-varying and depend on the temporal evolution of soil liquid water and ice content. After soil thermal properties are computed using available records of soil water content and the time-invariant properties are determined in any appropriate fashion, the surface ground heat flux can then be calculated using these derived soil thermal properties and the HOD model.

3.3 1-D freeze-thaw numerical model for saturated soil

The simulation of soil freezing and thawing processes is an active research topic in cold climate region studies due to their important roles in climate change, permafrost and hydrologic regimes, and cold region infrastructure sustainability (*Matzner and Borken*, 2008; *Painter et al.*, 2013; *Walvoord and Kurylyk*, 2016; *Painter et al.*, 2016; *Hjort et al.*, 2022). Various numerical models have been developed to solve the coupled heat and moisture dynamics of frozen soil (*Harlan*, 1973; *Taylor and Luthin*, 1978; *Mu and Ladanyi*, 1987; *Zhang et al.*, 2007; *Sheshukov and Nieber*, 2011; *Painter and Karra*, 2014). In this study, we adopt the solution for saturated frozen soil (*Sheshukov and Egorov*, 2002).

The one-dimensional (1-D) conservation equation of heat transfer can be described as:

$$\frac{\partial}{\partial t}(c_v T_S - \rho_i \phi L_f S_i) = \frac{\partial}{\partial z}(k \frac{\partial T_S}{\partial z}), \tag{3.11}$$

where the variables are defined in Table 1. The advective heat flux associated with the bulk movement of water in the soil in Sheshukov and Egorov's original formulation is neglected here considering that the heat transfer is dominated by conductive heat flux and phase change. In other words, the water content in the soil is viewed as nearly immobile. Under the fully saturated condition,

$$S_w + S_i = 1, \tag{3.12}$$

When soil temperature T is above the freezing point T_f, $S_w = 1$.

The freezing-thawing phenomena are considered to be similar to the drying-wetting phenomena (*Koopmans and Miller*, 1966) for which liquid water pressure head (ψ_w) and ice pressure head (ψ_i) and relative saturation S_w are described by a

Table 3.1: Nomenclature of parameters used to derive soil thermal properties

Parameter	Name	Unit	Model*
K_e	Kersten number (the weight for k_{sat} in Eq. (3.6) for soil thermal conductivity calculation)	-	A
k_{dry}	dry soil thermal conductivity	W m^{-1} K^{-1}	A
k_{sat}	saturated soil thermal conductivity	W m^{-1} K^{-1}	A
S_r	degree of saturation for total water content	-	A
$T_f = 273.15$	freezing temperature of water	K	A, N
k_s	soil solid thermal conductivity	W m^{-1} K^{-1}	A
θ_{sat}	saturated volumetric water content	-	A
θ_{liq}	liquid water content	-	A
θ_{ice}	ice content	-	A
$k_{liq} = 0.57$	thermal conductivity of liquid water	W m^{-1} K^{-1}	A
$k_{ice} = 2.3$	thermal conductivity of ice	W m^{-1} K^{-1}	A
c_h	volumetric heat capacity of soil constituents excluding water and ice	J m^{-3} K^{-1}	A
$c_{liq} = 4.2 \times 10^6$	volumetric heat capacity of liquid water	J m^{-3} K^{-1}	A
$c_{ice} = 2.0 \times 10^6$	volumetric heat capacity of ice	J m^{-3} K^{-1}	A
ϕ	porosity	-	A, N
$\rho_i = 918.0$	density of ice	kg m^{-3}	N
$\rho_w = 1000.0$	density of liquid water	kg m^{-3}	N
$L_f = 3.34 \times 10^5$	latent heat of freezing for water	J kg^{-1}	N
S_i	degree of saturation for ice	-	N
S_w	degree of saturation for water	-	N
b	reciprocal of pore size distribution index in the B-C model	-	N
ψ_s	air entry pressure head in the B-C model	m	N
$g = 9.8$	gravitational acceleration	m s^{-2}	N

*A: analytical model. N: numerical model.

water retention function $P_c(S_w)$ adopted from the Brooks and Corey (B-C) model (*Brooks and Corey*, 1966):

$$\psi_w - \psi_i = \omega P_c(S_w) = \omega \psi_s S_w^{-b}. \tag{3.13}$$

Assuming no frost heave presence, ψ_i is set to be zero in this study. The coefficient ω is the ratio of specific surface energies at the ice-water interfaces (*Koopmans and Miller*, 1966). Its value depends on soil saturation conditions and soil texture and is set to 1.0 in this study. Finally, if one assumes equilibrium at the ice-water interface, the generalized Clapeyron equation (*Miller*, 1980) is :

$$\frac{\psi_w}{\rho_w} - \frac{\psi_i}{\rho_i} = \frac{L_f}{\rho_w g T_f} T. \tag{3.14}$$

Under proper boundary and initial conditions, the temperature and moisture profile of the soil column can be simulated simultaneously by solving Eqs. (3.11)-(3.14). Uncertainties arise with this adapted model when the assumptions of restricted parameter settings (e.g., $\psi_i = 0$, $\gamma = 1.0$) and full saturation are applied. However, using the parameter inference process with the embedded uncertainty quantification machinery as described below, these uncertainties are implicitly included in the inferred parameter distributions.

The numerical model described above is validated using the experimental results provided by *Jame and Norum* (1980) and the details are presented in Appendix A.

3.4 UQ machinery

Mechanistic models including the HOD model are developed based on physical principles with certain assumptions and approximations to describe the behavior of physical systems. Both model inputs (spatial or temporal fields of states) and the parameters (e.g., proxies or intrinsic medium properties) have uncertainties. The UQ machinery has been widely used to quantify the uncertainty and objectively infer values for model parameters (*Malinverno*, 2002; *Ellison*, 2004; *Sraj et al.*, 2014; *Sargsyan et al.*, 2014, 2015; *Dwelle et al.*, 2019). In this section, several components of UQ are briefly introduced and an overview of the workflow is described at the end.

3.4.1 Bayesian inference

If one defines a 'forward problem' as the simulation of quantities of interest (QoIs) using a mechanistic model $\mathcal{M}$ with prescribed model parameter set $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \ldots, \gamma_r)$, where r is the number of parameters in $\mathcal{M}$, an 'inverse problem' can be also formulated: given available observations on QoIs and a set of modeled results using the uncertain set $\boldsymbol{\gamma}$, can one infer the *most likely* values in the parameter set $\boldsymbol{\gamma}$, given the need to match the simulated QoIs with observations?

The Bayesian inference is a widely used approach to the inverse problem (e.g., *Hou and Rubin*, 2005; *Sraj et al.*, 2014; *Sargsyan et al.*, 2015; *Dwelle et al.*, 2019). In general, let $\boldsymbol{Y}$ be a vector of available QoI data (from either observations or model simulations), and γ be the vector of uncertain model parameters. Then modeled outputs can be assumed to represent the QoI data::

$$\boldsymbol{Y} = \mathcal{M}(\boldsymbol{\gamma}), \tag{3.15}$$

The cornerstone of the Bayesian inference of model parameters is Bayes' theorem:

$$\Pi(\boldsymbol{\gamma}|\boldsymbol{Y}) \propto L(\boldsymbol{Y}|\boldsymbol{\gamma}) Pr(\boldsymbol{\gamma}), \tag{3.16}$$

where $Pr(\boldsymbol{\gamma})$ is the prior distribution of the set $\boldsymbol{\gamma}$ constructed from *a priori* information about the parameters (see Sec. 3.6.1); $L(\boldsymbol{Y}|\boldsymbol{\gamma})$ is the likelihood function quantifying the probability of producing observations on the QoIs by $\mathcal{M}$ given the set $\boldsymbol{\gamma}$. The likelihood function encapsulates assumptions about the distribution of discrepancy between the model and data. Assuming the discrepancy between $\boldsymbol{Y}$ and modeled results with N parameter sets $\gamma^n(n = 1, 2, \ldots, N)$, i.e., $\epsilon^n = \boldsymbol{Y} - \mathcal{M}(\boldsymbol{\gamma}^n)$, to be independent and identically distributed random variables following their corresponding marginal distributions Pr_ϵ^n, then:

$$L(\boldsymbol{Y}|\boldsymbol{\gamma}) = \Pi_{n=1}^{N} Pr_\epsilon^n(\boldsymbol{Y} - \mathcal{M}(\boldsymbol{\gamma}^n)), \tag{3.17}$$

Furthermore, we assume that the errors ϵ^n follow Gaussian distribution, $\epsilon^n \sim N(0, \sigma^2)$, leading to a likelihood function:

$$L(\boldsymbol{Y}|\boldsymbol{\gamma}) = \frac{1}{\sqrt{2\pi\sigma^2}} \Pi_{n=1}^{N} \exp\left(-\frac{(\boldsymbol{Y} - \mathcal{M}(\boldsymbol{\gamma}^n))^2}{2\sigma^2}\right), \tag{3.18}$$

where σ^2 can be determined as the variance of the QoI observational data noise.

Finally, assuming independent prior distributions, the final joint conditional posterior distribution can be written as:

$$\Pi(\boldsymbol{\gamma}|\boldsymbol{Y}) \propto \frac{1}{\sqrt{2\pi\sigma^2}} \Pi_{n=1}^{N} \exp\left(-\frac{(\boldsymbol{Y} - \mathcal{M}(\boldsymbol{\gamma}^n))^2}{2\sigma^2}\right) Pr(\gamma_1) Pr(\gamma_2) \ldots Pr(\gamma_r), \tag{3.19}$$

The joint posterior distribution in Eq. (3.19) is typically unavailable in closed form, and therefore we resort to Markov Chain Monte Carlo (MCMC) method (*Hastings*, 1970) that produces samples of the posterior distribution. Specifically, MCMC samples are filtered by removing the first 25% values of the MCMC parameter sequence and choosing every fifth value in the remaining set to reduce correlations in the chain. The reduced set of posterior parameter values is then used to evaluate the full distribution or its marginals using kernel density estimation (KDE).

3.4.2 Surrogate modeling

The sampling process in section 3.4.1 requires repeated simulations of the mechanistic model $\mathcal{M}$ for computing the likelihood function, which is computationally expensive. A feasible technique to improve computational efficiency is to replace the mechanistic model with a less complex, 'reduced-order', or 'surrogate' model. In this study, polynomial chaos (PC) surrogate construction is applied and briefly introduced below with more details in *Le Maître and Knio* (2010). The surrogate model ($\mathcal{M}^{PC}$) is built by expressing the outputs of a mechanistic model as a truncated polynomial expansion of the uncertain model parameters:

$$\boldsymbol{Y}(\boldsymbol{x},t) \approx \mathcal{M}(\boldsymbol{\gamma},\boldsymbol{x},t) \approx \mathcal{M}^{PC}(\boldsymbol{\gamma},\boldsymbol{x},t) = \sum_{l=0}^{P} c_l(\boldsymbol{x},t)\Psi_l(\boldsymbol{\gamma}), \tag{3.20}$$

where $c_l(\boldsymbol{x},t)$ are the polynomial coefficients, $\Psi_l(\boldsymbol{\gamma})$ multivariate orthogonal polynomials, and $P+1 = (r+o)!/(r!o!)$ is the number of terms in the PC expansion, where r is the number of stochastic dimensions (equals to the number of uncertain parameters), and o is the highest polynomial order of the expansion Eq. 3.20. The polynomial basis satisfy:

$$\int_{-\infty}^{\infty} \Psi_i(\boldsymbol{\gamma})\Psi_j(\boldsymbol{\gamma})\pi(\boldsymbol{\gamma})d\boldsymbol{\gamma} = 0 \quad \text{for all } i \neq j, \tag{3.21}$$

where $\pi(\boldsymbol{\gamma})$ is a weight function. The coefficients $c_l(\boldsymbol{x},t)$ are estimated from the multiple runs of the model $\mathcal{M}$ using random samples of the parameter set $\boldsymbol{\gamma}$. Model outputs of QoI are used to compute the coefficients $c_l(\boldsymbol{x},t)$ by using Gaussian quadrature projection (*Smolyak*, 1963), regression (*Berveiller et al.*, 2006; *Blatman and Sudret*, 2008, 2011), or Bayesian approaches such as Bayesian Compressive Sensing (which is applied in this study) (*Babacan et al.*, 2009; *Sargsyan et al.*, 2014).

3.4.3 Karhunen-Loève expansion of a stochastic process

A given QoI output of a mechanistic model $\mathcal{M}(\boldsymbol{\gamma}, \boldsymbol{x}, t)$ can be considered to represent a stochastic process, given the uncertainty associated with the parameter vector $\boldsymbol{\gamma}$. The size of the QoI output from the model $\mathcal{M}$ depends on the characteristics of the input domain (i.e., the physical dimensions of space and time duration of simulation) and is often large, consisting of $\mathcal{O}(10^3 - 10^5)$ spatial and temporal values. According to the Karhunen-Loève (KL) expansion (*Karhunen*, 1947; *Loeve*, 1948; *Le Maître and Knio*, 2010), a stochastic process (in this case, a QoI output of $\mathcal{M}(\boldsymbol{\gamma}, \boldsymbol{x}, t)$) can be represented by a sequence of uncorrelated variables ξ_j as:

$$\mathcal{M}(\boldsymbol{\gamma}, \boldsymbol{x}, t) = \bar{\mathcal{M}}(\boldsymbol{x}, t) + \sum_{j=1}^{\infty} \xi_j(\boldsymbol{\gamma}) \sqrt{\mu_j} \varphi_j(\boldsymbol{x}, t), \tag{3.22}$$

where $\bar{\mathcal{M}}(\boldsymbol{x}, t)$ is the mean of model outputs, $\xi_j(\boldsymbol{\gamma})$ are uncorrelated random variables called the KL coefficients, μ_j and $\varphi_j(\boldsymbol{x}, t)$ are the eigenvalues and eigenfunctions of the covariance matrix $cov[\mathcal{M}(\boldsymbol{\gamma}, \boldsymbol{x_i}, t_i), \mathcal{M}(\boldsymbol{\gamma}, \boldsymbol{x_j}, t_j)]$, respectively (the pairs $(\boldsymbol{x_i}, t_i)$ and $(\boldsymbol{x_j}, t_j)$ are points in the space and time domain). In practice, the first J eigenvalues and their corresponding eigenfunctions are selected with J be the minimum that satisfies the inequality $\sum_{j=1}^{J} \mu_j / \sum_{j=1}^{\infty} \mu_j > 0.99$ to capture 99% of variance contained in $\mathcal{M}(\boldsymbol{\gamma}, \boldsymbol{x}, t)$ output of QoI. A relatively small number J as the upper summation limit in Eq. (3.22) can make the estimation of the right-hand-side part of Eq. (3.22) sufficiently close to $\mathcal{M}(\boldsymbol{\gamma}, \boldsymbol{x}, t)$. Furthermore, if the mechanistic model is believed to accurately simulate the behavior of a real system and therefore can perform well with respect to the observations of QoIs, the eigenvalues and eigenfunctions inferred from Eq. (3.22) can be viewed as representative of the process governing the variation of QoIs as well. Therefore, observational data $\boldsymbol{Y}(\boldsymbol{x}, t)$ of QoIs can be also projected to the same eigenspace defined by μ_j and $\varphi_j(\boldsymbol{x}, t)$ of Eq. (3.22) as

$$\boldsymbol{Y}(\boldsymbol{x}, t) = \bar{\mathcal{M}}(\boldsymbol{x}, t) + \sum_{j=1}^{\infty} \eta_j \sqrt{\mu_j} \varphi_j(\boldsymbol{x}, t), \tag{3.23}$$

where η_j, the KL coefficients for $\boldsymbol{Y}(\boldsymbol{x}, t)$, are computed from Eq. (3.23). The observations $\boldsymbol{Y}(\boldsymbol{x}, t)$ can be approximately estimated by setting J as the upper summation limit in Eq. (3.23), i.e., using only the largest J eigenvalues and their corresponding eigenfunctions and KL coefficients η_j. The identified J is usually orders of magnitude smaller than the original size of the QoI output from the model $\mathcal{M}$ (*Ivanov et al.*, 2021). The mechanistic model parameters can then be inferred following the Bayesian inference approach: by using J coefficients η_j as “observations” Y that need to be reproduced by the model $\mathcal{M}(\boldsymbol{\gamma})$, given the uncertain coefficients $\xi_j(\boldsymbol{\gamma})(j = 1, 2, \ldots, J)$ that depend on the vector of uncertain model parameters $\boldsymbol{\gamma}$ (i.e., as in Sec. 3.4.1). In other words, the likelihood function described in Eq. (3.18) is now approximated by:

$$L(\boldsymbol{Y}|\boldsymbol{\gamma}) \approx L(\eta_j|\boldsymbol{\gamma}) = \frac{1}{\sqrt{2\pi\sigma^2}} \Pi_{j=1}^{J} \Pi_{n=1}^{N} \exp\left(-\frac{(\eta_j - \xi_j(\boldsymbol{\gamma}^n))^2}{2\sigma^2}\right), \tag{3.24}$$

An advantage here is that if the original QoI dimension is very large (e.g., $\mathcal{O}(10^3 - 10^5)$), a considerably smaller number J (e.g., $\mathcal{O}(10^0 - 10^1)$) of QoIs is obtained using this approach. Another advantage of this framework is that it allows the use of QoIs highly correlated in time and/or space.

As a result, a PC surrogate model can be built for each $\xi_j(\boldsymbol{\gamma})$ as:

$$\xi_j(\boldsymbol{\gamma}) \approx \xi_j^{PC}(\boldsymbol{\gamma}) = \sum_{l=0}^{P} c_{jl} \Psi_l(\boldsymbol{\gamma}), \tag{3.25}$$

Combing Eq. (3.25) and (3.22) leads to a joint KL-PC expansion for the mechanistic model:

$$\mathcal{M}(\boldsymbol{\gamma}, \boldsymbol{x}, t) \approx \bar{\mathcal{M}}(\boldsymbol{x}, t) + \sum_{j=1}^{J} \sum_{l=0}^{P} c_{jl} \Psi_l(\boldsymbol{\gamma}) \sqrt{\mu_j} \varphi_j(\boldsymbol{x}, t), \tag{3.26}$$

Only J surrogate models are needed, which thus substantially improves the computational performance of the UQ machinery (*Ivanov et al.*, 2021).

3.4.4 Overall workflow of UQ

The general procedure of parameter inference using UQ is described below. (i) The prior distributions of each parameter (i.e., the elements of the set $\boldsymbol{\gamma}$) are determined from *a priori* knowledge. (ii) Parameter values of the set $\boldsymbol{\gamma}$ are randomly generated from their priors N times, forming a set of N parameter vectors $\boldsymbol{\gamma}^n, n = 1, 2, \ldots, N$. (iii) N forward simulations of QoIs are conducted using the model $\mathcal{M}(\boldsymbol{\gamma}, \boldsymbol{x}, t)$ with $\boldsymbol{\gamma}^n$; such QoI outputs are called the *inference set.* (iv) The space-time dimensions of the QoIs from $\mathcal{M}(\boldsymbol{\gamma}, \boldsymbol{x}, t)$, and the data $\boldsymbol{Y}(\boldsymbol{x}, t)$ are reduced using the KL expansions in Eq. (3.22) and (3.23) resulting in independent variables $\xi_j(\boldsymbol{\gamma})$, and η_j respectively. (v) A PC surrogate model $\xi_j^{PC}(\boldsymbol{\gamma})$ is built to mimic each jth $\xi_j(\boldsymbol{\gamma})$ as in Eq. (3.25), leading to a joint KL-PC expansion in Eq. (3.26). (vi) In Eq. (3.15)-(3.19), the QoI observational data array $\boldsymbol{Y}$ is replaced by the set of J values of η_j and the model output on QoIs $\mathcal{M}(\boldsymbol{\gamma}, \boldsymbol{x}, t)$ is replaced by the joint KL-PC expansion, as developed in the right-hand-side of Eq. (3.26). (vii) The joint posterior distribution for the set $\boldsymbol{\gamma}$ in Eq. (3.19) is derived using the MCMC sampling. (viii) The maximum *a posteriori* (MAP) parameter vector $\boldsymbol{\gamma}^{MAP}$ is estimated from the mode of the multivariate joint posterior distribution, thereby inferring the effective parameter values (i.e., the elements of the set $\boldsymbol{\gamma}^{MAP}$). (ix) The marginal posterior distributions for each of the parameters of the set are developed as described in Sec. 3.4.1. (x) The final visual diagnostic is made by comparing the estimation of the QoI using the joint

KL-PC expansion in Eq. (3.26) with $\boldsymbol{\gamma}^{MAP}$ and the observation data with posterior predictive bounds representing the result uncertainty.

3.4.5 Posterior predictive construction

After the maximum *a posteriori* (MAP) parameter vector $\boldsymbol{\gamma}^{MAP}$ is derived, the QoI observation data $\boldsymbol{Y}(\boldsymbol{x}, t)$ can be estimated by

$$\boldsymbol{Y}(\boldsymbol{x}, t) \approx \mathcal{M}(\boldsymbol{\gamma}^{MAP}, \boldsymbol{x}, t) \approx \bar{\mathcal{M}}(\boldsymbol{x}, t) + \sum_{j=1}^{J} \xi_j^{PC}(\boldsymbol{\gamma}^{MAP}) \sqrt{\mu_j} \varphi_j(\boldsymbol{x}, t). \tag{3.27}$$

The variance of this estimation is coming from two sources: the variance of $\boldsymbol{\gamma}$ based on its posterior distribution (pushed-forward uncertainty, denoted as PF), and the data noise (DN) in the eigenspace (i.e., variance for η_j). The combined variance of PF+DN is known as the posterior predictive (PP) of $\mathcal{M}(\boldsymbol{\gamma}^{MAP}, \boldsymbol{x}, t)$. If the data noise is provided in terms of data variance σ^2, then PD can be computed by

$$PP = var[\sum_{j=1}^{J} (\xi_j^{PC}(\gamma') + \sigma\varepsilon) \sqrt{\mu_j} \varphi_j(\boldsymbol{x}, t)], \tag{3.28}$$

where γ' is the model parameter vector selected from the posterior distribution, and $\varepsilon \sim N(0, 1)$ is the random perturbation term. In this study, σ is fixed and determined by comparing the KL coefficients for observation data (η_j's) and the inference set (ξ_j's). PP can be visualized in the format of percentile bounds (5-95% in this study) of the KL-PC estimation ensemble constructed by randomly sampling γ' and ε from their distribution and plugging them in the following formula,

$$\bar{\mathcal{M}}(\boldsymbol{x}, t) + \sum_{j=1}^{J} (\xi_j^{PC}(\boldsymbol{\gamma'} + \sigma\varepsilon) \sqrt{\mu_j} \varphi_j(\boldsymbol{x}, t). \tag{3.29}$$

The constructed percentile bounds are named the posterior predictive bounds.

3.5 Surface ground heat flux reconstruction

3.5.1 Flux G_S data are available during snow-free season

Data on G_S can sometimes be obtained from field measurements but such data are mostly reliable during snowpack-free seasons. During a snow season, measurements can become noisy due to water phase changes in the vicinity of the heat flux plate and convective heat flux related to infiltration caused by snowmelt. Flux G_S can also be inferred from typical GCM outputs based on the surface energy balance (Sec. 3.1). However, as estimates of the change of the snowpack internal energy are never provided, this approach cannot be applied to derive G_S during snowpack seasons. In this case, the available G_S data during the snowpack-free season can be used as a reference to infer the parameters representing soil thermal properties in the HOD model. Once the parameters have been estimated, the continuous heat flux G_S can be reconstructed using near-surface temperature time series (more accurately measured and more readily available) following the procedure described below.

The prior distributions of the HOD model parameter set ($\boldsymbol{\gamma} = (c_h, k_{dry}, k_s)$) obtained using the parameter separation method are first derived from available information on soil properties such as SoilGrids, a digital soil mapping product based on a global compilation of soil profile data (*Hengl et al.*, 2017). Details for building prior distributions are provided in Sec. 3.6.1. After N arrays of the set $\boldsymbol{\gamma}$ are generated based on their prior distributions, a suite of N series of ground heat flux (G_1^{HOD}) at a shallow depth z_1 (usually not exceeding few cm and it was 5$\sim$10 cm in this study) are simulated using the time-series of near-surface soil temperature and moisture (if available) at the same depth z_1 using Eq. (3.1). In order to get a close approximation for the surface ground heat flux, the heat storage and water phase change transformations in the top soil layer between z_0 and z_1 should be accounted for. Because depth z_1 is deliberately chosen to be very shallow, ground heat flux $\tilde{G}_S$ is approximately proportional to G_1^{HOD}, i.e., with a scaling factor accounting for the effect of heat storage between the surface and z_1. To scale the N series of G_1^{HOD}, N values of a are sampled from a uniform prior distribution. Available series of the surface ground heat flux (from observations or GCM model estimates) for all snowpack-free seasons are then concatenated into a single series. They are used as available data on target QoIs (G_S^{TGT}) and the similarly concatenated simulated series of the heat flux $\tilde{G}_S$ are used as the inference set in the UQ machinery (see Sec. 3.4) to infer a joint posterior distribution of the HOD model parameter set $\boldsymbol{\gamma}$ and the scaling factor a. MAP estimates $\boldsymbol{\gamma}^{MAP}$ and a^{MAP} and the corresponding marginal posterior distributions are inferred from the joint distribution. Finally, the surface ground heat flux for both snowpack and snowpack-free seasons can be reconstructed with the HOD model using soil temperature and moisture (if available) data series at z_1 and the inferred MAP parameter estimates: $\tilde{G}_S^{MAP} = a^{MAP} G_1^{MAP}$, where $G_1^{MAP} = G_1^{HOD}(\boldsymbol{\gamma}^{MAP})$.

3.5.2 Flux G_S data are unavailable

When data or estimates on heat flux G_S data are unavailable, soil temperature $T_{S,2}$ at depth z_2 (below the depth z_1 at which soil temperature series must be available) is used as the target QoI (i.e., $T_S^{TGT} = T_{S,2}$). Similar to the process described in Sec. 3.5.1, the prior distributions for the parameter set $\boldsymbol{\gamma}$ are constructed first. After that, a suite of N soil temperature series at z_2 is simulated using Eq. (3.2) as the inference set (T_S^{HOD}). They are obtained using N arrays of the set $\boldsymbol{\gamma}$ derived from the prior distributions and soil temperature and moisture (if available) at the shallower depth z_1. The series T_S^{TGT} and T_S^{HOD} at z_2 are further used in the UQ machinery to derive the joint posterior distribution and the best parameter value vector $\boldsymbol{\gamma}^{MAP}$, which thus represents soil properties between z_1 and z_2. Soil heat flux at z_1 can then be simulated with the HOD model with $\boldsymbol{\gamma}^{MAP}$ (i.e., resulting in the flux series G_1^{MAP}).

Similar to the methodology in 3.5.1, a scaling factor a needs to be applied to approximate the surface ground heat flux from the flux at z_1: $\tilde{G}_S = aG_1^{MAP}$. Instead of inferring a as described in Sec. 3.5.1, a scaling factor a^{RMSE} is selected simply by minimizing the root mean square error (RMSE) between the available data on $T_{S,1}$ and modeled $T_{S,1}^{MAP}$ (i.e., the temperature series at z_1 estimated with Eq. (3.3) using aG_1^{MAP}). The final reconstructed surface ground heat flux is obtained as $\tilde{G}_S^{MAP} = a^{RMSE}G_1^{MAP}$.

Since there are no measurements available on G_S, there is no direct way to verify the approximated flux $\tilde{G}_S^{MAP}$. First, we note that the posterior joint distribution of the parameter set $\boldsymbol{\gamma}$ can yield a probabilistic expression of the surface ground heat flux. That is, instead of a single $\tilde{G}_S^{MAP}$ flux series inferred from the set $\boldsymbol{\gamma}^{MAP}$ and a^{RMSE}, an ensemble of realizations can be generated and thus provide a temporal evolution of the probabilistic structure of surface ground heat flux. Second, the estimated $\tilde{G}_S^{MAP}$ can be corroborated by setting it as the top boundary condition for the physically based numerical model described in Sec. 3.3 and comparing simulated soil temperatures at multiple depths with observations. As the parameters of the numerical model cannot be inferred directly from the set $\boldsymbol{\gamma}^{MAP}$ of the HOD model, they need to be inferred separately using the UQ machinery (see S4). The parameter inference may compensate for the $\tilde{G}_S^{MAP}$ flux inaccuracies but this can be deemed acceptable as long the inferred parameters are within a physically realistic range.

3.5.3 Soil moisture data are unavailable

When soil moisture data are unavailable, the HOD model parameters are determined assuming saturated soil with a single phase of water, i.e., $\theta_{liq} = \theta_{sat}$, $\theta_{ice} = 0$ for $T_S > 0$°C and $\theta_{liq} = 0$, $\theta_{ice} = \theta_{sat}$ for $T_S < 0$°C. For this case, the parameter k_{dry} is no longer needed in the computation, and the parameters to be inferred in the set $\boldsymbol{\gamma} = (c_h, k_s, \theta_{sat})$.

3.6 Application

Three types of data are used in this study to test the performance of the developed process of heat flux G_S reconstruction: First, we use soil temperature, moisture, and snowpack-free season G_S from in-situ field measurements at a shallow depth. Second, we use outputs from GCM simulations of soil temperature and moisture of the topmost soil layer and snow-free season G_S derived from the surface energy balance. Third, we use soil temperatures measured at two depths from a borehole site. An illustration of the general workflow for these applications is illustrated in Fig. 3.1.

3.6.1 Parameter prior distribution setup

In order to compute the posterior distributions for parameters using Baye's rule (Eq. (3.16)), one needs to provide the parameter prior distributions. In general, the priors can be built based on some existing (a priori) knowledge about the parameters used in the model, both quantitative (e.g., an empirical distribution of soil bulk density from field samples) and qualitatively (e.g., the material density must be positive). The selection of prior distribution remains a controversial topic because it relies on the experience and knowledge of the modeler (*Dwelle et al.*, 2019). In practice, common distributions such as uniform and Gaussian are usually selected to construct prior parameter distributions.

In Sec. 3.5.1, the uncertain parameters required to be inferred in inverse estimation are the soil thermal properties (i.e., c_h, k_{dry}, and k_s). Since field measurements are not available for such properties, information from the digital soil mapping (i.e., SoilGrids, 250 m grid cell resolution (*Hengl et al.*, 2017)), is used to build prior distributions. Specifically, soil sand (%soil), and clay (%clay) contents are downloaded for a $2°\times2°$area covering a considered site location. Then c_h, k_dry, and k_s are computed

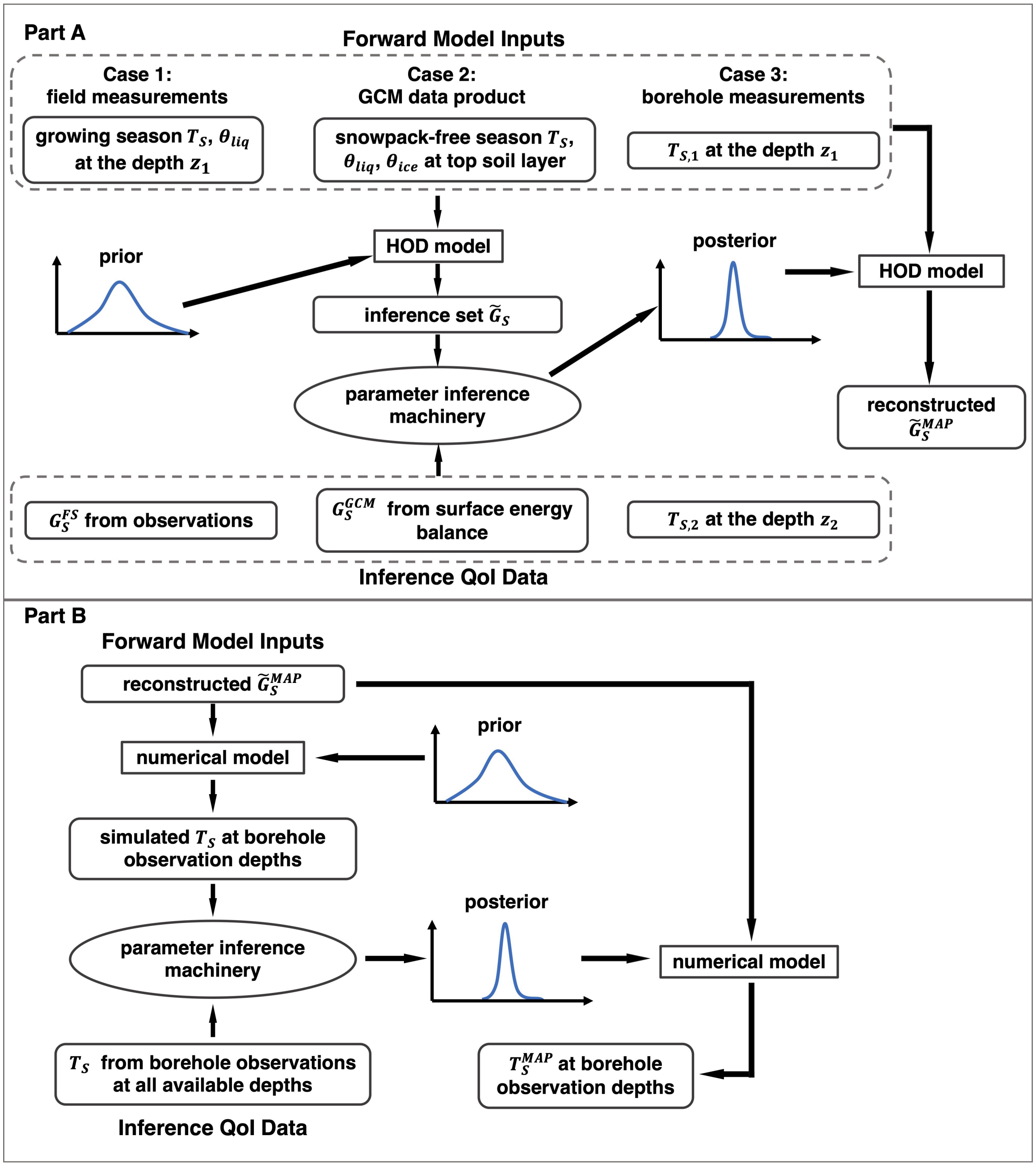

Figure 3.1: Part A: an overview of the heat flux G_S reconstruction workflow for the three cases presented in Sec. 3.6. Part B: corroboration for the reconstructed $\tilde{G}_S^{MAP}$ in case 3.

for each cell following *Lawrence and Slater* (2008):

$$c_h = [(1 - f_{om})c_{s,mine} + f_{om}c_{s,om}](1 - \theta_{sat}), \tag{3.30}$$

$$c_{s,mine} = \frac{2.128(\%sand) + 2.385(\%clay)}{\%sand + \%clay} \times 10^6, \tag{3.31}$$

$$k_{dry} = (1 - f_{om})k_{dry,mine} + f_{om}k_{dry,om}, \tag{3.32}$$

$$k_{dry,mine} = \frac{0.135\rho_d + 64.7}{2700 - 0.947\rho_d}, \tag{3.33}$$

$$\rho_d = 2700(1 - \theta_{sat,mine}), \tag{3.34}$$

$$k_s = (1 - f_{om})k_{s,mine} + f_{om}k_{s,om}, \tag{3.35}$$

$$k_{s,mine} = \frac{8.8\%sand + 2.92\%clay}{\%sand + \%clay}, \tag{3.36}$$

$$f_{om} = \frac{\theta_{sat} - \theta_{sat,mine}}{\theta_{sat,om} - \theta_{sat,mine}}, \tag{3.37}$$

$$\theta_{sat,mine} = 0.489 - 0.00126\%sand, \tag{3.38}$$

where $c_{s,mine}$, and $c_{s,om}$ are volumetric heat capacity of soil solid (J m^{-3} K^{-1}); f_{om} is

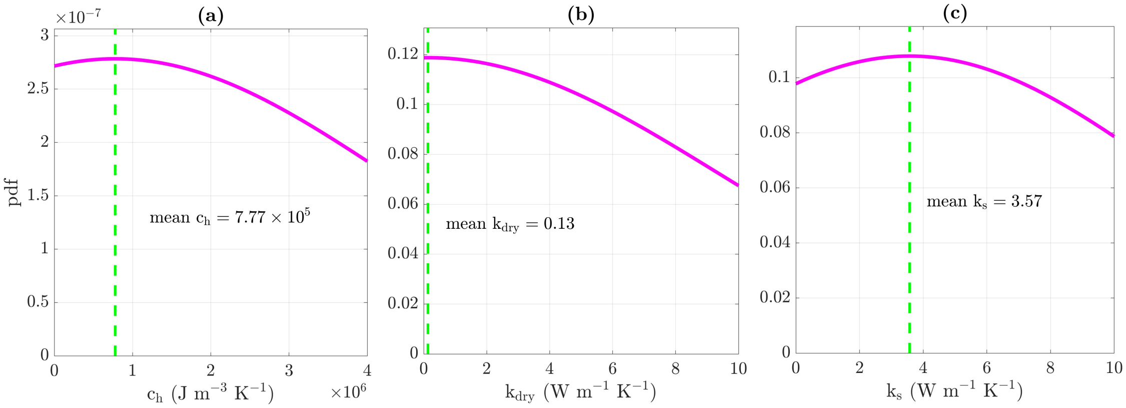

Figure 3.2: The truncated Gaussian prior distributions (magenta curves) for c_h, k_{dry}, and k_s with the mean derived from the SoilGrids product (green dashed lines) for the field site location (as described in Sec. 3.6.2.

the fraction of organic material in the soil layer; $\theta_{sat,mine}$, $\theta_{sat,om}$ (set to 0.9) are the saturated water content; $k_{dry,mine}$, and $k_{dry,om}$ are the dry thermal conductivity (W m^{-1} K^{-1}); ρ_d is the bulk density of mineral soil (kg m^{-3}); $k_{s,mine}$, and $k_{s,om}$ are thermal conductivity of soil solid (W m^{-1} K^{-1}). The subscript *mine* refers to mineral soil, and *om* to soil organic material. According to *Farouki* (1981), some typical values are set as constant: $c_{s,om} = 2.5 \times 10^6$ J m^{-3} K^{-1}, $k_{dry,om} = 0.05$ W m^{-1} K^{-1}, $k_{s,om} = 0.25$ W m^{-1} K^{-1}. The mean and standard deviation of c_h, k_{dry}, and k_s are then computed for all the grid cells.

As shown in Fig. 3.2 and 3.3, the truncated Gaussian type of prior distribution is used for c_h, k_{dry}, and k_s for the field site and the Marre-Sale borehole location, combining the information of mean and standard deviation values derived from SoilGrids and ranges obtained for actual materials (i.e., the horizontal axis bounds in Fig. 3.2 and 3.3). For the construction of the prior distributions, the standard deviations calculated from the SoilGrids product were enlarged by 3×10^3 times given the uncertainties inherent to the digital data product.

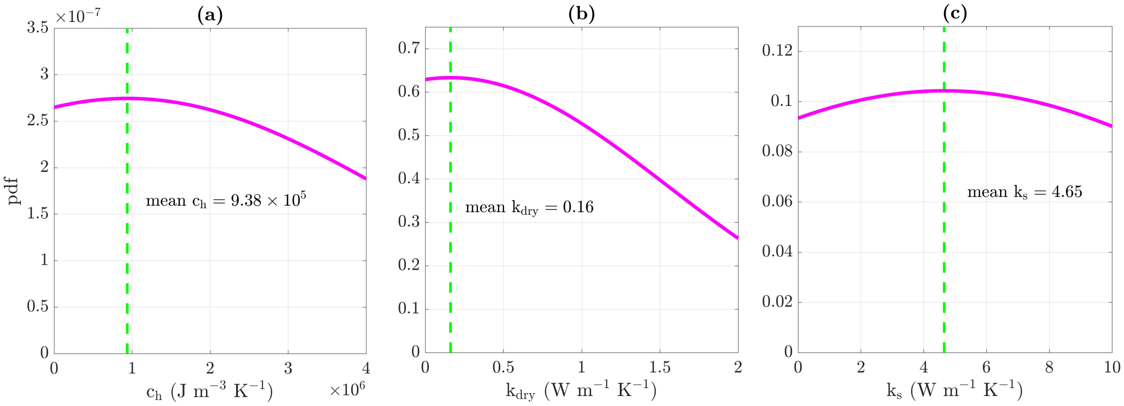

Figure 3.3: The same contents as in Fig. 3.2, for the Marre-Sale borehole site.

3.6.2 Model input: field measurements

Soil temperature, moisture, and ground heat flux were monitored at several plots located on the eastern slope of the Polar Urals of Russia since 2013. Data from one plot (LPTEG-TREE-1, 66°53’55” N, 66°45’27” E) are used to apply the developed process in Sec. 3.5.1. Fig. 3.4 shows the above-ground condition for this site. Specifically, ground heat flux was measured using soil heat flux plates (HuksefluxUSA Inc. HFP01 Heat Flux Sensor). The plate was placed horizontally in the soil and in tight contact with the soil medium. In most Arctic regions, a peat layer composed mainly of soil organic matter is present above mineral soil due to its slow rate of decomposition (*El Sharif et al.*, 2019). The porous peat layer of heterogeneous thickness and thermal properties makes it difficult to measure conductive heat flux due to poor contact of sensor-medium. Hence, the heat flux plate was installed at 6 cm below the mineral soil surface covered with a peat layer of 6 cm thickness. At the same location where the heat flux plate was buried, three temperature sensors (Campbell Scientific Inc. 109 and 107 thermistors, analog) were installed. Two thermistors were placed at 2 cm and 4 cm depth in the mineral soil to measure the temperature gradient and one thermistor was placed at 2 cm below the peat layer surface for peat layer temperature monitoring. The mineral soil temperature and moisture are measured

Figure 3.4: The surface condition of the LPTEG-TREE-1 site. The vegetation is composed of low shrubs and larch trees.

using a water content reflectometer (Campbell Scientific Inc. CS655, digital) with two prongs inserted at 1.5 cm and 4.5 cm below the mineral soil surface. Fig. 3.5 shows an example setup of some below-ground sensors in this field site. The rate of change of soil thermal energy storage above the heat flux plate ($\frac{dE_S}{dt}$, W m^{-2}) is computed by combining temperature and moisture observations for both peat and mineral soil: $\frac{dE_S}{dt} \approx \frac{\Delta E_S}{\Delta t} = \frac{\Delta T c_v \Delta z_S}{\Delta t}$, where ΔT is the hourly soil temperature change, Δz_S is the thickness of the mineral soil over the heat flux plate, and c_v is calculated by Eq. (3.10) with c_h estimated by Eq. (3.30). Ground heat flux at the surface for this plot (G_S^{FS}) is reconstructed as measured flux corrected by the rate of change of soil thermal energy storage (*Campbell Scientific*, 2016).

One issue with the field measurements is that liquid soil water content is available only when the soil temperature is above 0 °C. This is due to the physical design of the water content reflectometer based on the time domain principle (*Jones et al.*,

Figure 3.5: Setups for the CS655 water content reflectometer (white box on the left) and the HFP01 heat flux plate (horizontal plate on the right) placed in the mineral soil layer.

2002). Volumetric water content is derived based on the dielectric permittivity of the medium, so the sensor can only detect the signal of liquid water with high permittivity (*Campbell Scientific*, 2018). Due to this complication, ground heat flux is reconstructed using hourly temperature T_S at 2cm in the peat layer, and θ_{liq} at 2cm in the mineral layer, and G_S^{FS} during the growing season of 2019 when near-surface ice content $\theta_{ice} = 0$.

In the context of the method and notations from Sec. 3.5.1, measured G_S^{FS} is the target QoI, and the ensemble of $_S$ computed using $N = 500$ parameter sets represent the inference set. The number of eigenvalues and eigenfunctions identified in the KL expansion is $J = 2$, which is drastically smaller than the input size (1700 data points). The truncation order used for each polynomial surrogate model is $o = 4$. More details on the KL expansion estimates are presented in the Appendix B.

The reconstructed G_S^{FS} from heat flux plate measurements (black dashed line) and the reconstructed $\tilde{G}_S^{MAP}$ from temperature and moisture measurements (red solid line) at the field site are presented in Fig. 3.6(a). The variance of $\tilde{G}_S^{MAP}$ are presented in the format of 5-95% posterior predictive bounds (cyan area, see Sec. 3.4.5). The accuracy of the developed process is characterized by the RMSE = 6.447 and bias = 0.679 (W m^{-2}) as well as the scatter plot in Fig. 3.6(b). Fig. 3.6(c)-(f) shows the fitted marginal posterior distributions for c_h, k_{dry}, k_s and a (green curves) and the corresponding MAP estimates. The inferred MAP values fall within the typical ranges of soil substance properties: $c_h \sim 0.2$ to 1.4×10^6 (J m^{-3} K^{-1}), $k_{dry} \sim 0.05$ to 0.3 (W m^{-1} K^{-1}), and $k_s \sim 0.25$ to 9.0 (W m^{-1} K^{-1}) (*Lawrence and Slater*, 2008).

3.6.3 Model input: GCM data product

Global climate models are developed based on physical processes for simulating past, current, and future climate trajectories. The main components of the land surface energy budget including R_n, H, and λE, are provided explicitly in the GCM

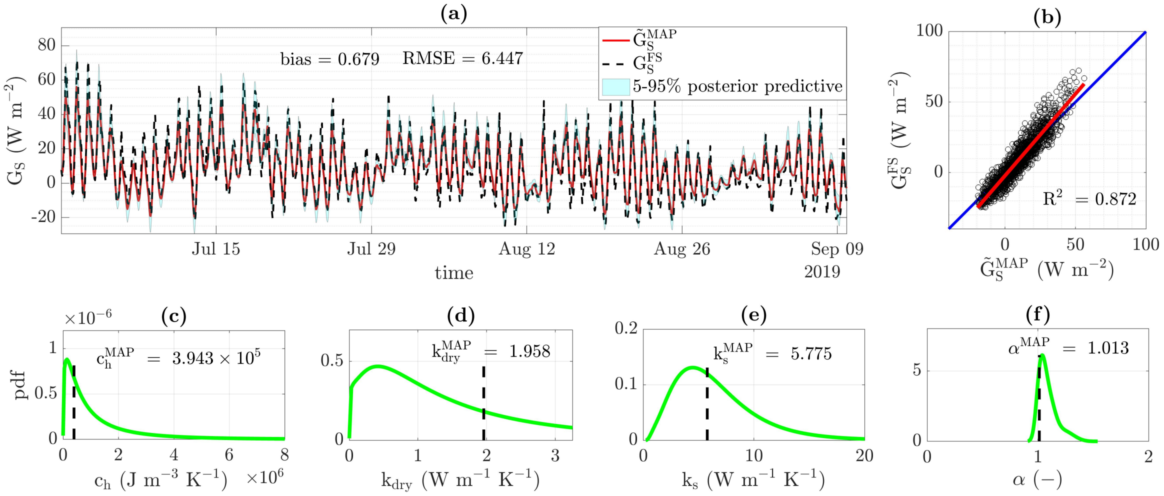

Figure 3.6: (a) Observed G_S^{FS} (black dashed line) and reconstructed $\tilde{G}_S^{MAP}$ (red solid line) heat flux series at the hourly resolution with 5-95% posterior predictive bounds (cyan area). The reconstruction is based on the process described in Sec. 3.5.1 using field-measured soil temperature and moisture (Sec. 3.6.4). (b) The scatter plot and linear regression (red solid line) fitted for $\tilde{G}_S^{MAP}$ and G_S^{FS} (the blue solid line is 1:1) with slope = 1.159 and intercept = -2.172. "R^2" is the coefficient of determination of the regression. (c)-(f) Kernel density estimation (KDE) of the marginal posterior distributions for c_h, k_{dry}, k_s, and α (green curves) derived as described in Sec. 3.6.4. Values of MAP parameter estimates (i.e., c_h^{MAP}, k_{dry}^{MAP}, k_s^{MAP}, and α^{MAP} are indicated by the vertical black dashed lines in (c)-(f). Parameter prior distributions are shown in Fig. 3.2.

outputs, but not G_S. Theoretically, the latter can be derived from the energy balance equation, i.e.,

$$G_S = R_n - H - \lambda E, \tag{3.39}$$

During snow-free summer period, heat flux derived using GCM outputs of surface energy budget components (as discussed in 3.1) can be considered as GCM-modeled heat flux at the ground surface (G_S^{GCM}). Conversely, the components of the snow surface energy budget are never explicitly provided in GCM outputs for the snowpack period, and therefore G_S cannot be derived from Eq. (3.39) during snowpack conditions. Nonetheless, when available, GCM simulated temperature of the topmost soil layer along with derived summer heat flux G_S^{GCM} can be jointly used to reconstruct G_S following the process described in Sec. 3.5.1 and illustrated in Fig. 3.1.

In this study, GCM models from the Coupled Model Intercomparison Project Phase 6 (CMIP6) (*Eyring et al.*, 2016) were selected based on the following criteria. First, the GCM model must provide near-surface soil temperature and water content in order to apply the parameter separation approach 3.2.2). Second, the model should provide outputs of components of the surface energy budget (i.e., R_n, H, and λE) to derive snowpack-free season G_S^{GCM} to infer soil thermal properties using the UQ machinery. Third, the model should also provide snow coverage conditions, such as snow depth, to separate the snowpack period from the rest of the year. After identifying the models, the required daily outputs (i.e., T_S, θ_{liq}, θ_{ice} in the top soil layer, R_n, H, and λE) for the historical period from 1981 to 2014 were downloaded from CMIP6 repository (see Sec. 4.2.2 for details).

In this study, the methodology is tested using daily historical outputs for the CMIP6 model: CESM2 (*Danabasoglu et al.*, 2020). The results are presented for the grid cell covering the location of the borehole site (Marre-Sale, see Sec. 3.6.4). Soil temperature and moisture of the 2-cm topsoil layer of the CESM2 are used to compute the inference set $\tilde{G}_S$. Summer period heat flux computed as the residual term of the surface energy budget (G_S^{GCM}) from Eq. (3.39) is used as G_S^{TGT} in the UQ machinery. For this case, the inference set size is $N = 500$, and the number of eigenvalues and eigenfunctions identified is $J = 2$ (i.e., much smaller than the input size of 3,889 data points, i.e., the number of snowpack-free days during 1981-2014). The truncation order of the surrogate polynomial is $o = 4$. Details of the KL expansion estimation are presented in Appendix B.

The concatenated series of snowpack-free seasons for reconstructed $\tilde{G}_S^{MAP}$ and G_S^{GCM} from CESM2 outputs were computed from 1981 to 2014 and Fig. 3.7(a) shows a time interval between 2010 and 2014 to illustrate a higher level of detail. Posterior predictive bounds display the variance of $\tilde{G}_S^{MAP}$. The scatter plot between $\tilde{G}_S^{MAP}$ and G_S^{GCM} in Fig. 3.7(b) shows a lower coefficient of determination (R^2) than that in Fig. 3.6(b). RMSE and bias are 9.762 and -6.083 (W m^{-2}), respectively. The marginal posterior distributions inferred for c_h, k_{dry}, k_s, and a and their corresponding MAP values are presented in Fig. 3.7(c)-(f). The MAP values of c_h, and k_{dry} are consistent with typical values for these soil properties, but that of k_s is one order of magnitude higher than what has been previously reported (see Sec. 3.6.2 for the ranges of these properties).

3.6.4 Model input: borehole measurements

Permafrost is defined as the ground frozen for at least 2 continuous years. In order to monitor the thermal state of the ground in cold regions, people usually drill narrow, deep pits into the ground, called boreholes, and place sensors in them properly to measure certain characteristics under the ground, e.g., temperature and moisture. Since the most recent International Polar Year (2007-2009), a global dataset of permafrost temperature has been developed based on borehole measurements. This data set is

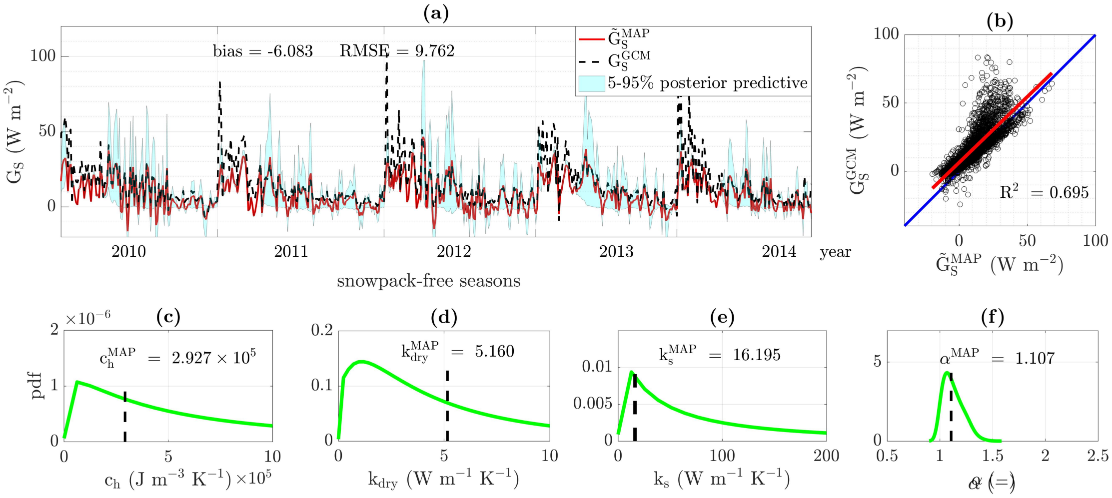

Figure 3.7: (a) Concatenated $\tilde{G}_S^{MAP}$ (red solid line) and G_S^{GCM} (black dashed line) series for snowpack-free seasons at the daily resolution with 5-95% posterior predictive bounds (cyan area). $\tilde{G}_S^{MAP}$ is reconstructed following the process in Sec. 3.5.1 using temperature and moisture of the topmost soil layer in CESM2 (Sec. 3.6.3). (b) The scatter plot and linear regression (red solid line) fitted for $\tilde{G}_S^{MAP}$ and G_S^{GCM} (the blue solid line is 1:1) with slope = 0.964 and intercept = 6.392. (c)-(f) KDE of the marginal posterior distributions for c_h, k_{dry}, k_s, and α (green curves) derived as described in Sec. 3.6.3. Vertical black dashed lines indicate values of MAP parameter estimates (i.e., c_h^{MAP}, k_{dry}^{MAP}, k_s^{MAP}, and α^{MAP}).

available from the Global Terrestrial Network for Permafrost (GTN-P) (*Biskaborn et al.*, 2015). Additional borehole soil temperature measurements in the GTN-P are also provided by the Circumpolar Active Layer Monitoring (CALM) program (*Brown et al.*, 2000). More details about these data sets are presented in Chapter IV. Borehole measurements indicate one of the fastest-warming rates of ground temperature in Northwestern Siberia (*Biskaborn et al.*, 2019). Measurements from the Marre-Sale borehole site (69°43' N,66°51' E) (*Melnikov et al.*, 2004; *Vasiliev et al.*, 2008; *Malkova et al.*, 2022) representing this region were selected to reconstruct G_S using the developed process in Sec. 3.5.1 (see Sec. 4.2.1 for data downloading details). In general, an 'ideal' set of characteristics for borehole locations includes (1) temperature measurements close to the ground surface (few cm), (2) continuous records and as long as possible, and (3) measurements for several depths. Daily soil temperatures are available for depths 5 cm, 50 cm, 1 m, 2m, 3m, 5m, and 10 m for this borehole. Time series segments between December 1, 2006, and December 31, 2010, with the most complete and continuous temperature measurements, are used henceforth. Missing data values in the observed $T_{S,1}$ and $T_{S,2}$ series are filled by linear interpolation, resulting in a total of 1492 data points. Since soil moisture was not measured in the borehole, assumptions in Sec. 3.5.2 are used.

Specifically, soil temperature $T_{S,1}$ measured at the shallowest depth $z_1 = 5$ cm is used to derive the inference set T_S^{HOD} at depth $z_2 = 50$ cm. Soil temperature measured at z_2 ($T_{S,2}$) is used as T_S^{TGT} in the UQ machinery with $N = 500$, $q = 4$, $J = 3$. The series G_1^{MAP} and $T_{S,2}^{MAP}$ are computed from $T_{S,1}$ using Eqs. (3.1) and (3.2) with the inferred c_h^{MAP}, $k_s MAP$, and θ_{sat}^{MAP}. A set of scaling factor a from 1 to 2 with an interval of 0.01 is used to compute $\tilde{G}_S$ and $a^{RMSE} = 1.17$ resulting in the minimum RMSE = 2.410 °C between $T_{S,1}$ and $T_{S,1}^{MAP}$, where $T_{S,1}^{MAP}$ is derived from $\tilde{G}_S$ using Eq. (3.3). The reconstructed heat flux $\tilde{G}_S^{MAP} = a^{RMSE} G_1^{MAP}$ is then verified by the numerical model. The numerical model simulation setups are described below.

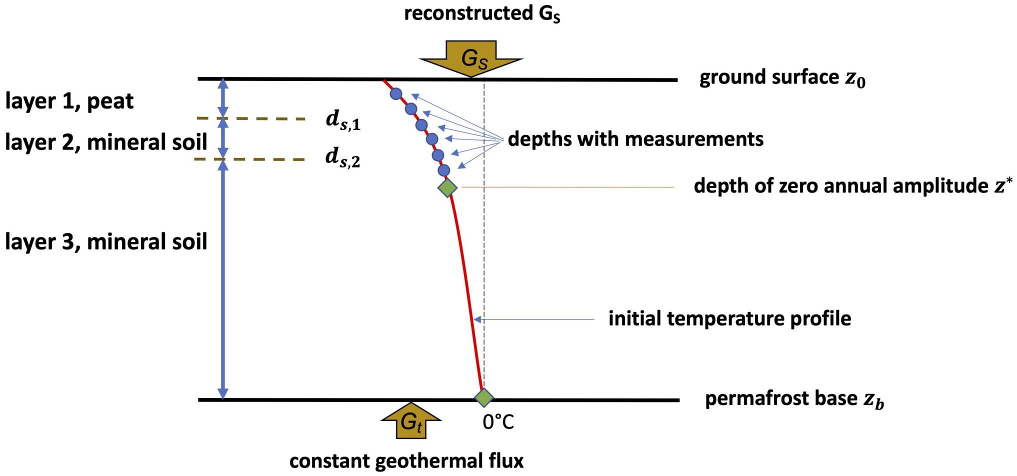

Figure 3.8: A diagram of soil column setup as used in the numerical model simulations.

Fig. 3.8 displays a conceptual diagram of the soil column setup based on three soil layers represented with the numerical model. Layer 1 represents the peat layer, while layers 2 and 3 represent mineral soils. The boundary condition at the ground surface is the reconstructed $\tilde{G}_S^{MAP}$ and at the permafrost base is a constant geothermal flux $G_t = 0.07$ (W m^{-2}) derived from a global earth surface heat flow map (*Davies*, 2013). The initial temperature profile (red curve) is constructed through a linear interpolation using borehole temperature measurements (blue dots), the temperature at the depth of zero annual amplitude z^* (green diamond), and temperature at the permafrost base z_b (0°C). In this study, z^* is set to 12 m, and the temperature at z^* is determined as the annual average of temperature from the deepest available measurements (10 m).

For a soil layer i, the uncertain parameters that need to be inferred are the thermal conductivity ($k_{s,i}$) and the volumetric heat capacity ($c_{s,i}$) of soil solid, the porosity (ϕ_i), the pore size distribution index (b_i), and the air entry pressure head ($\psi_{s,i}$) in the B-C model (see Sec. 3.3. In this study, three soil layers are considered in the numerical model (as presented in Fig. 3.8), thus there are two additional parameters:

the bottom depths of soil layer 1 ($d_{S,1}$) and 2 ($d_{S,2}$) need to be determined. The range for the uncertain parameter $d_{S,1}$ is [10, 30] cm and $d_{S,2}$ [300, 500] cm. The permafrost base z_b is set to 300 m (personal communication with research personnel at the Marre-Sale site). To sum up, a total number of $m = 17$ uncertain parameters are inferred following the process described in Sec. 3.4.

For each parameter, their respective minimum and maximum values are selected first based on a priori knowledge to construct a uniform prior distribution. The total of $N = 1500$ parameter sets are then generated using the developed uniform prior distributions. A set of N simulations are performed using the developed parameter sets. In this case, the QoIs are the soil temperatures at the depths where borehole measurements exist, and these are available for seven depths for the Marre-Sale site from Dec. 2006 to Dec. 2010. Infrequent missing data in the daily measurements are filled by linear interpolation.

In order to apply the KL expansion (Sec. 3.4.3, measurements at the different depths are concatenated head-to-tail, from shallow to deep locations, thereby forming a single series (black solid line in Fig. 3.9). For each simulation, temperatures at the seven depths are also concatenated (i.e., every single red dashed curve in Fig. 3.9). The posterior probability distributions and the MAP estimates for the uncertain model parameters are obtained using the ensemble of concatenated temperature simulations (i.e., the inference set) and the concatenated temperature measurements (i.e., the target QoI).

Figure 3.10 compares the observed $T_{S,2}$ at z_2 and $T_{S,2}^{MAP}$ simulated with Eq. (3.2) using $T_{S,1}$ series and MAP values of the parameters (indicated in the posterior marginal distribution in Fig. 3.10(c)-(e)) with posterior predictive bounds. The magnitudes of the c_h^{MAP}, and θ_{sat}^{MAP} estimates are within the order of values reported for soil medium properties previously. The estimation of the parameters along with the G_1^{MAP} series leads to a reconstruction $\tilde{G}_S^{MAP}$ (Fig. 3.11(a)) and its magnitude

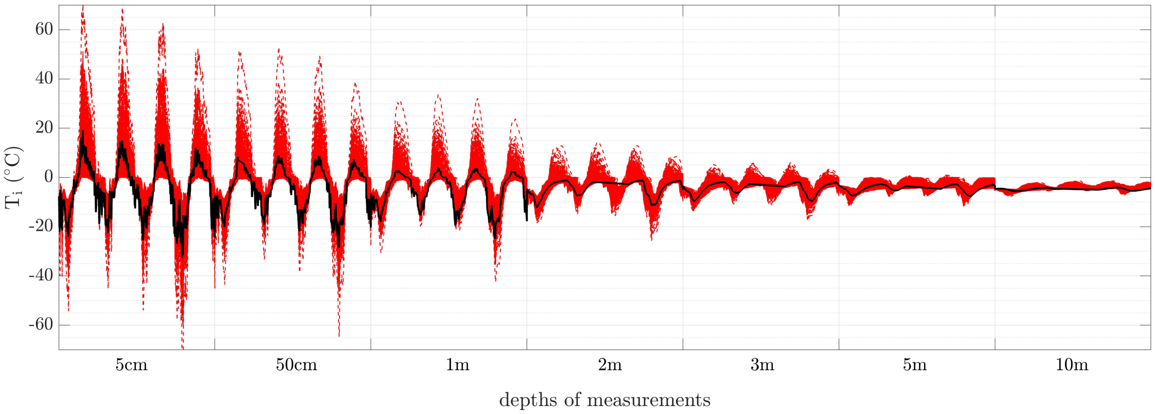

Figure 3.9: The head-to-tail, shallow-to-deep concatenated temperature measurements at the available depths (black solid curve) and the inference set simulations with the numerical model at these depths (red dashed curves ensemble). The individual segments for each depth are the temperature series from Dec. 2006 to Dec. 2010.

and dynamics are reasonable, as compared to field measured and GCM derived time series of G_S (e.g., in Fig. 3.6(a) and Fig. 3.7(a)).

Finally, Fig. 3.12 compares the simulated temperatures from the numerical model that uses $\tilde{G}_S^{MAP}$ as input and the borehole measurements at different depths down to 10 m with posterior predictive bounds representing the variance of simulation results. These results confirm the credibility of the reconstructed $\tilde{G}_S^{MAP}$ series. Fig. 3.13 displays the marginal posterior distributions for the 17 parameters used in the numerical model and their corresponding MAP estimated values. $k_{s,i}$ and $c_{s,i}$, $i = 1, 2, 3$ are in order of typical values (see Sec. 3.6.2 and Sec. 3.6.4). The porosities ϕ_i are also within the physical meaningful range (0,1). The estimates for parameters b_i, $\psi_{s,i}$ used in the B-C model are in order with values reported in the literature as well (e.g., *Morel-Seytoux et al.*, 1996; *Wang et al.*, 2002).

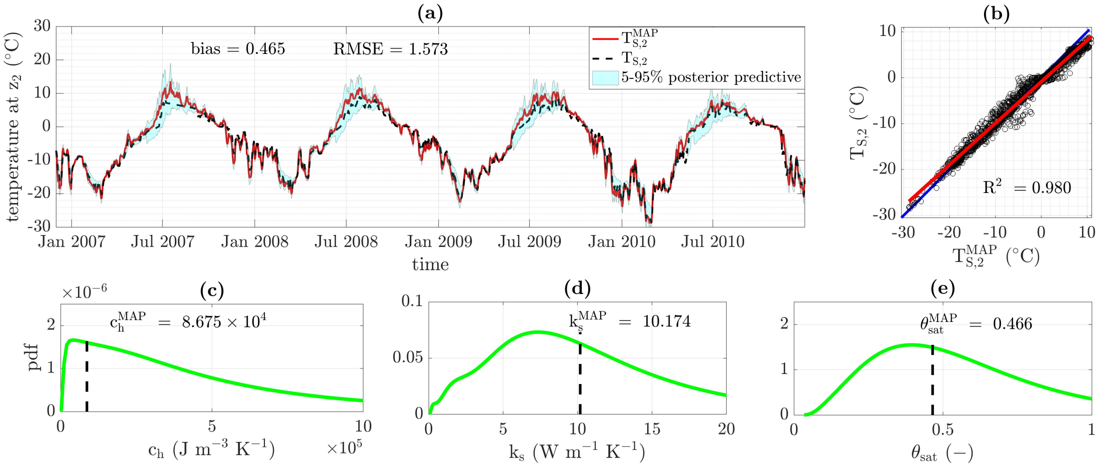

Figure 3.10: (a) The series of measured soil temperature $T_{S,2}$ (at $z_2 = 50$ cm, black dashed line) and $T_{S,2}^{MAP}$ (at $z_2 = 50$ cm) computed from $T_{S,1}$ measured at $z_1 = 5$ cm, as described in Sec. 3.6.4 with 5-95% posterior predictive bounds (cyan area). The observational data are for the Marre-Sale borehole site from 2007 to 2010 at the daily resolution. (b) The scatter plot and linear regression (red solid line) fitted for $T_{S,2}$ and $T_{S,2}^{MAP}$ (the blue solid line is 1:1) with slope = 0.903 and intercept = -0.891. (c)-(e) KDE of marginal posteriors for c_h, k_s, and θ_{sat} (green curves) derived in Sec. 3.6.4 and MAP estimations, c_h^{MAP}, k_s^{MAP}, and θ_{sat}^{MAP} (vertical black dashed lines).

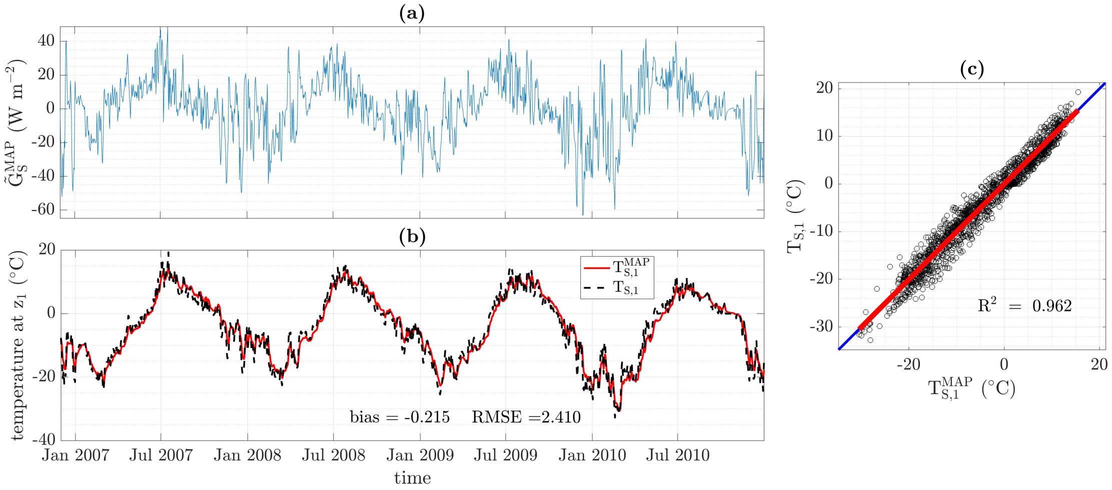

Figure 3.11: (a) The reconstructed $\tilde{G}_S^{MAP}$ series. (b) A comparison of soil temperature T_1 measured at $z_1 = 5$ cm and T_1^{MAP} derived from the $\tilde{G}_S^{MAP}$ series in (a) for the Marre-Sale borehole site from Dec. 2006 to Dec. 2010 at a daily scale. The time series of $\tilde{G}_S^{MAP}$ is computed as the product $\alpha^{RMSE} G_1^{MAP}$, where the time series G_1^{MAP} is calculated with Eq. (3.1) using the MAP parameter set (see Fig. 3.10) and the observed T_1 series. The minimum RMSE = 2.410 (°C) between T_1 and T_1^{MAP} is obtained by using the value of a scaling factor $\alpha^{RMSE} = 1.17$. (c) A scatter plot and linear regression (red solid line) between T_1 and T_1^{MAP} with slope = 1.000 and intercept = -0.023 (the blue solid line is 1:1).

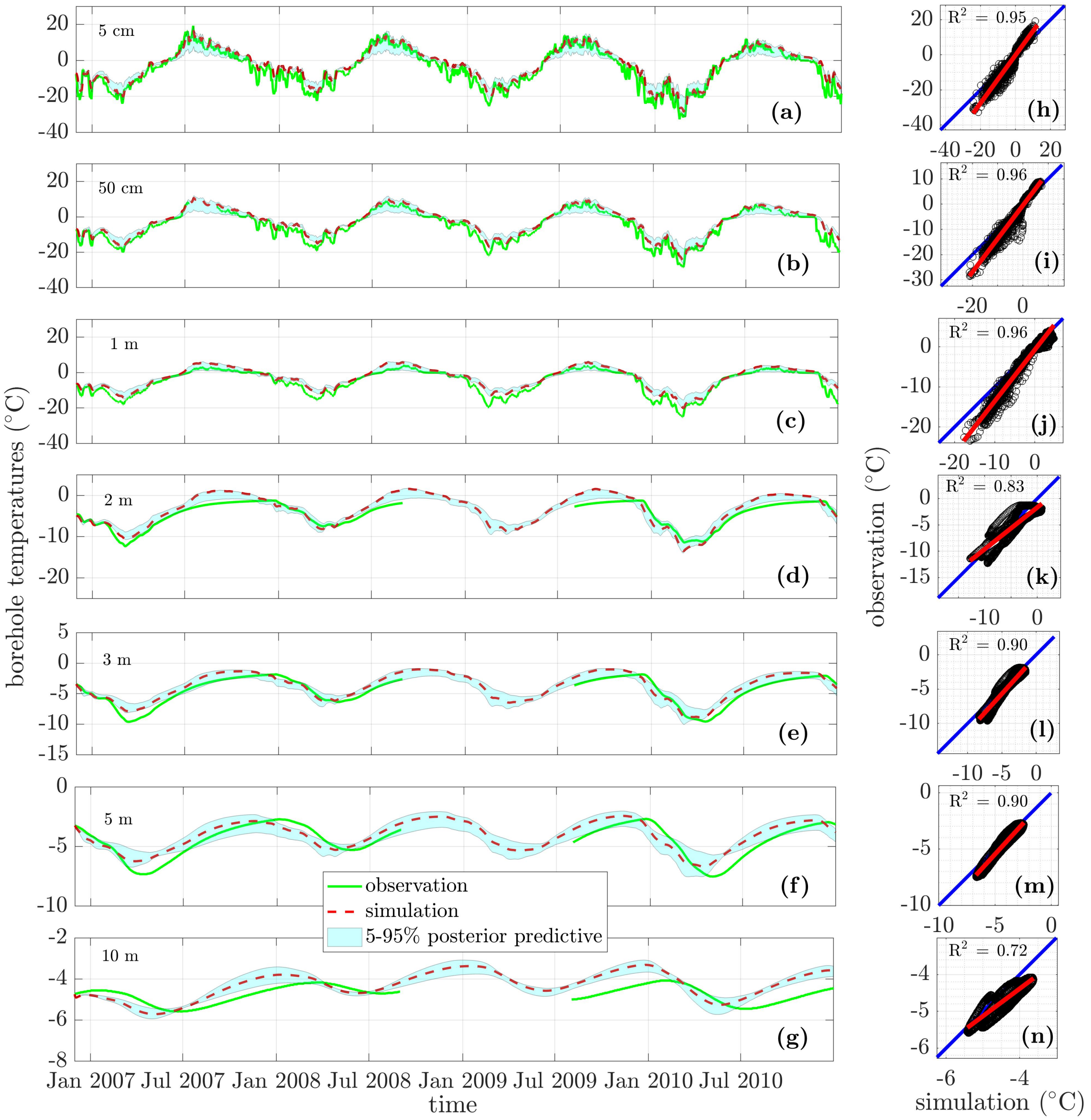

Figure 3.12: (a)-(g) A comparison of the temperature time series based on simulations using the numerical model described in Sec. 3.3 using $\tilde{G}_S^{MAP}$ (Fig. 3.11(a)) and the daily temperature measurements at different depths at Marre-Sale from 2007 to 2010 with 5-95% posterior predictive bounds (cyan area). (h)-(n) The corresponding scatter plots and linear regression (red solid line) between simulated and observed temperature.

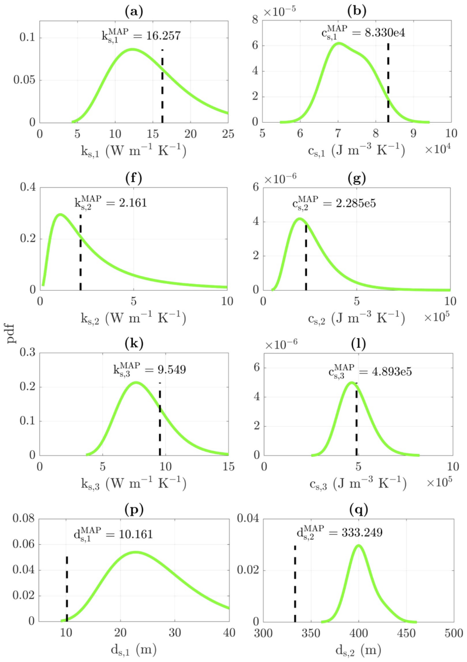

Figure 3.13: KDE of the marginal posterior distributions (green curves) and their MAP estimates (vertical black dashed lines) for numerical model parameters described in Sec. 3.6.4.

3.7 Discussion

One of the advantages of utilizing the UQ machinery in the developed process in this study is the significant improvement of computational efficiency. The contribution comes from the drastically decreased order of data size by KL expansion and reduced model complexity by PC surrogate modeling. For all the cases presented above, the data points needed in the parameter inference process are reduced from $\mathcal{O}\left(10^3\right)$ to around 2 to 3. Part of this extreme drop in data size is due to their high correlation in space and time. For example, soil temperatures for consecutive days are correlated through legacy effects, and for different depths are correlated through the heat transfer process. The observation data are well estimated by the KL-PC estimation and captured by its variance in terms of posterior predictive bounds for all types of data, which illustrate the accuracy of the developed process considering the information sacrifice while keeping the simulation fast.

The developed methodology permits probabilistic assessment of surface ground heat flux and parameters of models used in heat flux reconstruction. The MAP estimates of individual parameters are presented in Figs. 3.6(c)-(f), 3.7(c)-(f), 3.10(c)-(e), and 3.13 do not necessarily coincide with the modes of their marginal posterior distributions. This mismatch implies that the marginal posterior distributions of individual parameters are interdependent with respect to each other. In other words, if all marginal posterior distributions were entirely independent (and since the joint distribution, in this case, is the product of marginal distributions), then MAP estimates of individual parameters derived from the joint posterior distribution would correspond to the maxima of the corresponding marginal distributions. Another possible reason that causes this mismatch is the relative importance of different parameters. If the model is more sensitive to some parameters, their posteriors will be sharper, and other less important parameter posteriors will be flatter with MAP deviating from the mode.

When applying the developed process for GCM data products as shown in section 3.6.3, one inferred parameter is out of the typical range for soil medium properties. This may relate to a specific parameterization of heat flux processes implemented in the GCM. For most GCMs (including CESM2 used in this study), the integrated model is formed by coupling separate model components (e.g., atmosphere, land, and ocean models). Since the GCM data products used to reconstruct $\tilde{G}_S^{MAP}$ (e.g., soil temperatures and moisture) and G_S^{GCM} (e.g., R_n) are from different components, the parameter inference results may be affected by the coupling process assumed in a specific GCM. Nevertheless, in order to use either $\tilde{G}_S^{MAP}$ or G_S^{GCM} derived from global data products in the regional or local scale studies, a downscaling process is required to reduce model bias. Thus, the impact of the inferred parameter outside of the plausible range may not be crucial after correcting $\tilde{G}_S^{MAP}$ by using ‘ground truth’ in a downscaling process. The observed and reconstructed heat flux G_S at the field

(Sec. 3.6.2) or borehole (Sec. 3.6.4) sites are valuable as 'ground truth' and useful in the downscaling process.

One limitation of the HOD model is that the analytical solution of G_S does not take the latent heat of the phase change of water into account. However, the latent heat of the phase change within the considered shallow layer of 5 cm is negligible, as compared to the total accumulated seasonal heat into or out of the ground. For example, the latent heat of phase change can be approximately calculated as $(z_1 - z_0)\phi\rho_w L$, where ρ_w is the density of liquid water (kg m^{-3}) and L is the latent heat of freezing for water (J kg^{-1}). For the case of the borehole in Sec. 3.6.4, one can approximately estimate that the latent heat of phase change is less than 0.03% of the cumulative diffusive flux for all seasons (the calculation is not presented). Therefore, the reconstructed flux G_S is not affected by neglecting the phase change energy in the top shallow soil layer.

When soil moisture data are missing, assumptions about the soil water content is made, as described in Sec. 3.5.3. One can avoid using these assumptions by applying the HOD model in its original formulation in which only two temporally invariant variables (I, and D) are required, i.e., the formulation does not depend on the change in soil water content. However, the thermal properties of frozen and liquid water are different and need to be considered when reconstructing G_S at the daily resolution as well as in numerical simulations of even finer time steps. The assumptions are overall suitable for locations in the Arctic region with the active layer near saturation during the winter. During relatively short summer seasons, the reconstructed G_S may be overestimated under the assumption of full soil saturation assumption due to the higher thermal conductivity and diffusivity of liquid water as compared to air. Measurements of soil water content at a suitable frequency (daily or finer) concomitant with borehole temperature measurements may help overcome this issue.

3.8 Conclusion

Ground heat flux is a vital energy budget component in cold regions and its quantification is important for understanding surface mass-energy relations and subsurface thermal conditions. The process developed in this study provides a useful approach for reconstructing surface ground heat flux and its probability distribution from different types of shallow soil temperature data, with or without auxiliary data (i.e., soil moisture and snowpack-free season G_S data). This process of G_S reconstruction offers a suitable alternative to the laborious measurements of in-situ G_S in remote regions such as the Arctic. Another advantage of the proposed approach is that the flux G_S for all seasons can be derived as long as year-round (daily, or finer temporal resolution) soil temperature series are available, which are more straightforward to obtain in terms of relevant sensor deployment and maintenance in the field. In the case of future climate projections, the likelihood and severity of future ground heating in Arctic areas underlain by the permafrost can only be inferred from GCM projections of soil temperatures; however the validity of these projections is questionable since GCMs typically have a simplified representation of the belowground freeze-thaw dynamics, yet available outputs do not offer flux G_S for independent validation. Furthermore, in the experience of this author group, data requests to individual GCM centers to obtain G_S series are futile and therefore the developed methodology offers means for independent verification of the past heating and plausibility of the projected heating of the permafrost in the future.

The credibility of the reconstructed flux G_S is supported by a series of results presented in the application section. The UQ machinery applied in this study is useful for objective parameter inference and yields estimates in a plausible range of real-world soil properties. In the meantime, for data that are highly correlated in space and time, the developed process can significantly increase the simulation speed. This is especially beneficial for the case when complex models based on massive physical processes are applied. As discussed, the reconstructed flux G_S may be overestimated during summer due to the assumption of full saturation (when soil moisture data is not available). This assumption is acceptable for the Arctic regions considered in this study due to the relatively short summer period and low evapotranspiration potential. Suitable moisture measurements can facilitate a more accurate G_S reconstruction.

Importantly, the reconstructed flux G_S and its probability density function can be used as a top flux boundary condition in detailed numerical physical models to simulate the probability distribution of the subsurface thermal state in specific regional and built environment settings. This is important for understanding the future dynamics of the belowground environment. For example, the developed process can be applied to reconstruct plausible future heating scenarios of the permafrost, with respect to the Shared Socioeconomic Pathways as developed by the Intergovernmental Panel on Climate Change (IPCC). Aided by detailed physical modeling and data for a region of application, such scenarios can help better resolve the probabilistic structure of permafrost future impacts on land-surface hydrology, built environment, and livelihoods of the peoples of the North.

CHAPTER IV

Permafrost Fate under Projected Climate Change

4.1 Introduction

Permafrost degradation is defined as the decrease of permafrost in both thickness and area extent. It is one of the most significant phenomena across the circumpolar region under the influence of climate change (*Lunardini*, 1996; *Streletskiy et al.*, 2015). This change in permafrost condition will consequently impact the regional ecologic (*Jorgenson et al.*, 2001; *Osterkamp et al.*, 2009; *Jin et al.*, 2021) and hydrologic systems (*Jorgenson et al.*, 2013; *Walvoord and Kurylyk*, 2016; *Liljedahl et al.*, 2016), infrastructure stabilities (*Streletskiy et al.*, 2012; *Shiklomanov et al.*, 2017; *Hjort et al.*, 2018, 2022) in the Arctic, as well as the global climate (*MacDougall et al.*, 2012; *Schuur et al.*, 2013; *Schaefer et al.*, 2014).

Efforts have been made in previous studies to understand the change of permafrost thermal regime through field monitoring (*Melnikov et al.*, 2004; *Romanovsky et al.*, 2010a; *Vasiliev et al.*, 2008, 2020; *Malkova et al.*, 2022), and modeling *Smith and Riseborough* (1996); *Nicolsky et al.* (2007); *Riseborough et al.* (2008); *Nicolsky et al.* (2017); *Nicolsky and Romanovsky* (2018); *Teufel and Sushama* (2019).

Climate projections made from the Global Climate Models (GCMs) are often used as the forcing condition in studies concerning permafrost conditions in the future (*Lawrence and Slater*, 2005; *Nicolsky et al.*, 2007; *Zhang et al.*, 2008; *McGuire* *et al.*, 2018). GCMs contribute to the Coupled Model Intercomparison Project Phase 6 (CMIP6) (*Eyring et al.*, 2016) that is used more often in recent climate projection studies in the Arctic (*Burke et al.*, 2020; *Alexandrov et al.*, 2021) as it includes the latest comprehensive Earth system models with projected scenarios representing different Shared Socioeconomic Pathways (SSP) combined with updated Representative Concentration Pathways (RCP) (*Tokarska et al.*, 2020). Before utilizing the GCM outputs directly for further studies, their intrinsic biases need to be reduced with some optimized methods (*Raäisaänen*, 2007; *Knutti et al.*, 2017; *Xu et al.*, 2019) and the uncertainty associated with their coarse spatial resolution needs to be addressed as well (*Fatichi et al.*, 2011; *Kim et al.*, 2016). The Bayesian Weighted Averaging method (BWA) proposed by *Tebaldi et al.* (2005) is used as the stochastic downscaling strategy in this study. The weight for a specific GCM depends on its performance in terms of its difference to historical observation (bias criterion) and the overall alignment with other models in the future (convergence criterion) (*Giorgi and Mearns*, 2002; *Tebaldi et al.*, 2005).

The research questions that triggered the study in this chapter are: (1) how would the permafrost thermal regime (e.g., soil temperature profile) change under the projected climate in the future? And (2) how can we utilize historical observations about the permafrost condition to infer permafrost change? These questions are addressed by applying the uncertainty-informed framework developed in Chapter III and the BWA stochastic downscaling methods. Specifically, the future permafrost thermal regime is studied focusing on the Northwestern Siberia region in this chapter as observations indicate one of the fastest soil temperature warming rates here (*Biskaborn et al.*, 2019). Borehole observations and GCM outputs are combined in the downscaling process to provide projected surface ground heat fluxes for specific emission scenarios. The obtained ground heat fluxes are used as the boundary inputs and the permafrost temperature profile is simulated using the numerical model with parameters inferred through uncertainty quantification machinery.

4.2 Data

4.2.1 Soil temperature from boreholes in Northwestern Siberia

In order to study the impact of climate change on permafrost degradation in cold regions, it is crucial to understand the historical and ongoing permafrost thermal state. To fulfill this need, long-term monitoring of ground temperature distribution and seasonal ground thawing and freezing depth are recorded from borehole measurements. Despite the difficulty and challenges of conducting such in-situ measurements, many attempts have been made previously to monitor the permafrost state across the Arctic (*Harris et al.*, 2001, 2009; *Smith et al.*, 2005, 2010; *Romanovsky et al.*, 2010a,b; *Konstantinov et al.*, 2020; *Vasiliev et al.*, 2020). (*Brown et al.*, 2000). Based on the records, the permafrost thermal regime is usually characterized by the near-surface mean annual ground temperature (MAGT), active layer thickness (ALT), and the depth of zero annual amplitudes (DZAA) (*Smith and Riseborough*, 1996; *Malkova et al.*, 2022). The observed borehole temperature profile is also essential for validating mechanistic models that are applied to simulate and project the permafrost thermal state. Besides, important boundary conditions used in these models such as the surface ground heat flux (G_S) can also be derived from the available state data provided by these data sets (discussed in Chapter III).

In this study, data sets from two compilations of borehole thermal state based on international collaborative work are used to reconstruct the historical G_S flux: the Global Terrestrial Network for Permafrost (GTN-P) (*Biskaborn et al.*, 2015) and the Circumpolar Active Layer Monitoring (CALM) (*Brown et al.*, 2000) program. The main characteristics commonly monitored at the borehole sites are the soil temperatures at various depths and the active layer thickness (ALT). Following the three screening criteria listed in Sec. 3.6.4, vertical temperature profiles from three 'ideal' borehole sites (listed in Table 4.1) in northwestern Siberia: Marre Sale (MS), Vaskiny Dachy (VD), and Bolvansky (BV) are selected to conduct the G_S reconstruction process. Fig. 4.1 displays their locations. The MS site is located on the west side of the Yamal Peninsula and near the coast; the VD site is situated in the middle of the Yamal Peninsula; the BV site is located in the Pechora River delta. Brief descriptions of the surface cover condition, qualitative soil properties, as well as data-collecting procedures for each of the borehole sites are available on the CALM website. Data are downloaded directly from the publicly open data websites for GTN-P (http://gtnpdatabase.org/boreholes) and CALM (https://www2.gwu.edu/ calm/data/north.htm).

Soil temperature at different depths combining the available measurements from the two data sets for the three boreholes is presented in figure 4.2. Specifically, temperature measurements above 3 m are from CALM, and below from GTN-P. Although for these three boreholes, temperature records over at least 2 years are most complete and continuous compared to other boreholes listed in the GTN-P and CALM data sets, there are still many gaps that exist in the temperature series. Among the three boreholes, the MS site has the longest records for shallow depths. For the VD site, there are even no data available below 1.5 m depth (blank plots at the bottom in the second column of Fig. 4.2.

Table 4.1: Boreholes information

Borehole site	Coordinates	Measurement depths	Reference
Marre Sale	69°43'N 66°51'E	5 cm, 50 cm, 1 m, 2 m, 3 m, 5 m, 10 m	*Melnikov et al.* (2004), *Vasiliev et al.* (2020)
Vaskiny Dachy	70°17'N 68°54'E	1 cm, 50 cm, 1 m, 1.5 m	*Leibman et al.* (2015)
Bolvansky	68°18'N 54°30'E	10 cm, 50 cm, 1 m, 1.5m, 3 m, 5 m, 10 m	*Mazhitova et al.* (2004), *Vasiliev et al.* (2020)

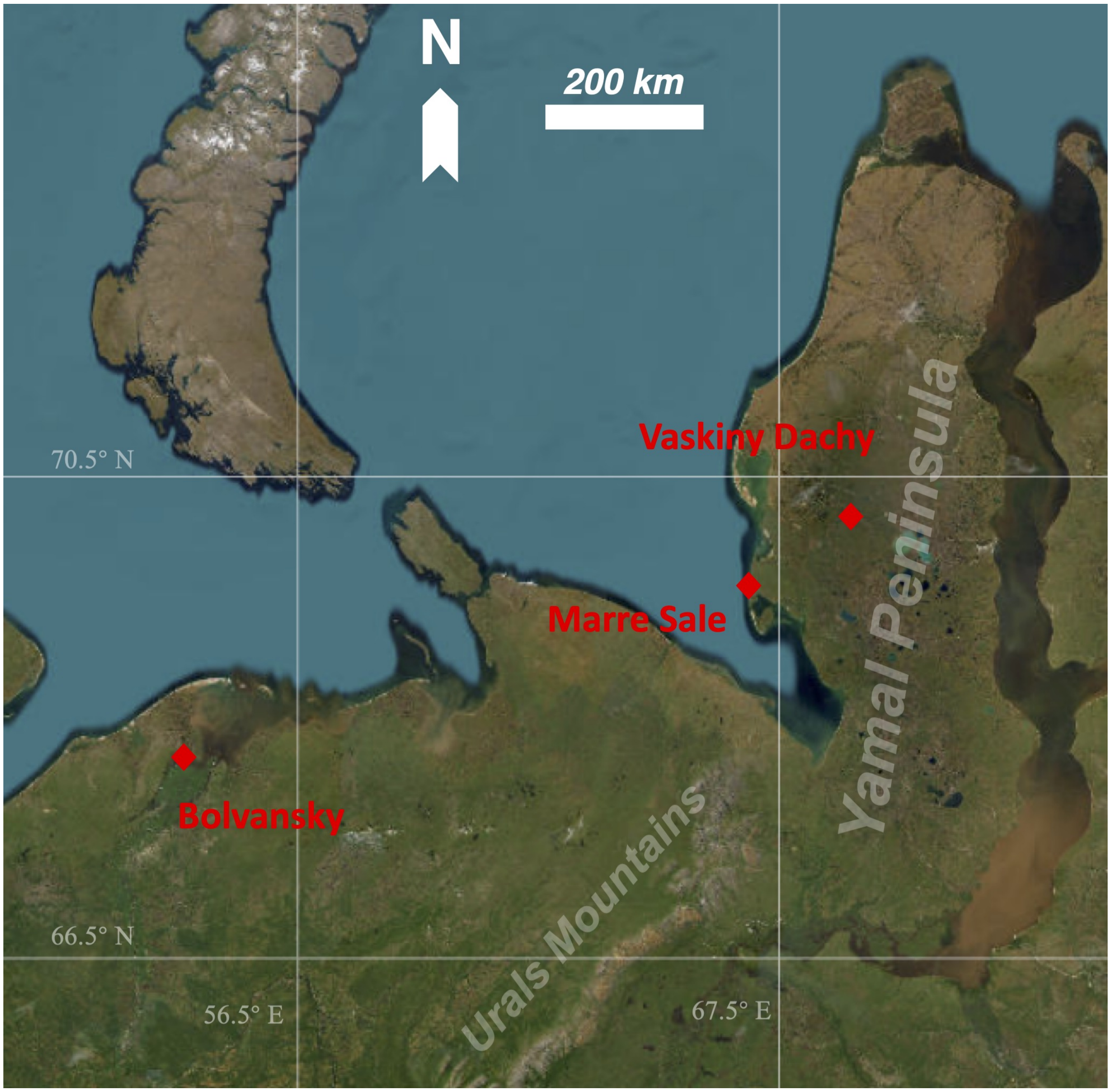

Figure 4.1: Geographic map including 3 borehole locations listed in Table 4.1.

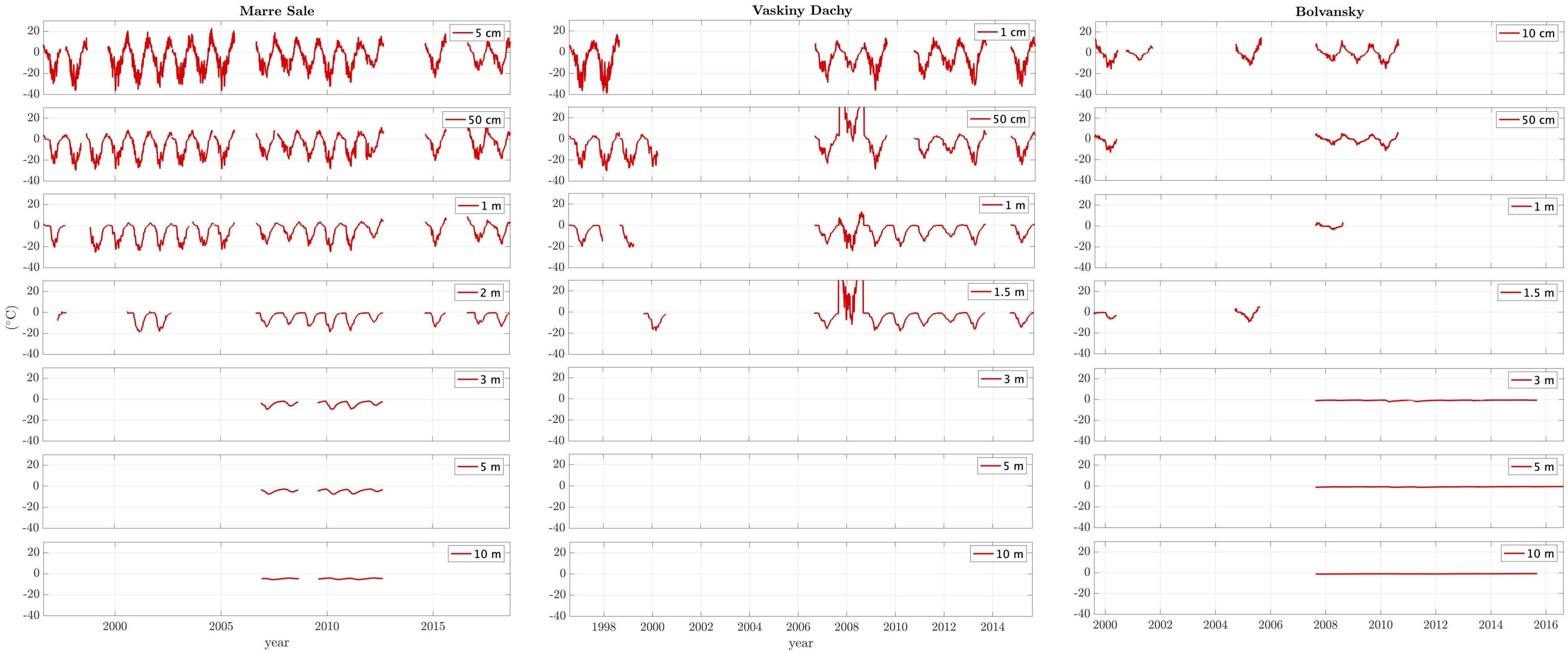

Figure 4.2: Raw temperature measurements at various depths for the boreholes listed in Table 4.1

4.2.2 Historical and future GCM outputs

In order to project the permafrost thermal state in the future using the numerical model described in Sec. 3.3, the boundary condition, i.e., G_S flux in the future, needs to be derived first. The details of the G_S reconstruction process are described below in the methods section (see Sec. 4.3.2). The required GCM outputs are net radiation (Rn), sensible (H) and latent (λE) heat fluxes for the historical period, soil temperature (T_S), and moisture (θ, if available) at the top soil layer for both the historical and future period. In this study, the required data are retrieved at daily resolution from 9 selected GCMs (listed in Table 4.2) that satisfied the screening criteria listed in Sec. 3.6.3 from the CMIP6 model archive (*Eyring et al.*, 2016). The historical simulation outputs are downloaded from 1981 to 2014. The CMIP6 projections are made on different scenarios which combine shared socioeconomic pathways (SSPs) with updated representative concentration pathways (RCPs). In this study, projected outputs from 2015 to 2100 under two of these scenarios are used: SSP2-4.5 (represents the medium of plausible future forcing pathways, or 'mild warming') and SSP5-8.5 (represents the high end of plausible future forcing pathways, or 'intensive warming'). Data are downloaded from the CMIP6 repository website (https://esgf-node.llnl.gov/search/cmip6/). Fig. 4.3 displays the 9-model averaged historical and projected (under two warming scenarios) air temperature at 2 m above the ground surface and precipitation from GCM outputs at the grid cells that cover the borehole sites MS, VD, and BV. Both the air temperature and the precipitation display a generally increasing trend since 1980 for all borehole sites. For the projected era, this trend is more significant under the SSP5-8.5 scenario (red lines) compared to the SSP2-4.5 scenario (blue lines).

GCM outputs are usually at a relatively coarse spatial resolution ($\sim$ 100 km). In order to conduct the downscaling process, GCM outputs are retrieved only at the grid cells that cover the locations of the three boreholes. GCM outputs for a grid cell

Table 4.2: CMIP6 models used in this study

No.	Model	Institute	Land model	No. layers	Soil depth (m)	References
1	CESM2	NCAR	CLM5	25	42.0	*Danabasoglu et al.* (2020)
2	HadGEM3-GC31-LL	MOHC	JULES	4	2.0	*Williams et al.* (2018)
3	IPSL-CM6A-LR	IPSL	ORCHIDEE	18	65.6	*Boucher et al.* (2020)
4	MPI-ESM1-2-LR	MPI-M	JSBACH3.20	5	7.0	*Mauritsen et al.* (2019)
5	UKESM1-0-LL	MOHC	JULES	4	2.0	*Sellar et al.* (2019)
6	ACCESS-ESM1-5	CSIRO	CABLE2.4	6	2.9	*Ziehn et al.* (2020)
7	EC-Earth3	EC-Earth-Consortium	HTESSEL	4	1.9	*Döscher et al.* (2022)
8	MIROC-ES2L	JAMSTEC	MATSIRO6.0	6	9.0	*Hajima et al.* (2020)
9	NorESM2-LM	NCC	CLM5	25	42.0	*Seland et al.* (2020)

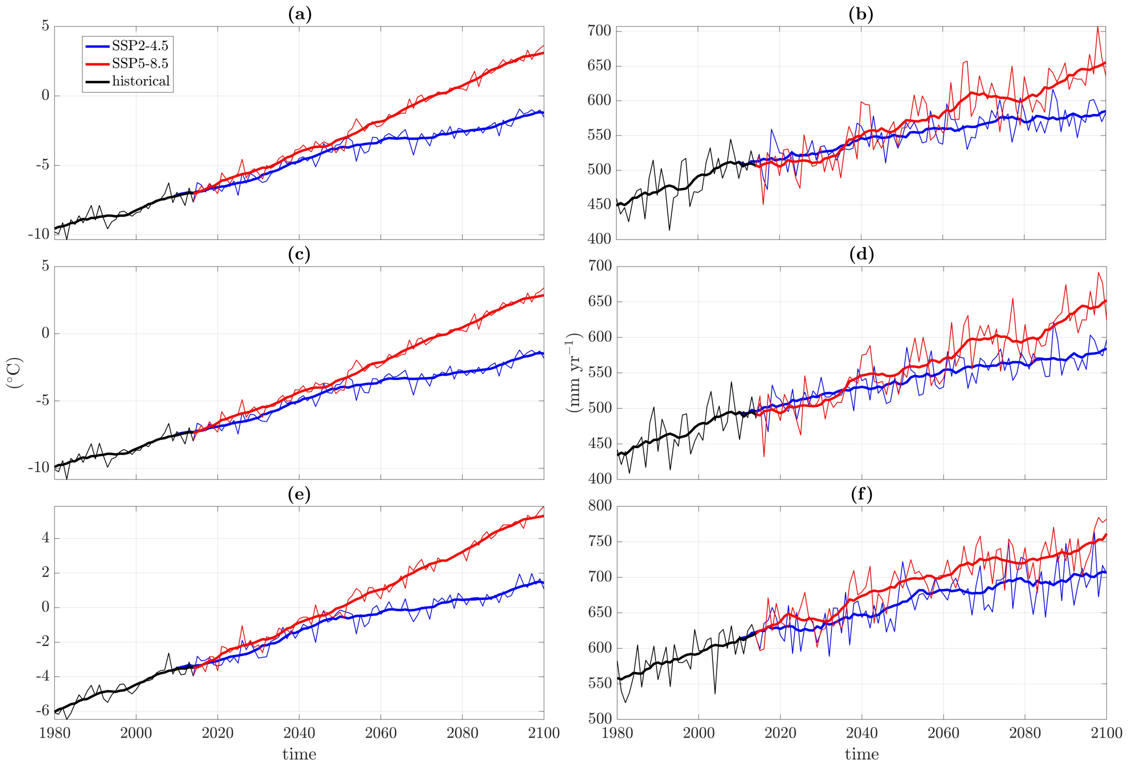

Figure 4.3: (a), (c), (e) Historical (black lines) and projected (blue and red lines) air temperature at 2 m height and (b), (d), (f) precipitation for MS, VD, and BV site, respectively. Results are averaged over the outputs retrieved from the 9 models listed in Table 4.2. The thinner and more drastically oscillated lines are annual mean values, while the thicker and smoother lines are moving averaged values with a 10-year averaging window.

at the coast that consists of partial land and partial ocean are usually a composition of results on both land and sea. Since only land surface results should be used in the G_S reconstruction process, for some GCMs, if the grid cell in which a borehole is located covers a large portion of the sea, an adjacent grid cell that covers over 90% of land is selected for data retrieving to keep data accuracy (e.g., λE will be affected by water surface area). The exact grid cell used for each borehole and its corresponding land cover portion in all 9 models is provided in Appendix C. In CMIP6, every GCM often provides an ensemble of outputs with different initial conditions, forcing conditions. Each ensemble member is given a variant id in the format 'r1i1p1f1', where the numbers are configuration indices of r-realization, i-initialization method, p-physics, and f-forcing. In this study, the first available ensemble member of each model is selected.

4.2.3 ERA5-Land reanalysis historical climate data

The temperature time series recorded in all three boreholes are filled with missing data and continuous measurements are usually available only for several years as presented in Fig. 4.2. The permafrost condition reflected from these short-length measurements may be affected by extreme climate events and deviate from the actual long-term average thermal regime. To get a more representative view of the historical thermal state of the permafrost, continuous soil temperature data at a climatic scale (over several decades) need to be provided. The data from climate reanalysis products have been widely used as long-term historical climate reference (*Graham et al.*, 2019; *Demchev et al.*, 2020; *Tarek et al.*, 2020; *Jiang et al.*, 2021). Consistent and continuous climate variables such as air and ground surface temperature, precipitation, and snow depth are derived by combining observations from the past with dynamic models. The fifth generation European Centre for Medium-Range Weather Forecasts (ECMWF) atmospheric reanalysis of the global climate and weather (ERA5) data set is one of the latest reanalysis products which provides high spatial and temporal resolution information covering the time period from 1950 (*Hersbach et al.*, 2020). Recently, the ERA5-Land has been produced which provides climate variables at a higher spatial resolution (0.1°×0.1°) over the land surface. In this study, hourly air temperatures at 2 meters above ground (T_{2m}) from the ERA5-Land data set at the grid cells that cover the borehole location (or adjacent grid cell if the original cell doesn't have data due to coverage of ocean) from 1981 to present (*Muñoz Sabater*, 2019) are downloaded directly from the Copernicus Climate Change Service Climate Data Store website (https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land?tab=overview). The downloaded hourly reanalysis air temperature data are then averaged to daily resolution and used to derive a long-term soil temperature at shallow depths combining the incomplete daily temperature measurements from the borehole sites (see section 4.3.1 for details).

4.3 Methods

4.3.1 Historical G_S reconstruction from borehole measurements

The reconstruction process details for this case are similar to the example described in Sec. 3.6.4 where MS site data are used. First of all, soil temperatures measured at the top two depths ($T_{S,1}$ at z_1, $T_{S,2}$ at z_2) are used to infer the MAP estimates for the HOD model parameters for each borehole site. To avoid encountering large data gaps (a few months) in the G_S reconstruction as well as the following calibration process (see Sec. 4.3.3), only the data for the time period in which measurements at all available depths are most continuous are used. Specifically, $T_{S,1}$ and $T_{S,2}$ data in [12/01/2006, 12/31/2010], [11/24/2010, 08/21/2013], and [01/10/2008, 08/12/2010] are used for the MS, VD, and BV site, respectively. The remaining occasional missing data for less than a few days are filled by linear interpolation.

The reconstructed G_S using $T_{S,1}$ and the inferred parameters with the HOD model for these time periods (G_S^{short}) are further used as the top boundary input in the numerical model calibration process.

After that, long term daily $T_{S,1}$ from 1981 to around 2015 ($T_{S,1}^{long}$) are derived from the ER5-Land daily T_{2m} series. Studies have shown that the surface ground temperature is highly correlated to the near-surface air temperature (*Tsilingiridis and Papakostas*, 2014; *Luo et al.*, 2018) at a daily scale and also related to other factors like topography and surface cover condition (*Kang et al.*, 2000). In this study, available $T_{S,1}$ measurements are divided into calibration and test periods at first. Then a simple linear regression between $T_{S,1}$ measurements in the calibration period and their corresponding T_{2m} (on the same day when $T_{S,1}$ is measured) is applied. $T_{S,1}^{long}$ is then derived based on the fitted linear relationship using the continuous T_{2m} data. $T_{S,1}$ measurements in the test period is then compared to the fitted $T_{S,1}^{long}$ to verify the linear regression quality.

Finally, the long-term surface ground heat flux from 1981 to 2014 (G_S^{long}) is obtained by applying the HOD model with $T_{S,1}^{long}$ and the MAP parameter estimates. G_S^{long} is further used as the historical observation data in the downscaling process (see Sec. 4.3.4).

4.3.2 Historical and future G_S reconstruction from GCM outputs

Since GCMs don't commonly provide G_S as a direct output, the reconstruction process established in Sec. 3.5.1 needs to be applied. Specifically, historical (1981-2014) net radiation (Rn), sensible (H) and latent (λE) heat fluxes, soil temperature (T_S), and moisture (θ, if available) are required during snowpack-free seasons (considering the effect of snow cover on surface energy balance, see Sec. 3.6.3 for details) in order to infer the MAP estimates for parameters used in the HOD model. For GCM # 4, #7, and #8 in Table 4.2, soil moisture data are unavailable, so the full saturation assumption described in Sec. 3.5.3 is applied. For the future time period (2015-2100), only T_S projection (and θ projection if available) at the top soil layer is needed from GCMs to compute the G_S projection for each warming scenario using the HOD model with the MAP parameter estimates from the historical period. The reconstructed daily G_S for each borehole site is averaged over several decades to get a representative climate for both the historical ($\bar{G}_S^{GCM,HIST}$) and future period ($\bar{G}_S^{GCM,FUT}$) and are used in the downscaling process (Sec. 4.3.4). Specifically, $\bar{G}_S^{GCM,HIST}$ is averaged over the period [1981, 2014]. While the future period is divided into three sections that are comparable to the historical period length and with three corresponding annual mean G_S: $\bar{G}_S^{GCM,FUT\,1}$ for [2015, 2040], $\bar{G}_S^{GCM,FUT\,2}$ for [2041, 2070], and $\bar{G}_S^{GCM,FUT\,3}$ for [2071, 2100].

4.3.3 Numerical model calibration

To simulate the future ground thermal state with the numerical model, its parameters need to be calibrated in advance using soil temperature measurements. The parameter inference process described in section 3.4.1 is used here to calibrate the numerical model parameters for each borehole site listed in 4.1. The same subsurface soil layer setups described in Sec. 3.6.4 and presented in Fig. 3.8 are used. The reconstructed G_S^{short} series from $T_{S,1}$ is applied as the boundary input. The calibration process is briefly reviewed hereafter. First, a suite of uncertain parameter sets used in the numerical model (see Sec. 3.6.4) is selected from their corresponding uniform priors. Second, the inference set is built from simulations using the developed parameter sets. Third, the observation and the inference set are concatenated head-to-tail for all available depths from the top to the bottom. Finally, the posterior distribution for the uncertain parameters is inferred by the UQ machinery using the concatenated series for the inference set and the observation. The MAP estimates derived from the posteriors are treated as the calibrated model parameters and then used throughout the projection simulations.

4.3.4 Bayesian Weighted Averaging

The BWA stochastic downscaling approach proposed by *Tebaldi et al.* (2005) is applied in this chapter to reduce the bias and uncertainty inherited in GCMs outputs. The original BWA model is briefly described in this section. Let X_i and Y_i be the outputs for the variable of interest from the ith GCM for the historical and future period, respectively. In this study, X_i is the historical annual mean ground heat flux ($\bar{G}_S^{GCM,HIST}$) , and Y_i is for one of the future periods ($\bar{G}_S^{GCM,FUT\,1}$, $\bar{G}_S^{GCM,FUT\,2}$, or $\bar{G}_S^{GCM,FUT\,3}$). These outputs are assumed to follow the Gaussian distribution:

$$X_i \sim N(\mu, \lambda_i^{-1}), \tag{4.1}$$

$$Y_i \sim N(\nu, (\theta\lambda_i^{-1})), \tag{4.2}$$

where μ and ν are the true values of historical and future annual mean G_S. λ_i is the reciprocal of the variance and named as the *precision* of the distribution of X_i and is related to the weight for the ith GCM. θ is a common parameter for all GCMs to allow a different precision of Y_i and X_i.

Let X_0 be the historical observations for the same variable of interest (i.e.,) and follows Gaussian distribution:

$$X_0 \sim N(\mu, \lambda_0^{-1}), \tag{4.3}$$

The precision λ_0 represents the natural variability to the observation X_0 and is related to its weight, while λ_i depends on specific GCM configurations such as parameterization, grid resolution, and numerical solutions for physical processes. In this study, λ_0

is a fixed constant and determined from historical data. Eqs. (4.1)-(4.3) describe the distribution of the GCM results and the observation providing the parameters μ, ν, θ, λ_i, and λ_0, which are essentially the *likelihood* functions in the Bayesian approach.

Next, uninformative *prior* distributions for the parameters are selected as follows:

$$\mu, \nu \sim U(-\infty, \infty), \tag{4.4}$$

$$\lambda_i \sim Ga(a, b), \tag{4.5}$$

$$\theta \sim Ga(c, d), \tag{4.6}$$

where U is the uniform distribution and Ga the Gamma distribution, with $a = b = c = d = 0.01$. The joint *posterior* distribution can then be derived by applying Baye's theorem:

$$posterior \propto prior \times likelihood. \tag{4.7}$$

Since the posterior distribution cannot be explicitly derived from the analytical form calculated by Eq. 4.7 (Eq. (4) in *Tebaldi et al.* (2005)), the Markov chain Monte Carlo (MCMC) method is used to sample the empirical posterior distribution. Details of the MCMC simulations are provided in the Appendix in *Tebaldi et al.* (2005). Specifically, a total of 75,000 iterations was simulated with 25,000 of them used as the 'burn-in' period in this study. The sampling results are saved every 50 iterations to reduce the correlation between successive samples, leading to 1000 samples which are used to generate the posterior distribution at the end for each parameter.

An approximation of the expected values for λ_i can be described as:

$$E(\lambda_i) \approx \frac{a+1}{b + \frac{1}{2}[(X_i - \tilde{\mu})^2 + \theta(Y_i - \tilde{\nu})^2]}, \tag{4.8}$$

where $\tilde{\mu}$ and $\tilde{\nu}$ are the mathematical expectation of μ and ν, respectively. As a result, the quantities $\|X_i - \tilde{\mu}\|$ and $\|Y_i - \tilde{\nu}\|$ correspond to the *bias* and *convergence* criteria, respectively (*Tebaldi et al.*, 2005). The statistical model described in Eqs. (4.1)-(4.3) is made more robust by introducing a linear correlation between the historical and future climate response within a GCM by rewriting Eq. (4.2) as:

$$Y_i = \nu + \beta_x(X_i - \mu) + \xi_i/\sqrt{\theta}, \tag{4.9}$$

where $\beta_x \sim U(-\infty, \infty)$ is the correlation factor between $X_i - \mu$ and $Y_i - \nu$, and $\xi_i \sim N(0, \lambda_i^{-1})$. The expectation of λ_i in Eq. (4.8) is updated as:

$$E(\lambda_i) \approx \frac{a+1}{b + \frac{1}{2}\{(X_i - \tilde{\mu})^2 + \theta[(Y_i - \tilde{\nu} - \beta_x(X_i - \tilde{\mu})]^2\}}. \tag{4.10}$$

By setting Y_i to $\bar{G}_S^{GCM,FUT\,1}$, $\bar{G}_S^{GCM,FUT\,2}$, and $\bar{G}_S^{GCM,FUT\,3}$ separately, one can derive the posterior distribution for each future period section ($\bar{G}_S^{post,FUT\,1}$, $\bar{G}_S^{post,FUT\,2}$, and $\bar{G}_S^{post,FUT\,3}$). In addition, by selecting the representative projected $\bar{G}_S$ at the same percentile (e.g., 50th percentile or the median) from these 3 posteriors, a continuous annual mean G_S representing the future climate from 2015 to 2100 can then be constructed by linear interpolation between the three percentile based selections of $\bar{G}_S$.

4.4 Results

4.4.1 Linear fitted long-term shallow borehole temperatures

Subplots in the first, second, and third rows in Fig. 4.4 present the long-term borehole temperature construction results for the borehole sites MS, VD, and BV, respectively. The $T_{S,1}$ measurements and the fitted long term soil temperature $T_{S,1}^{long}$ at z_1 are compared in Fig. 4.4(a)-(c). The calibration periods (green segments) defined for all borehole sites are MS [08/31/2006, 08/28/2012], VD [09/12/2010, 08/23/2013], and BV [08/12/2007, 08/13/2008]. Fig. 4.4(d)-(f) compares the air temperature at 2 m height from ERA5-Land data products T_{2m} and the $T_{S,1}$ observations in the calibration period in scatter format along with their corresponding linear regression (red lines). Fig. 4.4(g)-(i) tests the quality of the linear regression constructed in the calibration periods by comparing $T_{S,1}$ and $T_{S,1}^{long}$ in the test periods (red segments in Fig. 4.4(a)-(c)) in both scatter and linear regression (red lines) format. For MS and VD sites, the linear fitted long-term soil temperature at z_1 is close to the actual measurement indicated by a high value of R^2 for the data in the test periods. For the BV site, the linear fitting methods don't perform as well as for the other two sites. This issue may due to the relatively fewer data points for this site, or the highly non-linear relationship between air temperature and ground temperature (*Shati et al.*, 2018).

4.4.2 Reconstructed historical G_S

Results of historical G_S reconstruction and parameter inference for the MS site are presented previously in Sec. 3.6.4 (Fig. 3.10 and 3.11). Similar results for the VD and BV sites are presented in this section. Subplots (c)-(e) in Fig. 4.5 and 4.7 present the MAP estimates for parameters used in the HOD model. Their values are in the typical ranges of soil properties (see Sec. 3.6.2 for details of specific values). The

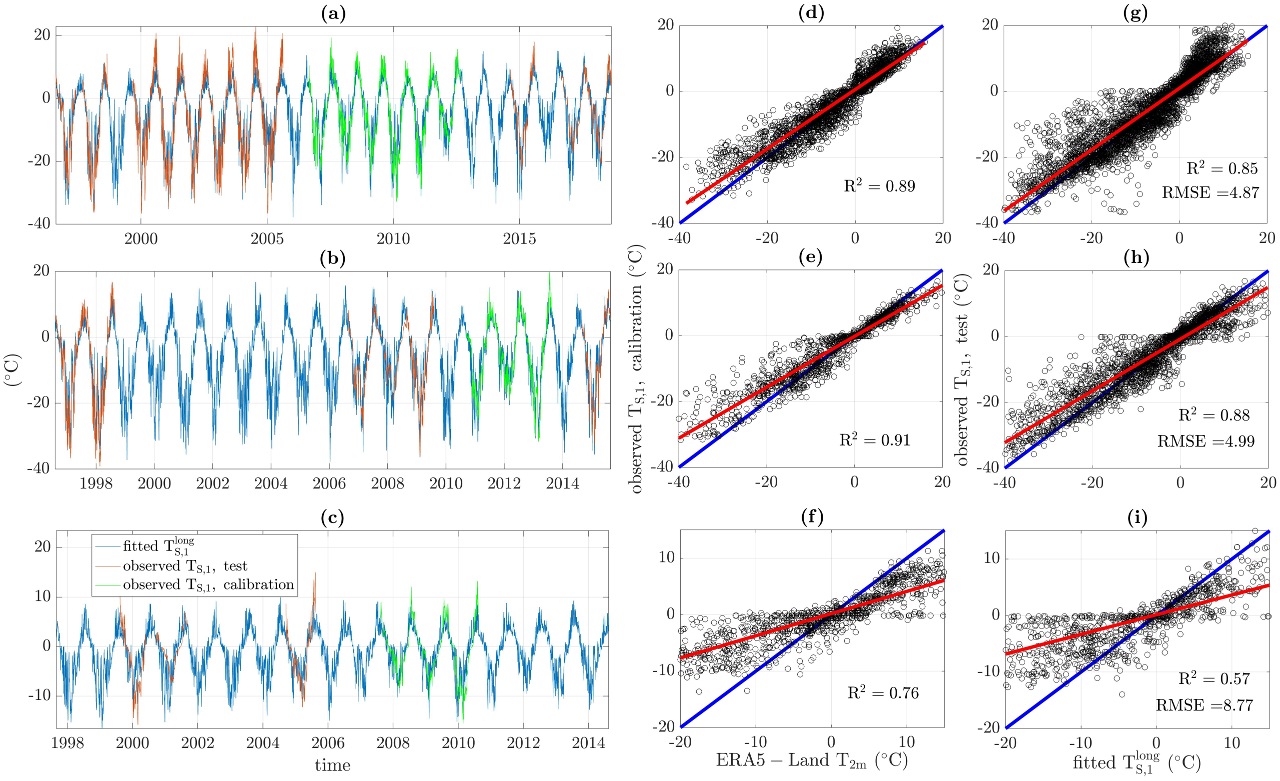

Figure 4.4: (a)-(c) Comparison between the observed soil temperature $T_{S,1}$ linear fitted soil temperature $T_{S,1}^{long}$ (blue lines) at z_1 for both the calibration (green lines) and the test (red lines) periods for the MS, VD, and BV site, respectively. Subplots (d)-(f) compare the ERA5-Land products of air temperature at 2 m height T_{2m} and the $T_{S,1}$ during the calibration period for each site. The linear regression (red lines) information for each site: (d)-MS, slope = 0.90, intercept = 0.60; (e)-VD, slope = 0.77, intercept = -0.18; (f)-BV, slope = 0.39, intercept = 0.19. "R^2" is the coefficient of determination of the regression. Subplots (g)-(i) compare the fitted soil temperature $T_{S,1}^{long}$ and observed $T_{S,1}$ in the test period in scatter format along with their corresponding linear regression (red lines). "RMSE" is the root mean square error between $T_{S,1}^{long}$ and $T_{S,1}$.

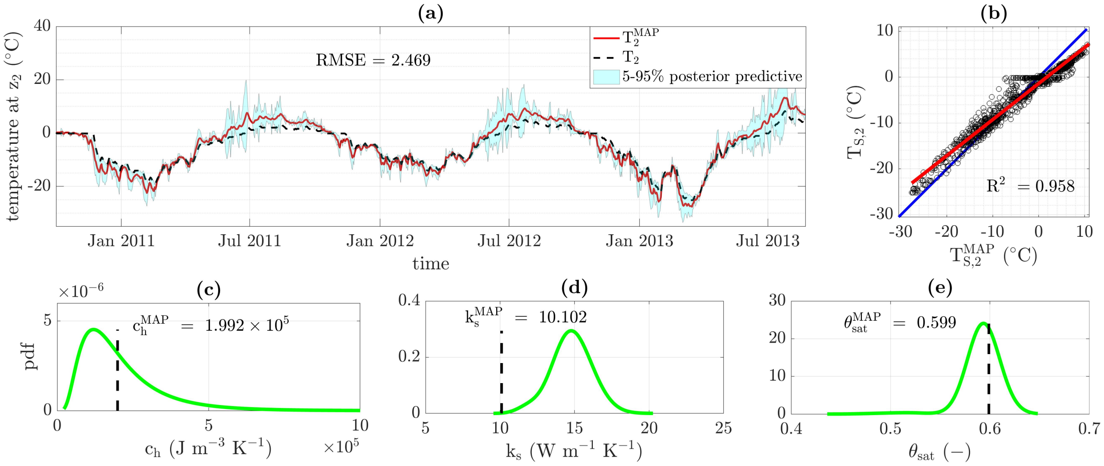

Figure 4.5: Similar results in Fig. 3.10 but for VD site. (a) The series of measured soil temperature T_2 (at $z_2 = 50$ cm, black dashed line) and T_2^{MAP} (at $z_2 = 50$ cm) computed from T_1 measured at $z_1 = 1$ cm. The observational data are from Nov. 2010 to Aug. 2013 at the daily resolution. (b) The scatter plot and linear regression (red solid line) fitted for T_2 and T_2^{MAP} (the blue solid line is 1:1) with slope = 0.786 and intercept = -1.386. (c)-(e) Kernel distribution of marginal posteriors for c_h, k_s, and θ_{sat} (green curves) and MAP estimations, c_h^{MAP}, k_s^{MAP}, and θ_{sat}^{MAP} (vertical black dashed lines).

long-term historical ground heat flux $\tilde{G}_S^{MAP}$ from 1981 to 2014 is computed using the HOD model with these MAP parameter estimates and displayed in Fig. 4.6(a) and 4.8(a). The $\tilde{G}_S^{MAP}$ results are zoomed in the time period when T_1 measurements are available (e.g., black dashed lines in Fig. 4.6(b) and 4.8(b)) to show more details and to make results comparable. Besides, $\tilde{G}_S^{MAP}$ in this time period (G_{short} in Sec. 4.3.1) is used as the top boundary input for the numerical model parameter calibration (see Sec. 4.4.3 for results).

4.4.3 Calibrated numerical model

Fig. 4.9 and 4.10 compare the simulated and observed soil temperatures at all available measurement depths down to 10 m within the selected time period (see Sec. 4.3.1) for the VD and BV sites. Results provide a reference for the credibility of the reconstructed $\tilde{G}_S^{MAP}$. The linear regression between simulated and observed

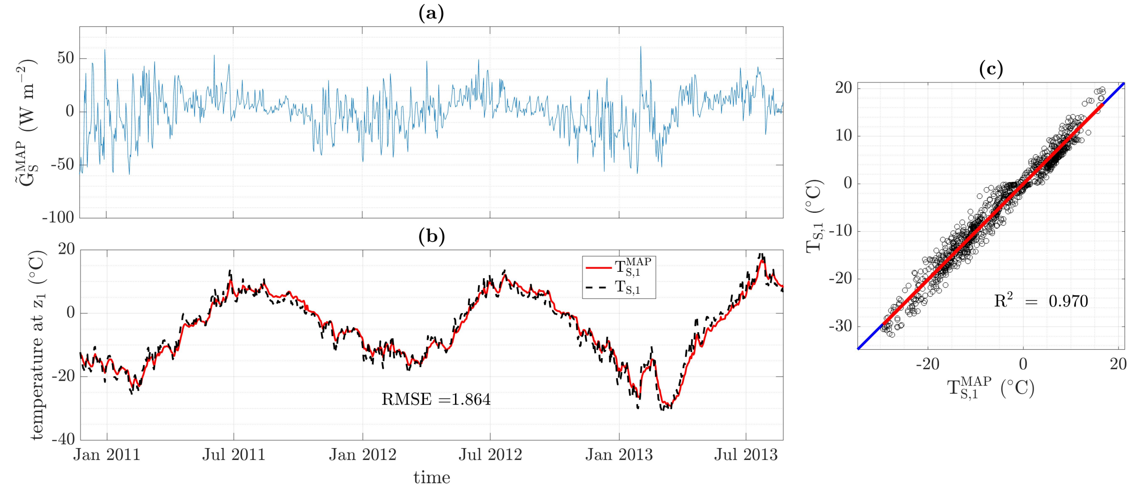

Figure 4.6: (a) The reconstructed $\tilde{G}_S^{MAP}$ series for VD site. (b) A comparison of soil temperature T_1 measured at $z_1 = 1$ cm and T_1^{MAP} derived from the $\tilde{G}_S^{MAP}$ series in (a) for the VD borehole site from Nov. 2011 to Aug. 2013 at a daily scale. The time series of $\tilde{G}_S^{MAP}$ is computed as the product $\alpha^{RMSE} G_1^{MAP}$, where the time series G_1^{MAP} is calculated with Eq. (3.1) using the MAP parameter set (see Fig. 4.5) and the observed T_1 series. The minimum RMSE = 1.864 (°C) between T_1 and T_1^{MAP} is obtained by using the value of a scaling factor $\alpha^{RMSE} = 1.138$. (c) A scatter plot and linear regression (red solid line) between T_1 and T_1^{MAP} with slope = 1.004 and intercept = -0.036 (the blue solid line is 1:1).

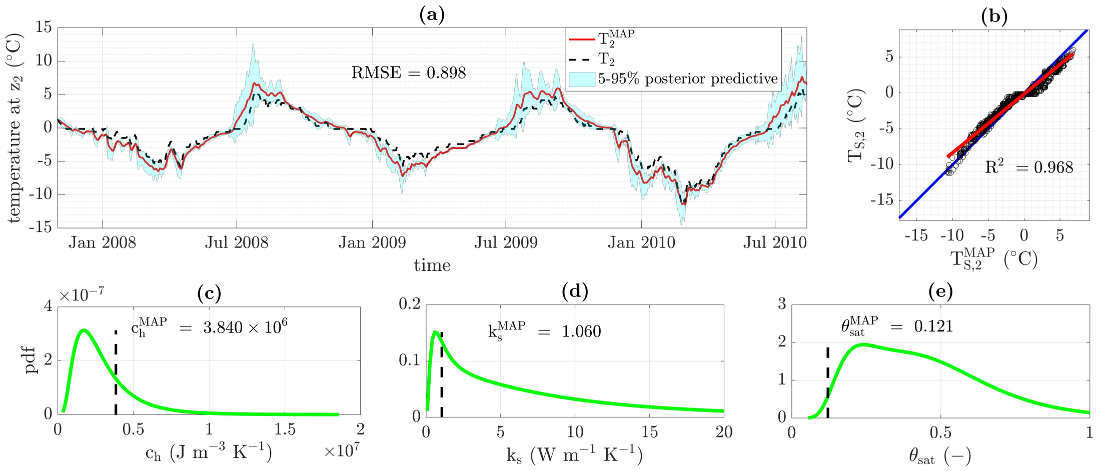

Figure 4.7: Same type of results as in Fig. 4.5 for the BV borehole site. Soil temperature T_1 in (a) is measured at $z_1 = 10$ cm from Jan. 2008 to Aug. 2010. Linear regression (red solid line) in (b) with slope = 0.821 and intercept = -1.334.

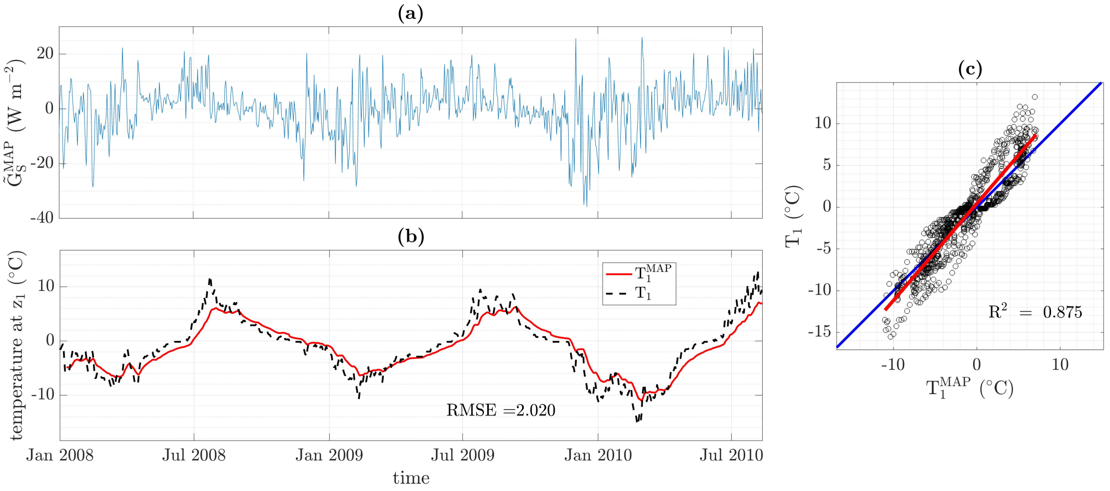

Figure 4.8: Same type of results as in Fig. 4.6 for the BV borehole site. T_1 in (b) is measured at $z_1 = 10$ cm. Linear regression (red solid line) between T_1 and T_1^{MAP} in (c) with slope = 1.161 and intercept = -0.444.

temperatures at 1 m at the BV site (Fig. 4.10(i)) doesn't have a high R^2 value due to the small number of observations. For results at 3 to 5 m, the low R^2 values in Fig. 4.10(j)-(k) could be related to the nearly constant observations, since the R^2 value is more sensitive to little oscillations in the simulation for this case. The inferred numerical model parameters here are used in the long-term projection simulation runs in the future.

4.4.4 BWA projection of annual mean $\bar{G}_S$

The BWA method (described in Sec. 4.3.4) is applied to project the annual mean $\bar{G}_S$ for the future combining the long-term historical G_S^{long} reconstructed from borehole measurements ($\tilde{G}_S^{MAP}$ in Fig. 3.11(a), 4.6(a), and 4.8(a)) and historical $G_S^{GCM,HIST}$ and future $G_S^{GCM,FUT}$ reconstructed from GCM data products under different future climate scenarios (see Sec. 4.3.2). After applying the G_S reconstruction process, 4 GCMs with a worse match between the G_S from GCM outputs and HOD model simulations are not used in the stochastic downscaling process The GCM outputs used for downscaling are from model #1 to #5 listed in Table 4.2. Figs. 4.12-4.17 present

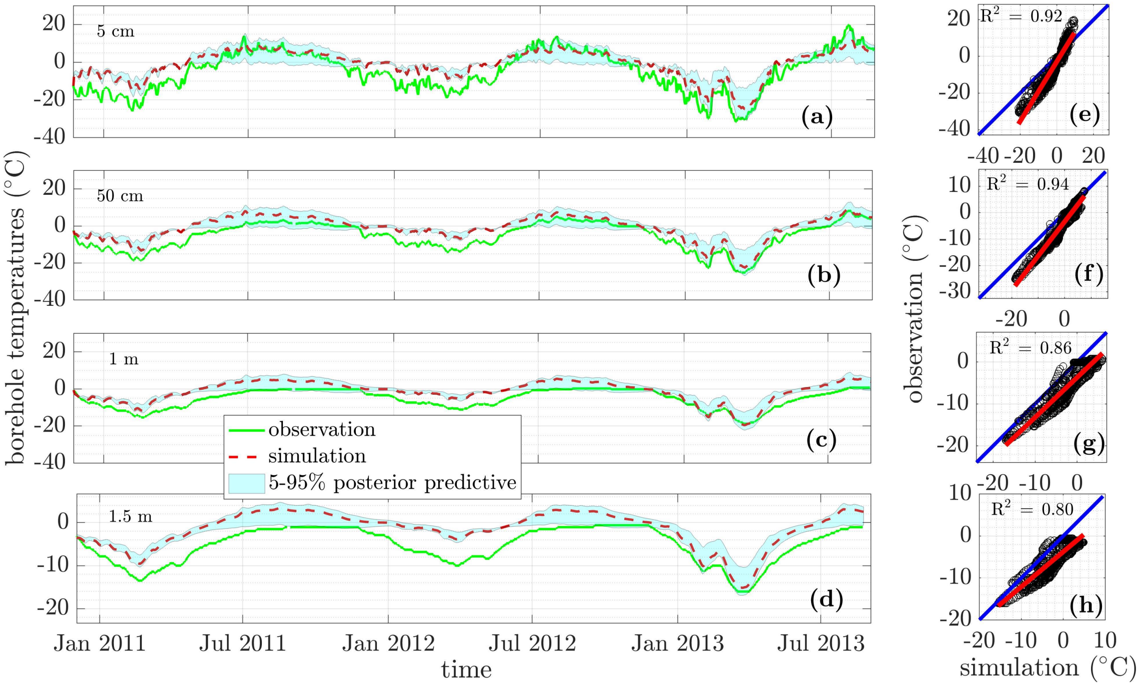

Figure 4.9: (a)-(d) A comparison of the temperature time series based on simulations using the numerical model described in Sec. 3.3 using $\tilde{G}_S^{MAP}$ in Fig. 4.6(a) and the daily temperature measurements at different depths at Vaskiny Dachy from Nov. 2010 to Aug. 2013. The cyan areas are 5-95% posterior predictive bounds. (e)-(h) The corresponding scatter plots and linear regression (red solid line) between simulated and observed temperature.

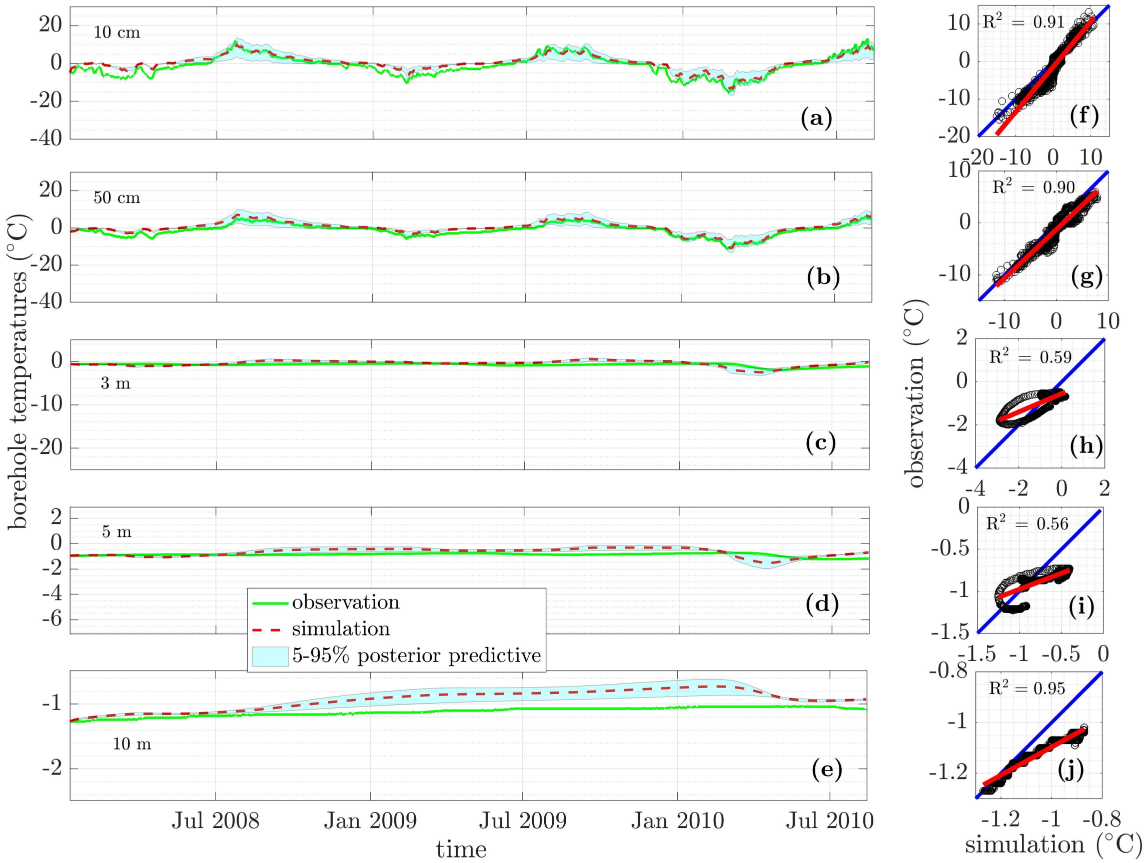

Figure 4.10: The same contents as in Fig. 4.9 for the Bolvansky borehole site from Jan. 2008 to Aug. 2010.

the BWA posterior distribution of the annual mean $\bar{G}_S$ for historical (subplots (a)) and three future time periods (subplots (b)-(d)) under the SSP2-4.5 and SSP5-8.5 projected future climate scenarios.

Continuous $\bar{G}_S$ from 2015 to 2100 for both future scenarios are then generated through linear interpolation and extrapolation with representative statistics values (every 10th percentile) selected from these posterior distributions. Results are presented in Fig. 4.18. The red lines in all subplots are the interpolated $\bar{G}_S$ based on the median values (i.e., 50th percentile) of the posteriors for the three future periods at the same borehole site under the same projected future scenario. The grey dashed lines from the bottom to the top are generated similarly for the 10th to 90th percentile to show the $\bar{G}_S$ variability related to the downscaling approach. Projection of $\bar{G}_S$ gradually decreases for the MS and BV site for the SSP2-4.5 scenario, while in other cases it decreases first and then starts to increase. In the SSP2.4-5 scenario at the MS site, the $\bar{G}_S$ even drops below 0 W m^{-2}, which means the permafrost depth will begin to increase. This can be partially explained by the general slowing down warming trends projected for the air temperature for the SSP2-4.5 scenario in Fig. 4.3(a), (c), and (e). Such a phenomenon is not observed for the projected air temperature under the SSP5-8.5 scenario. The future CO_2 emissions for different SSP scenarios in the IPCC report (*Masson-Delmotte et al.*, 2021) provide indirect evidence of such a decrease in the warming trend of SSP2-4.5 (Fig. 4.11).

Next, the long-term future projection of the permafrost thermal regime is conducted by running the numerical model with the interpolated $\bar{G}_S$ corresponding to each percentile value as the top boundary input.

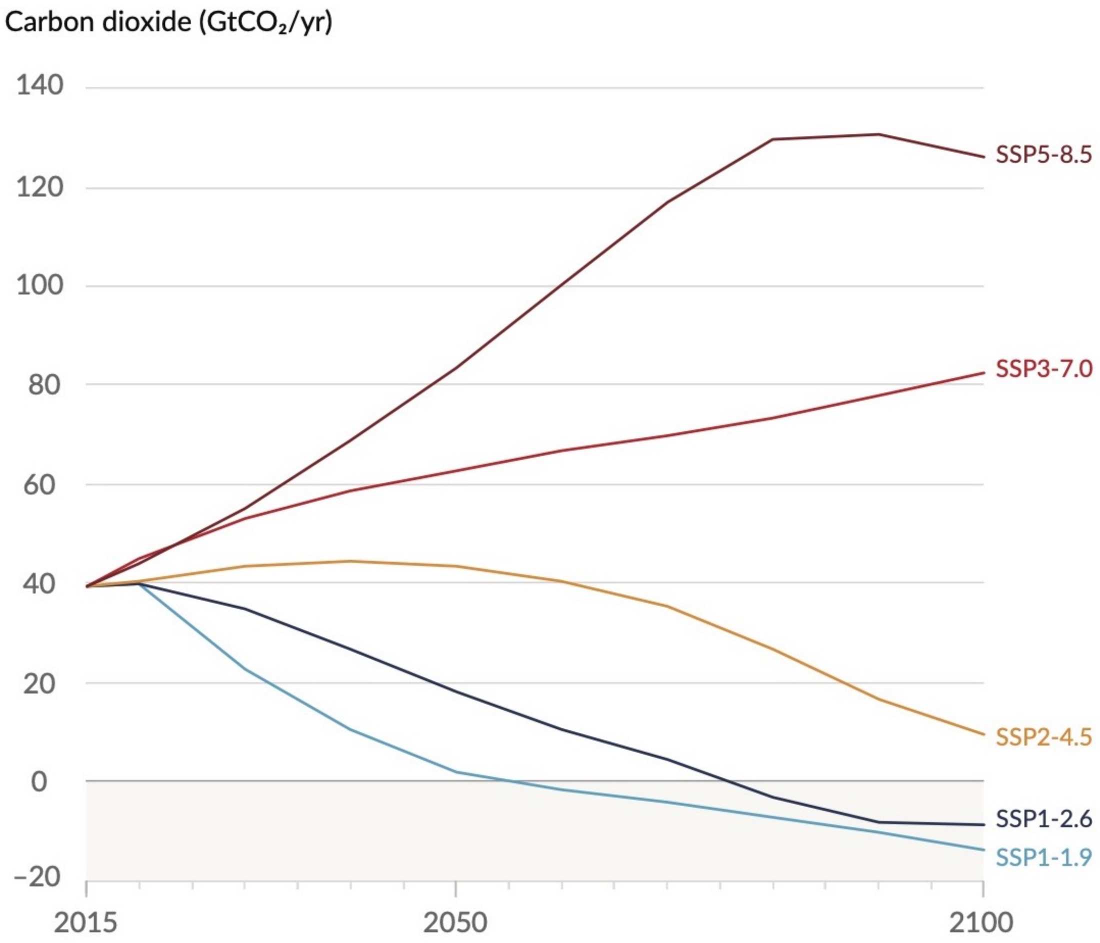

Figure 4.11: Future annual emissions of CO_2 across five illustrative scenarios (*Masson-Delmotte et al.*, 2021).

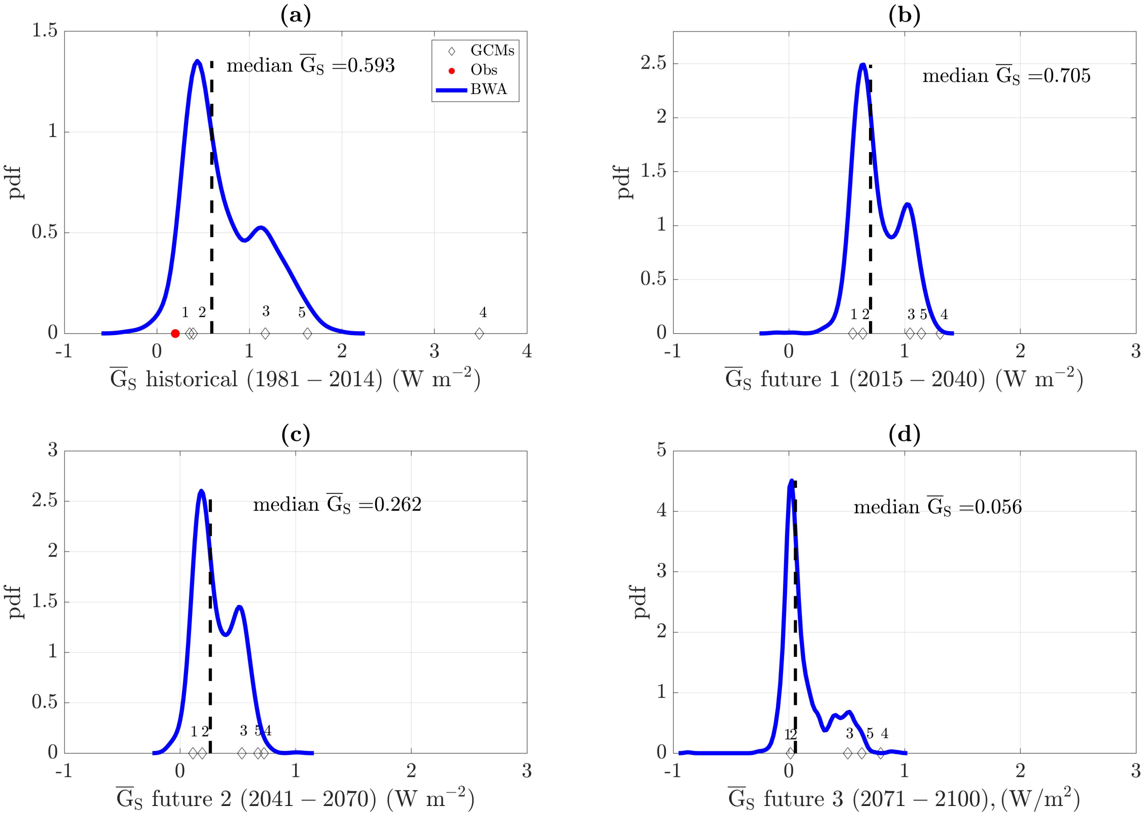

Figure 4.12: The posterior distribution generated by the BWA method for the annual mean $\bar{G}_S$ (blue curves) for (a) the historical from 1981 to 2014 and future (b) from 2015 to 2040, (c) from 2041 to 2070, and (d) from 2071 to 2100 period under the projected SSP2-4.5 future scenario for the MS site. The median values (vertical black dashed lines) for each distribution are indicated as representative statistics. Historical observation (red dot) and historical and future GCMs outputs (void diamonds) for $\bar{G}_S$ are plotted on the x-axis. The number above the void diamond corresponds to the same GCM model number in Table 4.2.

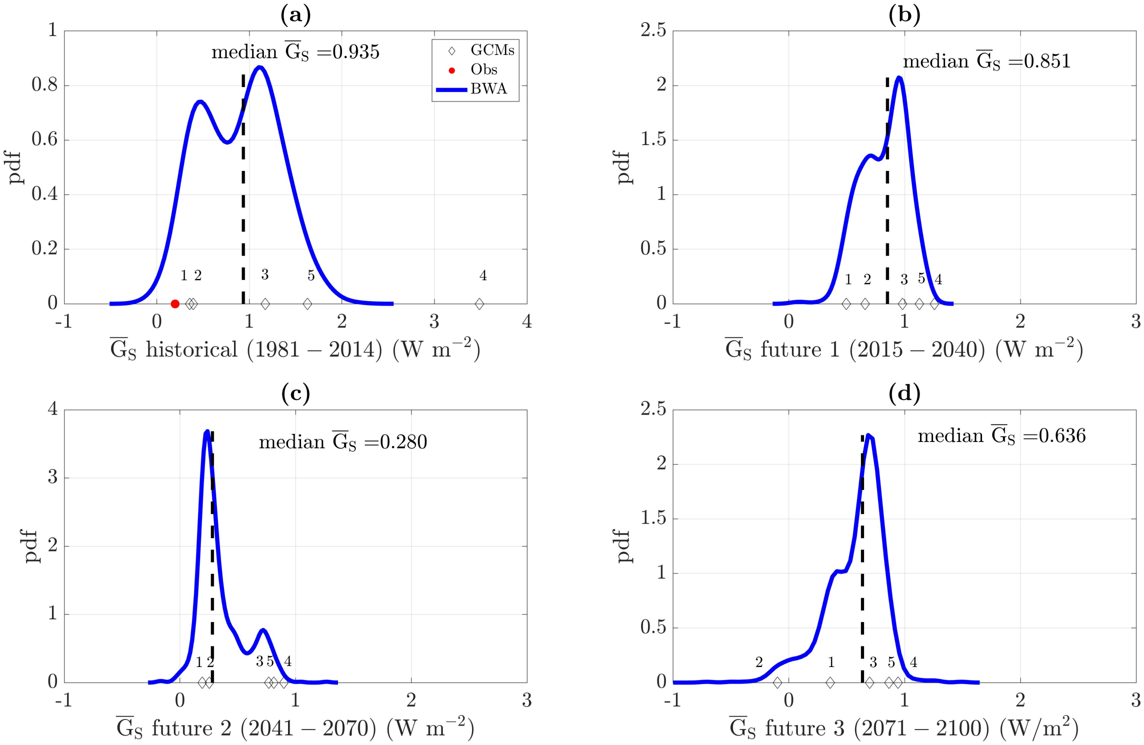

Figure 4.13: Same contents as in Fig. 4.12 for the MS site under the SSP5-8.5 scenario.

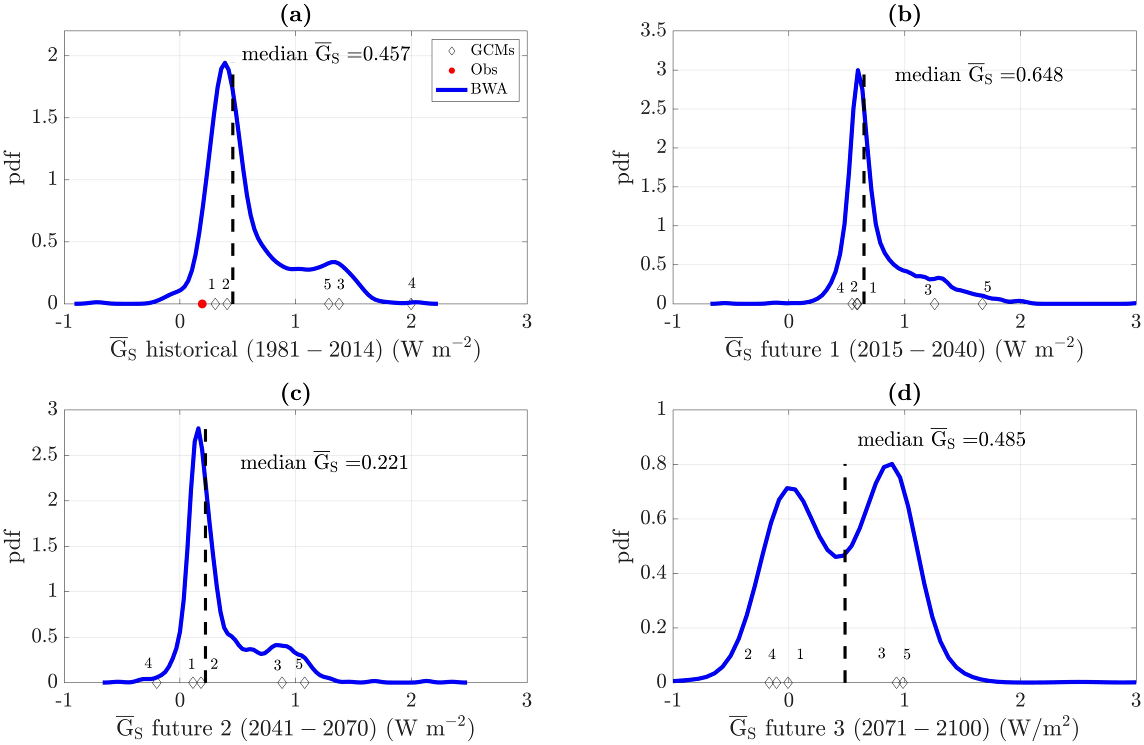

Figure 4.14: Same contents as in Fig. 4.12 for the VD site under the SSP2-4.5 scenario.

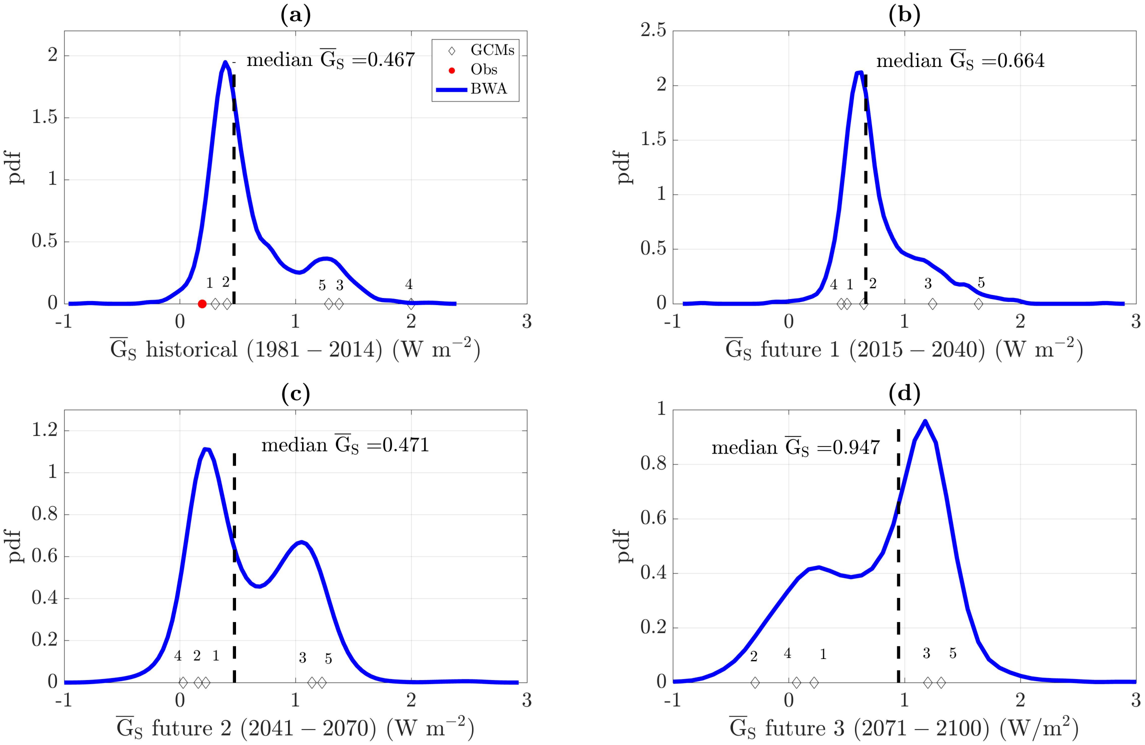

Figure 4.15: Same contents as in Fig. 4.12 for the VD site under the SSP5-8.5 scenario.

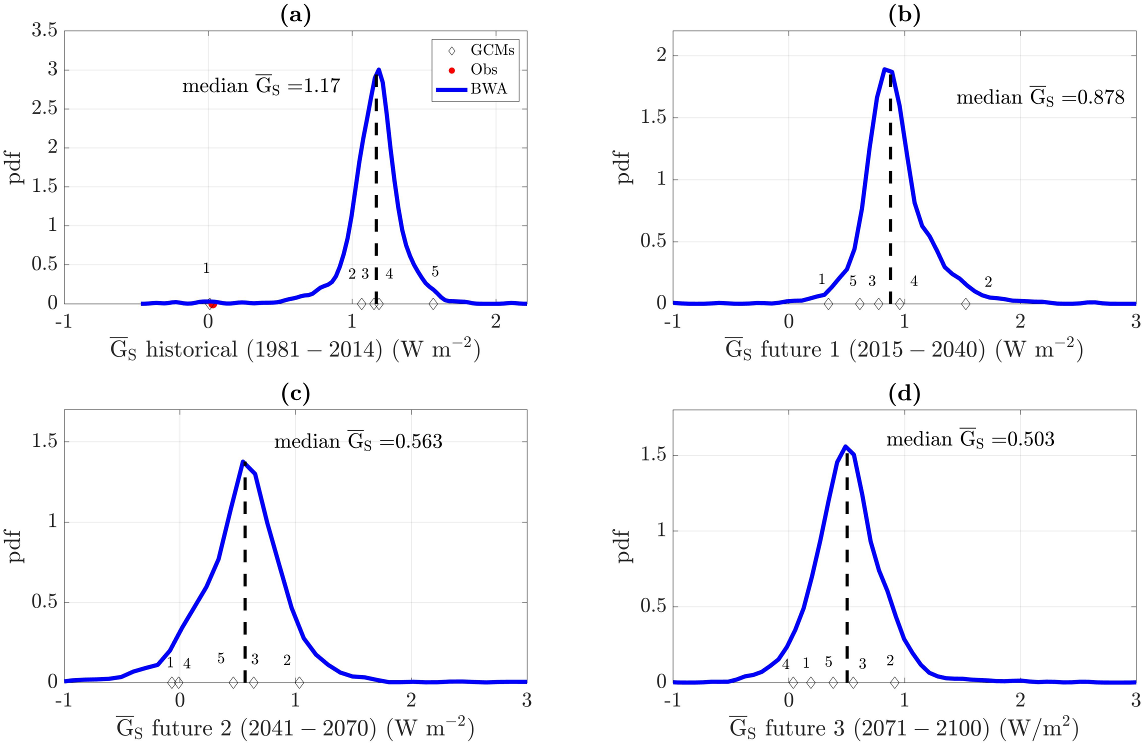

Figure 4.16: Same contents as in Fig. 4.12 for the BV site under the SSP2-4.5 scenario.

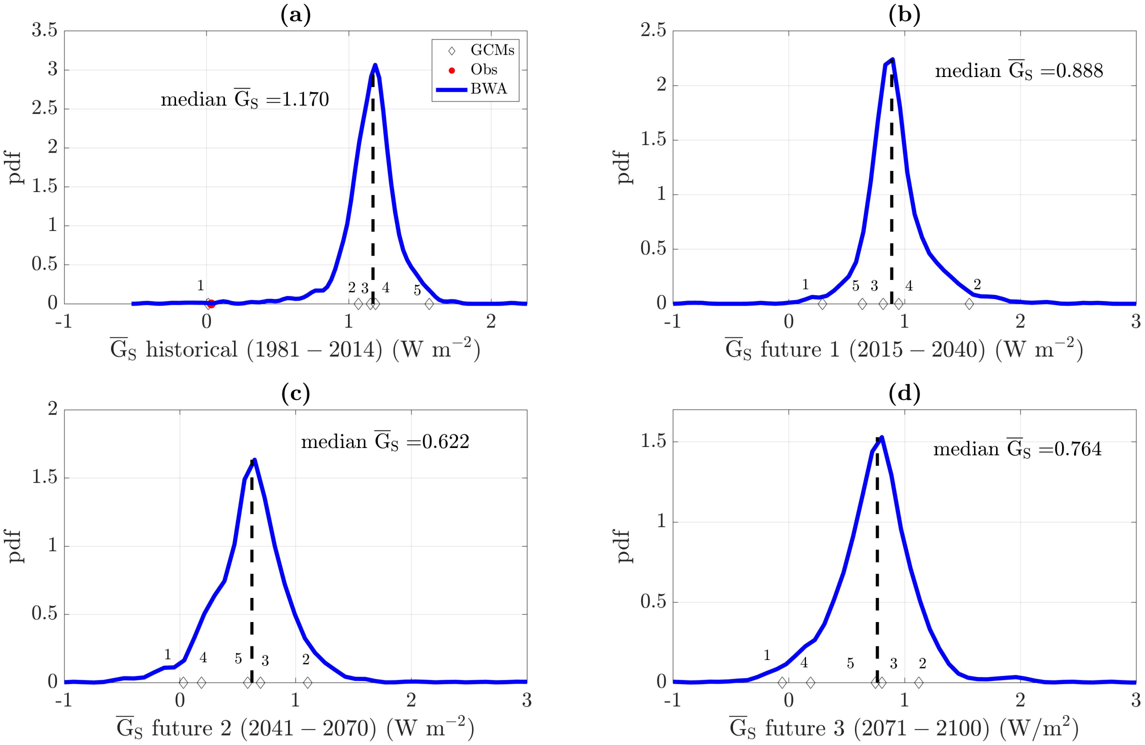

Figure 4.17: Same contents as in Fig. 4.12 for the BV site under the SSP5-8.5 scenario.

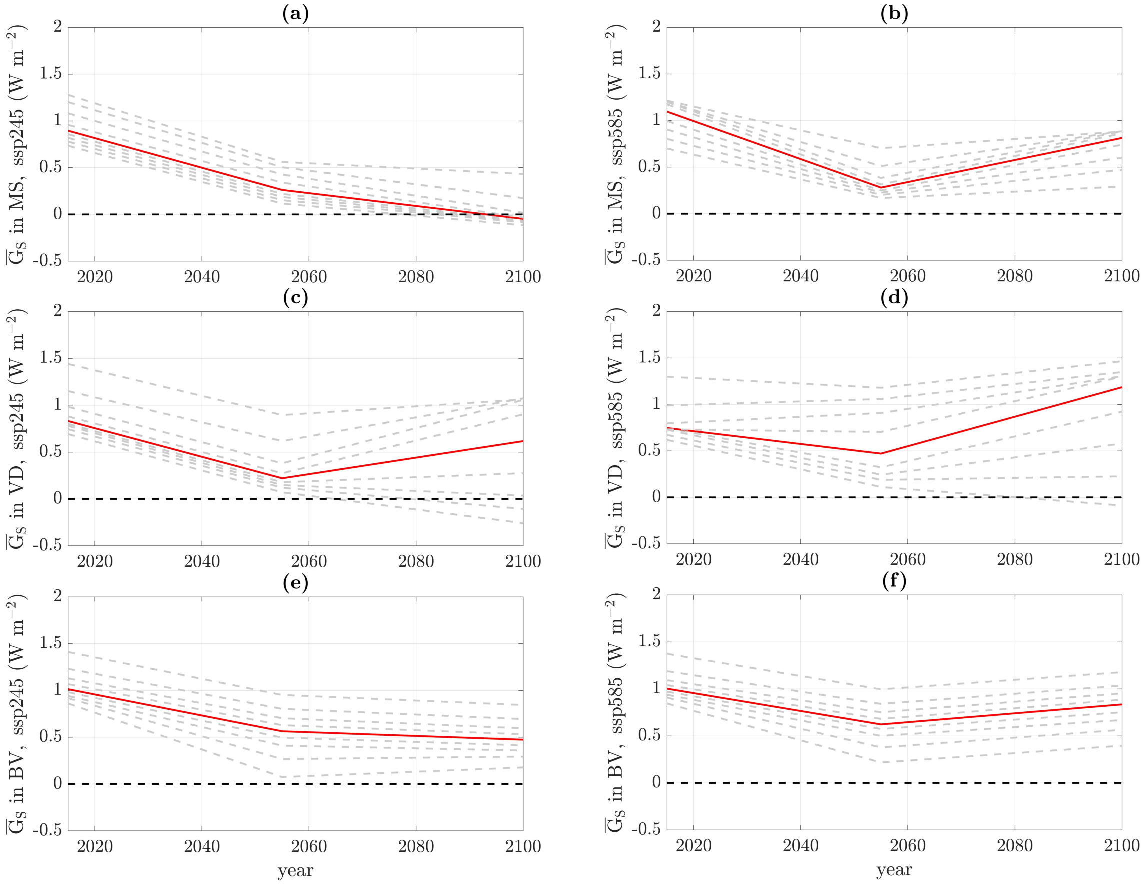

Figure 4.18: Interpolated annual mean $\bar{G}_S$ for every 10th percentile from the BWA posteriors (Fig. 4.12-4.17) on three future periods at the (a)-(b) MS, (c)-(d) VD, and (e)-(f) BV sites. The left column corresponds to the results for the SSP2-4.5 scenario, and the right column is for SSP5-8.5. Red lines in all subplots represent the interpolated $\bar{G}_S$ based on the 50th percentiles from corresponding BWA posteriors. Grey lines are the 10th to 90th percentile results. Black lines are the 0 W m^{-2} line.

4.4.5 Simulation of future subsurface temperature profiles

Fig. 4.19 presents the subsurface soil temperature change based on numerical model results using the median reconstructed annual mean $\bar{G}_S$ (red lines in Fig. 4.18) as the top boundary input from 2015 to 2100 at the (a)-(b) MS, (c)-(d) VD, and (e)-(f) BV sites. Results in the left column represent the SSP2-4.5 scenario and the right column SSP5-8.5 scenario. Based on the simulations, the most significant increase in the soil temperature happens in the soil located at the top 100 m. Due to a more intensive warming climate projection, the surface soil temperature increase is larger in the SSP5-8.5 scenario compared to the results for SSP2-4.5. Among the three borehole sites, surface soil temperature at the BV site increases at a faster rate and reaches the highest value of 6.49 °C at the surface in 2100 under the SSP5-8.5 scenario, while for the MS and VD sites, these values are 0.86 and 3.16 °C, respectively.

Series of the surface soil temperature ($\bar{T}_{S,0m}$) and the soil temperature at 10 m ($\bar{T}_{S,10m}$) at the (a) MS, (b) VD, and (c) BV borehole sites based on model results are presented in Fig. 4.20 and 4.21, respectively. Soil condition at around 10 m depth is important for infrastructure foundation construction and hence $\bar{T}_{S,10m}$ is presented here as a proxy. Results in the case where the median $\bar{G}_S$ is used as the boundary input (red and blue lines) are considered hereafter as representative. Overall, $\bar{T}_{S,0m}$ and $\bar{T}_{S,10m}$ increase more for the SSP5-8.5 scenario than SSP2-4.5. At the MS site, $\bar{T}_{S,0m}$ doesn't go$\bar{T}_{S,10m}$ above 0 °C for SSP2-4.5, and $\bar{T}_{S,10m}$ is almost certain to be under 0 °C for both scenarios. While for the BV site, both $\bar{T}_{S,0m}$ and $\bar{T}_{S,10m}$ are almost certain to be larger than 0 °C. The surface soil temperature increases $\Delta\bar{T}_{S,0m}$ from 2015-2100 at the three sites are: MS, 1.14, VD, 3.58, and BV, 3.87 °C for the SSP2-4.5 scenario and MS, 3.84, VD, 6.28, and BV, 6.79 °C for SSP5-8.5. The results variance at the VD site is larger compared to that at the other two sites displayed by the wider range of shaded areas.

Fig. 4.22-4.24 show the vertical annual mean soil temperature profile change from 2015 to 2100 using the median $\bar{G}_S$ as simulation input. The annual mean soil temperature at top soil layers increases to above 0 °C at the end of this century for all sites, and all scenarios, except at the MS site for the SSP2-4.5 scenario. Specifically, the depth of 0 °C annual mean soil temperature are VD, 7.5, and BV, 34.8 m for SSP2-4.5; MS, 4.5, VD, 27.4, and BV, 40.0 m for SSP5-8.5. To derive the active layer thickness (ALT) or the seasonal thaw depth (STD), information on the seasonal oscillation of G_S needs to be provided. ALT and STD cannot be derived directly due to the lack of such data in this study. However, considering the seasonal variation

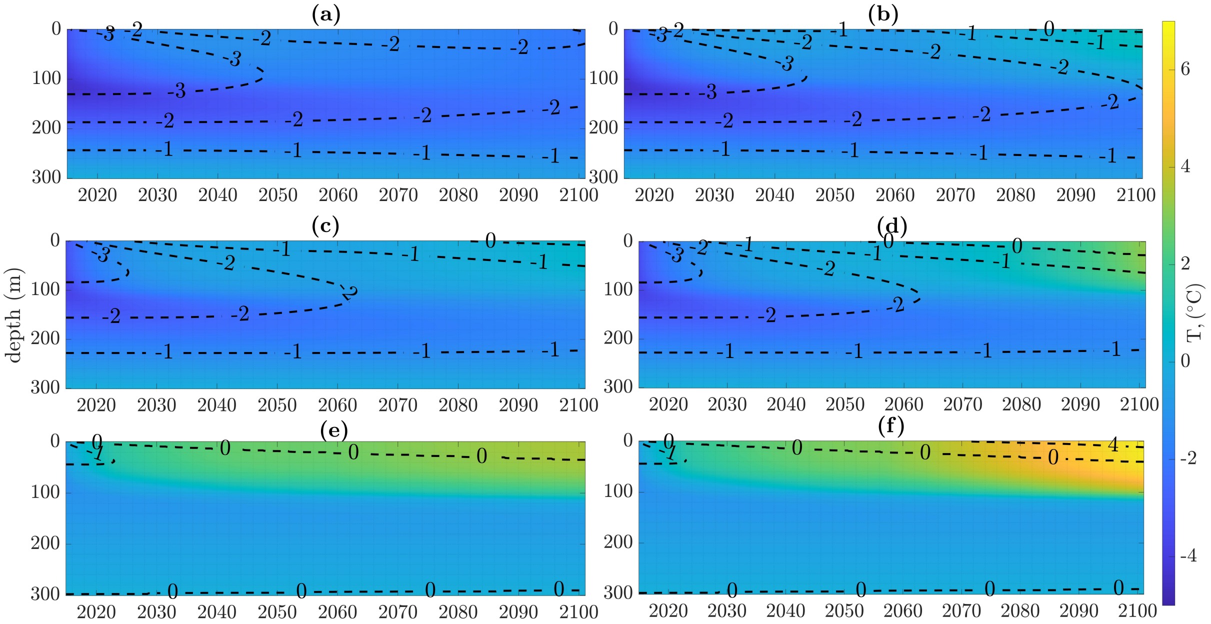

Figure 4.19: The change of soil temperature profile from 0 to 300 m in the projection period (i.e., 2015-2100) for the (a)-(b) MS, (c)-(d) VD, and (e)-(f) BV sites. Results in the left column correspond to SSP2-4.5 and the right column is for SSP5-8.5. Contour lines represent the same temperatures across years.

of the soil temperature, the ALT or STD will only be larger than these depths. The soil temperature change at the bottom of the permafrost is a mixed result of the constant positive geothermal flux input and the heat diffusion solution applied in the numerical model.

4.4.6 Uncertainty analysis

For the permafrost thermal regime projections carried out in this study, uncertainty could come from the following sources: (1) imperfection of surrogate or mechanistic model and approximation made in the KL expansion used in the parameter inference process for G_S reconstruction; (2) MAP estimates generated from posteriors for parameter inference; (3) variability of GCM outputs after the downscaling procedure; (4) defect of specific stochastic downscaling methodology; (5) impacts of extreme climate events on observations with limited length in time. (1)-(3) are ad-

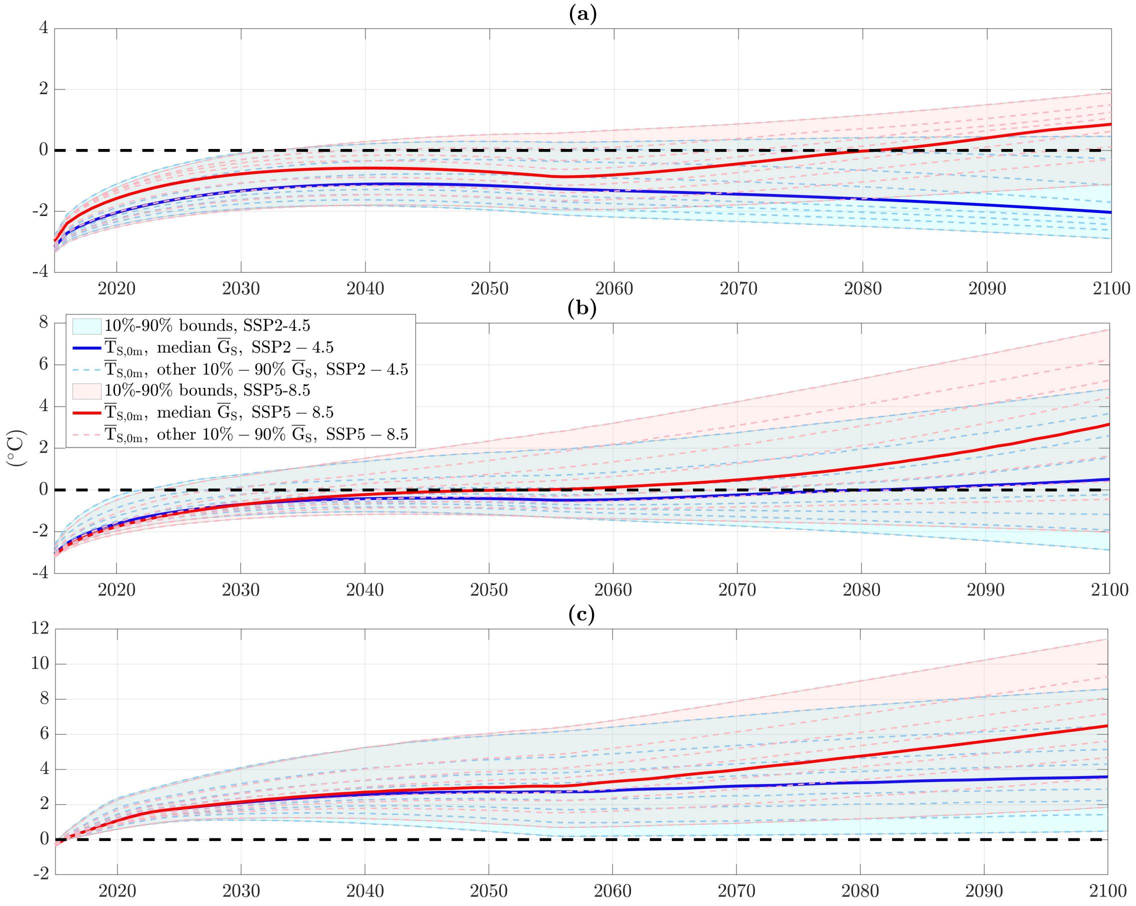

Figure 4.20: The surface soil temperature ($\bar{T}_{S,0m}$) series for the projection period (i.e., 2015-2100) at the (a) MS, (b) VD, and (c) BV sites under the SSP2-4.5 (blue lines) and SSP5-8.5 (red lines) scenarios. The shaded areas are bounded by $\bar{T}_{S,0m}$ series derived from the 10th and 90th percentile $\bar{G}_S$. $\bar{T}_{S,0m}$ derived from other percentile $\bar{G}_S$ are plotted as dashed lines. Black lines are the 0 °C line.

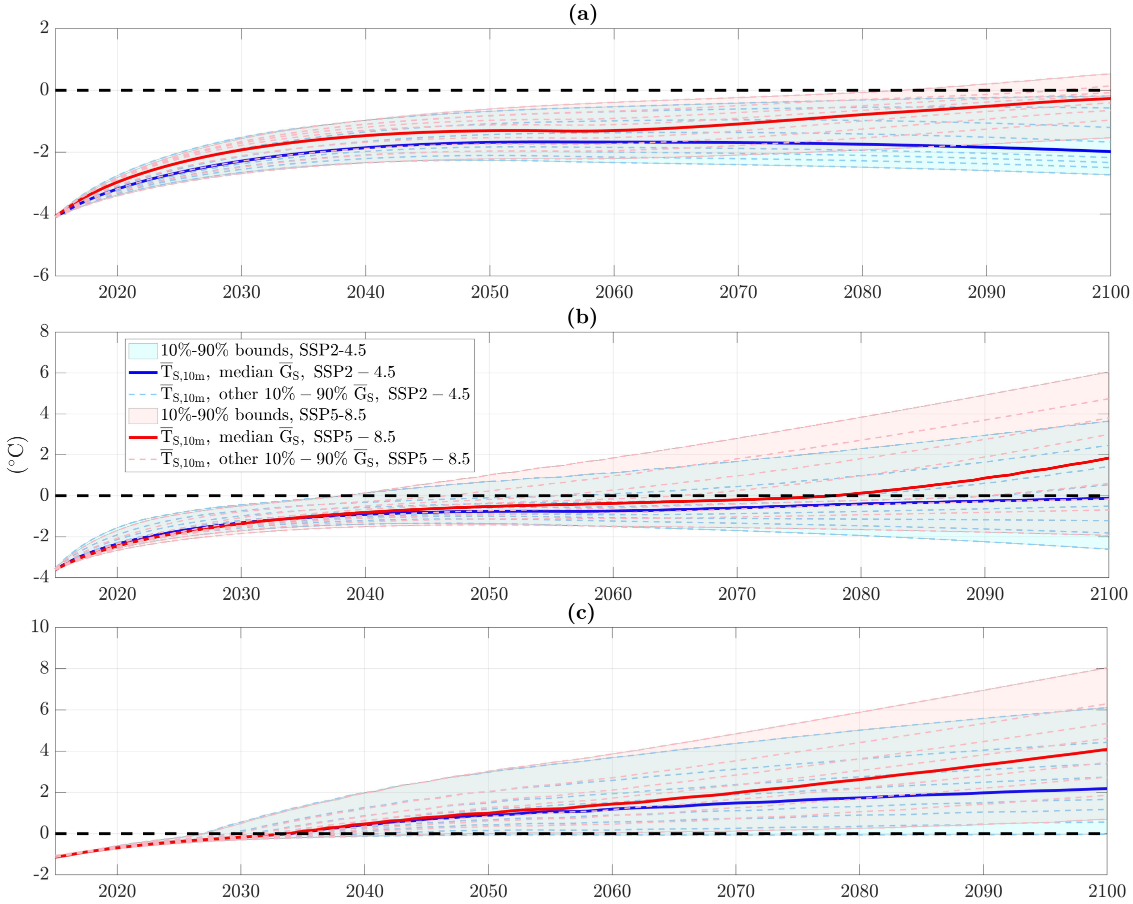

Figure 4.21: The same content for the soil temperature at 10 m ($\bar{T}_{S,10m}$) series for the projection period (i.e., 2015-2100) at the (a) MS, (b) VD, and (c) BV sites under the SSP2-4.5 (blue lines) and SSP5-8.5 (red lines) scenarios.

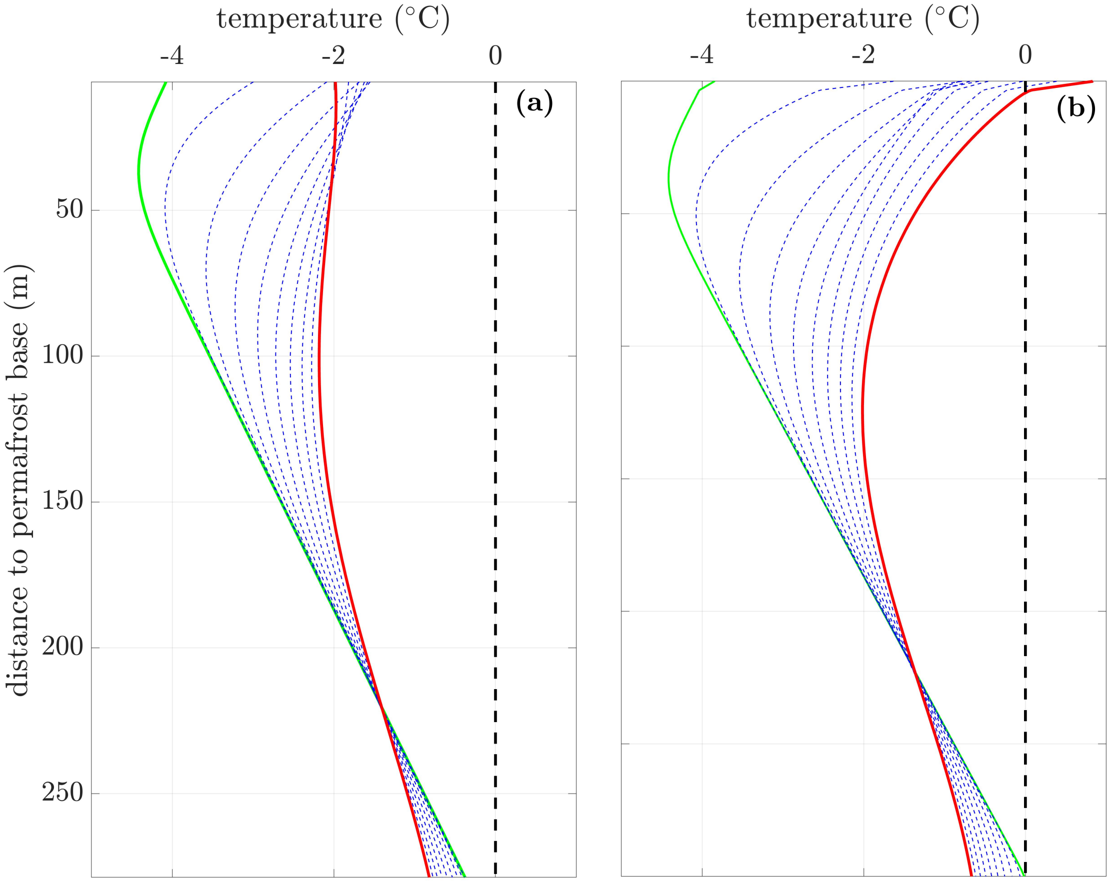

Figure 4.22: Simulated soil temperature profile along the depth from 0 to 300 m depth (permafrost base) at the MS site for (a) SSP2-4.5 and (b) SSP5-8.5 using the median projected $\bar{G}_S$ as the boundary condition. Green curves are profiles in 2015, red curves are in 2100, and blue dashed curves are in 2020 to 2090 with a 10 years step. Black vertical lines are the 0 °C line.

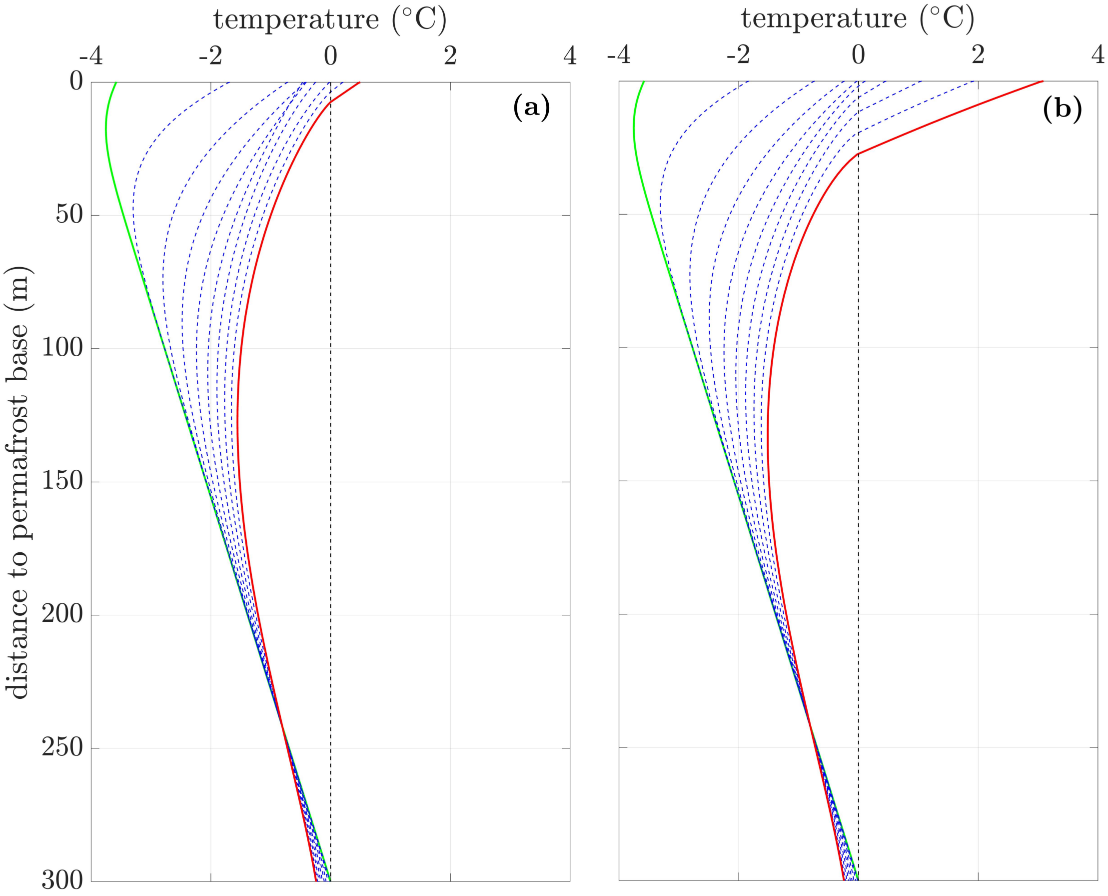

Figure 4.23: The same content in Fig. 4.22 for VD site.

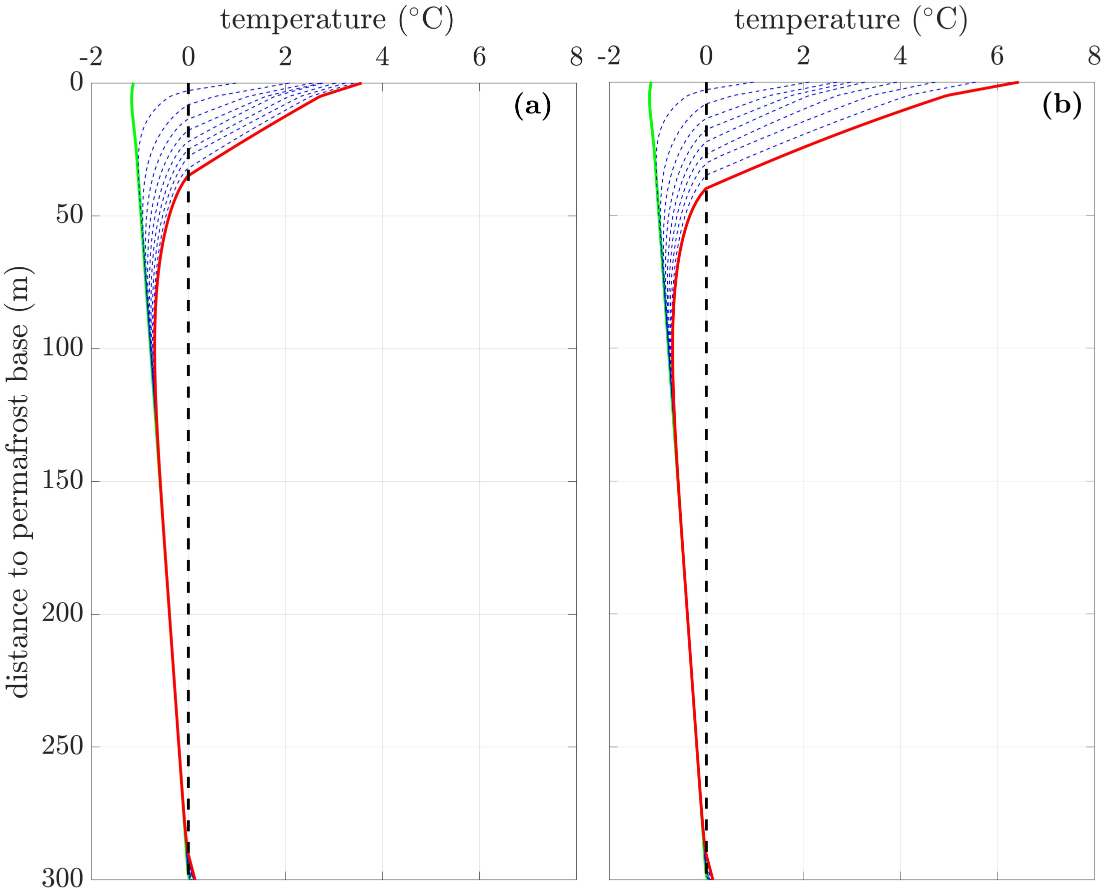

Figure 4.24: The same content in Fig. 4.22 for BV site.

dressed below. (4) and (5) along with other types of uncertainty sources such as personnel or instrument bias inherent in borehole temperature measurements, configuration of different GCMs, and emission forcing uncertainty for different SSP scenarios are beyond the scope of this study.

The quality of KL expansion in terms of the bias between the original observation and estimation computed by Eq. 3.26 is presented in Appendix B. It relies on the assumption of viewing the mechanistic model simulations as pure stochastic processes (see Sec. 3.4.3), and also depends on the mechanistic model performance in mimicking the actual physical process. The surrogate models built for the mechanistic model in the parameter inference process to improve computational efficiency generally provide good results in this study. More details about the surrogate modeling uncertainties are discussed in the literature (e.g., *Sargsyan et al.*, 2014; *Dwelle et al.*, 2019). The stand-alone performance of the HOD models (described in Sec. 3.2) is tested in *Wang and Bras's* work (1999) and the 1-D freeze-thaw model (described in 3.3) is validated in Appendix A.

After the parameter inference process, a posterior distribution for mechanistic model parameters is constructed. A single Maximum A Posterior (MAP) set of estimates is often selected for the following application using the model. The variability associated with the posterior is presented in Fig. 3.12, 4.9, and 4.10 for the numerical model calibration process described in Sec. 3.6.4. If a smaller range of percentile (e.g., 5-95%) is selected, these bounds will be even narrower and almost not distinguishable from the results with MAP parameter sets.

The shaded area in Fig. 4.20 and 4.21 represents the variability of $\bar{T}_{S,0m}$ and $\bar{T}_{S,10m}$ considering the uncertainty carried in the top boundary input $\bar{G}_S$ which is constructed by interpolating percentile values from the BWA posterior distributions for three future time periods. The uncertainty bounds increase as the model projects to the future and the impacts of the initial setting gradually disappear. The uncertainty related to the structure of downscaling methodology is discussed in the literature (*Chen et al.*, 2011; *Kim et al.*, 2016; *Xu et al.*, 2019) and is not addressed here.

4.5 Discussion

The soil temperature measurements used in this study are usually available only for several years and contain a lot of missing data (see Fig. 4.2). These observations could be impacted by annual or seasonal extreme climate perturbations and thus cannot represent the actual long-term thermal state of the permafrost in the past. A simple linear regression is fitted in Sec. 4.4.1 between the ERA5-Land reanalyzed air temperature and available soil temperature data at shallow depths. The regression works well at the MS and VD sites, but not at BV (Fig. 4.4). Previous studies have stated that the correlation between the air and ground temperatures can be highly nonlinear (*Tsilingiridis and Papakostas*, 2014; *Luo et al.*, 2018; *Shati et al.*, 2018), thus the linear regression may not be a good method to rebuild the long-term soil temperature. Other potential ways to construct the soil temperature from above-ground information such as machine learning (*Feng et al.*, 2019; *Alizamir et al.*, 2020) need to be tested and applied. Additional measurements of continuous soil temperature and moisture profile along the depth in this region are suggested in the future to make a more comprehensive and convincing analysis.

The numerical model parameters are calibrated using the limited historical soil temperature from borehole observations in order to project the permafrost thermal condition to the future (Sec. 4.3.3). The calibrated model parameters are then treated as constant throughout the simulation runs assuming no change in the soil properties until the end of this century. This may be a strong assumption considering the fast-changing ambient conditions above-ground in the Arctic and the resulting shift of the underground hydrologic regime at the time scale considered in the analysis. Besides, the relationship between soil properties such as thermal conductivity to the change of temperature is also not included in this study (*Johansen*, 1977; *Romanovsky and Osterkamp*, 2000). Notice that for the model calibration results at the BV site, the simulated soil temperatures below 50 cm are always higher than the observations. This mismatch is a mixed result of such a lack of measurements and improper setup of the numerical model that makes the inference set unable to cover the observation. Such overestimation of temperature indicates that the future projection of soil temperatures at the BV site can be overestimated as well, which adds another source of uncertainty to the results.

To apply the BWA downscaling method, an ensemble consists independent GCMs needs to be formed first (*Tebaldi et al.*, 2005). Since the variables (e.g., daily soil temperature) required to reconstruct the surface ground heat flux in this study are not available for most of the existing GCMs in the CMIP6 program, only 9 models are selected for data retrieving. Furthermore, based on the performance after applying the developed G_S reconstruction process (described in Sec. 3.5) for the obtained GCM outputs, only 5 models are used to form the ensemble used in the downscaling procedure. In fact, this number may be too small to make it a representative GCMs ensemble for climate projection purposes. Consequently, the innate bias carried by these GCMs may not be properly addressed. Notice that GCM #2 and # 5 are from the same institute using the same land model component, the assumption on model independence may be invalid. Besides, the soil layer settings for these GCMs may not be suitable for permafrost analysis. For example, GCM # 2 and # 5 only have 4 soil layers with a maximum depth of 2.0 m, which is way shallower than the actual permafrost thickness. Hence the bias and uncertainty of the soil temperature simulations derived from these models could be extreme.

In Fig. 4.18, the 10-90th percentile for projected $\bar{G}_S$ at the BV site is always above 0 W m^{-2}, which indicates that the permafrost will almost certainly decrease at the BV site for both emission scenarios. This high certainty of permafrost degradation in the BV site can be explained by its lower latitude than the other sites (Fig. 4.1). The range of the 10-90th percentile for $\bar{G}_S$ at the VD site is wider than the other two sites as a result of the wider posterior distributions generated in the downscaling process (Fig. 4.12-4.17), which means larger uncertainty for the projection results at this site. The possibility of permafrost refreezing at the VD site implied from the negative 10-30th percentile $\bar{G}_S$ in (c) should take this high uncertainty into consideration. The negative $\bar{G}_S$ in Fig. 4.18 is also a result of the interpolation methods used to create the long-term projection. If a single $\bar{G}_S$ averaged over the 86 years (2015-2100) or a step function consisting of the median of the three future periods is used, then the projected annual average ground heat flux would always be positive, which would result in a slow down of permafrost warming rather than a refreeze.

Among the three sites, the annual mean surface temperature $\bar{T}_{S,0m}$ at the BV site increases at a much higher rate than the other two sites and reaches 0 °C first before 2020 for both scenarios. At the VD site, $\bar{T}_{S,0m}$ reaches 0 °C at 2082 and 2054 for the SSP2-4.5 and SSP5-8.5 scenario, respectively. For the MS site, $\bar{T}_{S,0m}$ doesn't even go above 0 °C for the SSP2-4.5 scenario. As a matter of fact, the temperature profile at the BV site is warmer than the other two sites in 2015 (Fig. 4.22-4.24), which could be a result of the already higher warming rate before 2015 due to its low latitude. The exact reason that causes the faster warming rate of $\bar{T}_{S,0m}$ at the VD site than that at MS still needs to be analyzed. Compared to the MS site, the VD site is located in a more interior place, so a possible cause would be the influence of the ocean.

If the seasonal change of soil temperature profile is neglected (which is proper considering the time scale of the problem), almost one-seventh of the permafrost at the BV site will disappear in 2100 under the SSP5-8.5 scenario (median projected $\bar{G}_S$). The permafrost thickness decreases as well in a smaller amount at the other sites for all scenarios, except at the MS site for SSP2-4.5. However, the annual mean surface temperature still increases by 1.14 °C in that case.

4.6 Conclusion

In this chapter, the projection of the permafrost thermal state in Northwestern Siberia is conducted by applying a physical-based numerical model under different emission scenarios with uncertainty being constrained. Projected annual mean surface ground heat flux $\bar{G}_S$ inferred from a framework adopting uncertainty quantification machinery with historical borehole observations and outputs from GCMs is used as the boundary input for the numerical model. To reduce the bias and uncertainty contained in GCM outputs, a stochastic downscaling method is applied. Historical $\bar{G}_S$ reconstructed from borehole temperature measurements is treated as an observational reference in the downscaling process. The largest changes of annual mean surface soil temperature based on the simulation for two of the borehole sites under the SSP5-8.5 scenario are higher than the global surface temperature change of around 4.7 °C provided in the IPCC report (*Masson-Delmotte et al.*, 2021) for the same scenario in 2100, which indicates the sensitivity and vulnerability of the Arctic over other regions under the impact of the climate change. Almost one-seventh of the permafrost is going to thaw at one of the borehole sites in 2100 under the worst carbon emission scenario, while permafrost at other sites also disappears at lower levels. Specifically, permafrost at the site with lower latitude is undergoing a thawing rate about 1.1 to 3.4 times faster than other high-latitude sites. This significant loss of permafrost will cause huge greenhouse gas emissions into the atmosphere, destabilize the foundation of infrastructure and change the above and below-ground hydrologic and thermal dynamics. The simulation results in this study provide a reference for a plausible future of the permafrost thermal regime and the associated feedback of this change and emphasize the urgent need to control the warming trend in the Arctic.

CHAPTER V

Research Summary and Future Studies

5.1 Summary of research

The Arctic hydrological and ecological systems are experiencing drastic shifts for the past few decades due to climate change. How would the Arctic change in the future influence the global climate as well as the low latitudinal regions due to its crucial role in global climate circulation (*Hinzman et al.*, 2005)? The answer essentially relies on historical and ongoing trends of these changes and is complex to disentangle considering the yet unclear interactions between various components in the Arctic terrestrial system. Before making a conclusion for this question in a comprehensive and systematic way, it is important to address the changes happening in each individual component and understand their underlying cause by analyzing the mechanistic processes. Specifically, efforts are made in this thesis for a better understanding of two of the most significant processes in the Arctic: vegetation dynamics above-ground, and subsurface permafrost state and related processes.

Evidence from historical observations has shown a vegetation encroachment into the Arctic tundra area non-homogeneously happening in space and time. Besides the driving factors such as increasing summer temperature and growing season length, other factors need to be considered to address the nonuniform characteristic of this process at the regional or local scale. In Chapter II, the relationships between tall vegetation expansion and two important factors, topography and snow distribution, are discussed based on available long-term historical data from field campaigns at a plot scale. Tall vegetation shows a preference to grow in locations with divergent surface characteristics, which are explained in terms of combining the positive effects of well-drained soils and more favorable climate over negative effects such as windswept land surfaces and summer drying for tall vegetation survival. The change of clustered tree expansion patterns presented in the study plot is also explained through its interaction between snow distribution shifts and the teleconnection of tree growth in different spaces and times.

Observation from borehole measurements indicates the warming trend in permafrost across the circumpolar region for decades. One of the key factors that control the subsurface freezing and thawing processes, the surface ground heat flux G_S, has received little discussion in the literature for future climate scenarios. Direct measurements for G_S are challenging and cumbersome in harsh environments such as the Arctic. Besides, large-scale products from GCMs don't provide direct outputs for G_S and its estimate through surface energy balance is only good for the snowpack-free season. A framework is established in Chapter III to reconstruct G_S and its probability distribution for all seasons from shallow soil temperature measurements which are more commonly measured with other auxiliary data (e.g., summer time G_S observations and soil moisture). The framework combines a heat-transfer-based analytical model which relates ground heat flux to soil temperature and the state-of-art UQ machinery in model parameter inference. The reconstructed G_S from the developed process is a useful boundary input for long-term subsurface thermal regime simulation.

To make a projection of the future permafrost thermal state in the study area utilizing the framework proposed in Chapter III, outputs for required variables are retrieved from qualified CMIP6 GCMs under two projected future climates with different SSP and emission scenarios. The bias and uncertainty of the G_S derived from the selected GCM ensemble are reduced by applying the BWA stochastic downscaling method combining historical observational G_S reconstructed from borehole temperature measurements. Simulation results using a physical-based numerical model with the downscaled G_S as top boundary forcing are presented in Chapter IV for three boreholes in Northwestern Siberia until the end of this century. The projected average ground surface temperature for two of the boreholes in the study region is generally higher than the global surface temperature projection provided in the IPCC reports (*Masson-Delmotte et al.*, 2021) in 2100 for the same SSP scenario. Results in this study emphasize the importance of ongoing permafrost monitoring as well as the urgent need to control carbon emissions and constrain the global warming trend.

5.2 Research assumptions and limitations

The assumptions and limitations applied for each chapter are listed below.

- Chapter II
  - Change in the stem diameter is used as a proxy for tree growth. This is based on the high correlation between sapwood area and assume sapwood is a proper indicator of biomass.
  - Only current DEM is used to derive surface topography features with the assumption of no significant geomorphology change in the study area since the beginning of the campaigns.
  - Tree distribution is assumed to be not limited by seed dispersion in the study plot. This is supported by field observation.
  - The snow distributions mapped from the three campaigns may not be representative of long-term snow conditions. Additional continuous snow mapping is suggested for future works.

- Chapter III
    - The HOD model used in G_S reconstruction is derived from analytical solutions of the heat transfer equation assuming a semi-infinite soil column with uniform soil properties and excluding phase change energy.
    - The surrogate and mechanistic model errors are neglected when evaluating the performance of KL expansion estimates using a few eigenvalues and eigenfunctions.
    - Data noise in constructing the KL-PC estimates is assumed to be fixed. A more proper way would be inferring it within the MCMC process.
    - In the case where soil moisture data are unavailable (e.g., from borehole measurements), the soil column is assumed to be fully saturated with water that exists in a single phase (either liquid or ice).
- Chapter IV
    - GCMs selected from CMIP6 are assumed to be independent of each other in order to apply the stochastic downscaling method. However, GCM #2 and # 5 are from the same institute and the same land model is used, which may violate this assumption.
    - Natural variability of the observation data (borehole temperature measurements) is assumed to be fixed in the downscaling method due to the small amount of available data points.
    - The numerical model parameters used for long-term permafrost thermal state projection is calibrated and inferred from observations for only a few years, which may not be representative of the actual historical soil properties.

  - Historical soil temperature at shallow depths is derived from ERA5-Land air temperature with a simple linear regression. The actual relationship between air and ground temperature may be highly nonlinear.
  - Impacts of surface conditions such as a change in vegetation cover, snow distribution, and increase of precipitation in the future on the subsurface thermal processes are excluded in the projection simulation runs.

5.3 Future work: Impacts of vegetation change on permafrost thermal regime

In Chapter II, and Chapter III-IV, the vegetation dynamics and permafrost physical processes are analyzed separately and independently, and the crucial interactions between them like the impacts of increased vegetation on permafrost thermal regime are not included. Such impacts are still debated in the literature. Specifically, on one hand, increased vegetation growth in the circumpolar region can cause warmer winter ground temperature and deeper permafrost thaw depth due to insulation effects of increased snowpack retained by taller vegetation (*Sturm et al.*, 2001; *Myers-Smith and Hik*, 2013; *Frost et al.*, 2018; *Wilcox et al.*, 2019). On the other hand, the ground surface temperature may be reduced by the cooling effects of increased vegetation cover (*Blok et al.*, 2010; *Myers-Smith and Hik*, 2013; *Frost et al.*, 2018). The combined impacts of these positive and negative feedbacks need to be considered when simulating the subsurface thermal process. *Way and Lapalme* (2021) have studied these impacts with an analytical model applied for below-ground temperature simulation.

The uncertainty-informed framework developed in this thesis could provide a new view in simulating the future subsurface thermal condition under the complex impacts of increased vegetation. For this purpose, a more comprehensive numerical model needs to be built in the future by coupling the snow and vegetation dynamics into the

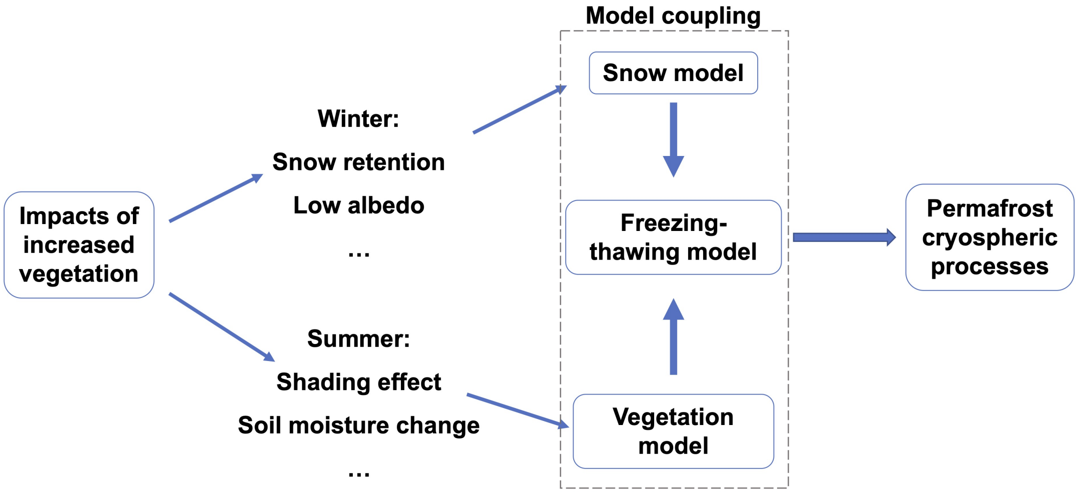

Figure 5.1: Dynamics of increased vegetation on permafrost and model coupling scheme.

current subsurface freeze-thaw model. Available models that represent the vegetation and snow dynamics are provided in literature (e.g., *Bartelt and Lehning*, 2002). A possible scheme is presented in Fig. 5.1.

APPENDICES

APPENDIX A

One Dimensional Freeze-Thaw Numerical Model Validation

The numerical model which solves soil thermal and moisture state (described in Sec. 3.3) is validated using laboratory experiment results from *Jame and Norum* (1980). The experiment setup is described briefly here. Several horizontal soil columns with initially uniform temperature and unsaturated moisture profiles were frozen at one end of the columns, while at the other end, the temperature was kept equal to the initial temperature. Water flux was inhibited at both ends, i.e, the total water content in the column was unchanged during the test. Temperature and total water content (liquid water and ice) were then recorded at some time intervals at several points horizontally distributed along the column.

Specifically, data from test 2 in Jame and Norum's work is used in this study. In this test, a lucite tube 30 cm in length, and 10 cm in diameter was filled with a #40 silica flour, of which 72% passed a 325 sieve (0.044-mm openings). The dry density of the material is around 1.33 g cm^{-3}. The initial temperature profile was uniformly set to 4.2 °C, and the temperature at the cold end to freeze the soil column is -6 °C, leading to an overall 1/3 °C/cm temperature gradient. The initial total water content is 0.2. The experiment was carried out for 72 hours, with measurements taken at the 6, 12, 24, 48, and 72 hours.

The original formulation for the freeze-thaw model in *Sheshukov and Egorov* (2002)'s work considering moisture movement in unsaturated conditions was used here. Compared to the formulation assuming fully saturated condition and no moisture movement as described in Sec. 3.3, the conservation equation of moisture is described as:

$$\phi\frac{\partial}{\partial t}(\rho_w S_w + \rho_i(1 - S_w)) + \rho_w\frac{\partial q_w}{\partial x} = 0, \tag{A.1}$$

Most parameter names are listed in Table 3.1, while q_w is the liquid water flux (m s^{-1}). According to Darcy's law including gravitational potential head,

$$q_w = -K\frac{\partial \psi_w}{\partial x} + K, \tag{A.2}$$

where K is the hydraulic conductivity (m s^{-1}) given by:

$$K = K_{sat}S_w^{2b+3}10^{-E\phi(1-S_w)}, \tag{A.3}$$

with K_{sat}, the saturated hydraulic conductivity (m s^{-1}), and E the impedance factor introduced by *Taylor and Luthin* (1978). The impedance effect of liquid flow due to ice blockage is neglected here by setting $E = 0$. Since the soil column in the test case used here is set horizontally, the gravitational potential head will disappear in Eq. (A.2) leads to:

$$q_w = -K\frac{\partial \psi_w}{\partial x}, \tag{A.4}$$

The temperature and moisture profiles in the soil column during the freeze-thaw process can then be solved by combining the conservation equations described here and those in Sec. 3.3. Parameter values and units used in the validation run are listed in Table A.1.

Table A.1: Model parameters and their corresponding values used for the *Jame and Norum* (1980) experiment test*

Parameter	Name	Values	Unit
ϕ	porosity	0.5	-
b	reciprocal of pore size distribution index in the B-C model	2	-
ψ_s	air entry pressure head in the B-C model	-2	m
K_{sat}	saturated hydraulic conductivity	1.5×10^{-6}	m s^{-1}
c_h	volumetric heat capacity of soil constituents excluding water and ice	1.12×10^{6}	J m^{-3} K^{-1}
k_s	soil solid thermal conductivity	1.5	W m^{-1} K^{-1}

*Other constant parameters values are listed in Table 3.1.

Fig. A.1 compares the simulated and measured soil temperature and water content for the test 2 case at 72 hours. The overall good match of results validates the feasibility of the numerical model used in this study to simulate the freeze-thaw process in the soil column.

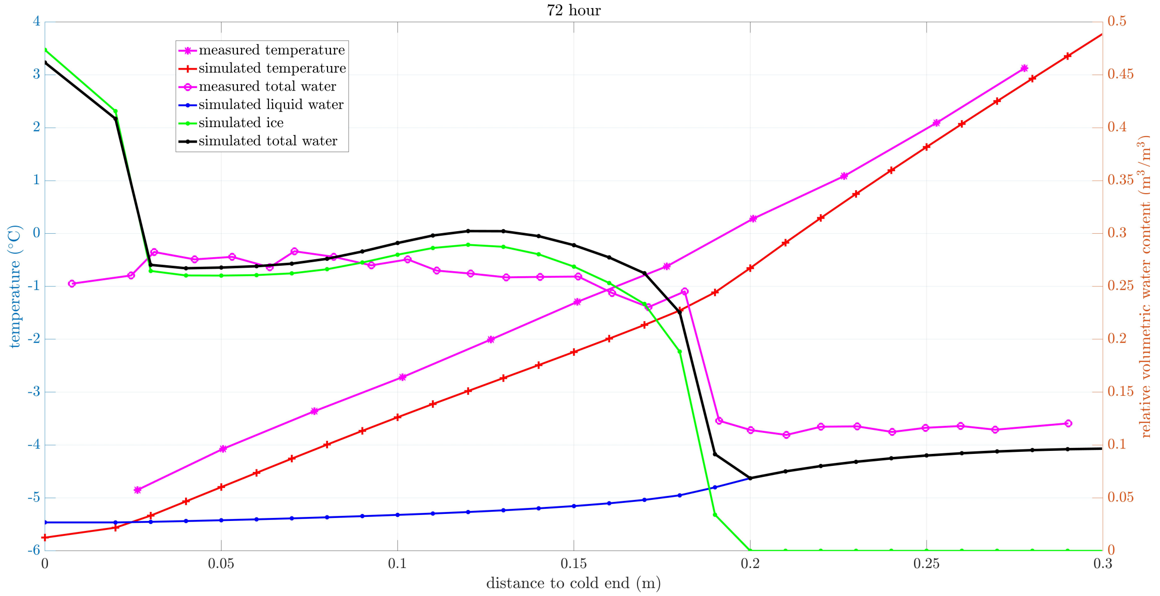

Figure A.1: Comparison between the measured temperature(magenta line with circle marker) and total water content (magenta line with asteroid marker) and simulated temperature (red line with cross marker) and total water content (black dotted line) at the 72 hours for the test 2 case in *Jame and Norum* (1980) experiment. The simulated liquid water (blue dotted line) and ice (green dotted line) content are also plotted for reference.

APPENDIX B

Selection of Eigenvalues Number J in the KL Expansion

For a set of given QoIs output from the model $\mathcal{M}$ of N data points, a set of eigenvalues μ_j, $j = 1, 2, \ldots, N$ and corresponding eigenfunctions can be derived after applying the KL expansion (Eq. (3.22)). The observational data $\boldsymbol{Y}$ on QoIs can be projected to the defined eigenspace using Eq. (3.23) and generating N KL coefficients η_j, $j = 1, 2, \ldots, N$. If using the largest J eigenvalues, and their corresponding eigenfunctions and KL coefficients, $\boldsymbol{Y}$ can be approximately estimated as:

$$\boldsymbol{Y}(\boldsymbol{x}, t) \approx \bar{\mathcal{M}}(\boldsymbol{x}, t) + \sum_{j=1}^{J} \eta_j \sqrt{\mu_j} \varphi_j(\boldsymbol{x}, t), \tag{B.1}$$

The larger J is, the more eigenvalues, eigenfunctions, and KL coefficients will be used, and the better the estimation will be. In the extreme scenario, when $J = N$, the KL expansion (right-hand-side of Eq. (B.1), denoted as RHS) will be exactly equal to the observation (left-hand-side of Eq. (B.1), denoted as LHS). The relationships between J and the absolute bias ($\frac{\sum abs(\bar{\mathcal{M}}(\boldsymbol{x},t)+\sum_{j=1}^{J} \eta_j \sqrt{\mu_j} \varphi_j(\boldsymbol{x},t)-\boldsymbol{Y}(\boldsymbol{x},t))}{N}$) and the root mean square error RMSE ($\sqrt{\frac{\sum(\bar{\mathcal{M}}(\boldsymbol{x},t)+\sum_{j=1}^{J} \eta_j \sqrt{\mu_j} \varphi_j(\boldsymbol{x},t)-\boldsymbol{Y}(\boldsymbol{x},t))^2}{N}}$) are plotted for the 3 kinds of model inputs (described in Sec. 3.6.2-3.6.4) in Fig. B.1(a)-(c) and numerical model corroboration (described in Sec. 3.6.4) in Fig. B.1(d) at log scale. The inset plots in Fig. B.1(a)-(d) show the relationship between the explained variance ($\sum_{j=1}^{J} \mu_j / \sum_{j=1}^{\infty} \mu_j$) and number of eigenvalues J up to 10 (20 in (d)) since the explained variance reaches 100% very quickly. The absolute bias and RMSE decrease from $J = 1$ to $J = 2$ or 3. Then they remain unchanged (Fig. B.1(a)-(b)), or gradually decrease (Fig. B.1(c)-(d)) until J is close to N and abruptly decrease to 0 at $J = N$. Fig. B.2(a)-(d) presents the comparison between the original observation (LHS in Eq. (B.1)), black solid lines) and the KL expansion estimation (RHS in Eq. (B.1)) with $J = 2$ or 3 (blue solid lines) and $J = N$ (red dashed lines) for 3 types of model inputs in Sec. 3.6 and numerical model corroboration in Sec. 3.6.4.

To keep the high accuracy of the KL estimation to the original QoI observation (i.e., small absolute bias and RMSE), one may want to select a J close to N. However, the purpose of applying the KL expansion in Chapter III is to reduce the size of input for the UQ machinery from N (number of data points in $\boldsymbol{Y}$) to J (the number of eigenvalues) and hence the improvement of computational efficiency. The selection of J based on the criteria in Sec. 3.4.3 is a trade-off between accuracy and efficiency, yet still reasonable because the accuracy doesn't increase much as J increases until close to N. The accuracy of the KL estimation also depends on the eigenvalues and eigenfunctions derived from the set of QoI simulations using model $\mathcal{M}$. Thus, another way to increase the KL estimation accuracy is the amelioration of model $\mathcal{M}$. Conversely, if the performance of the model $\mathcal{M}$ is poor, the developed process presented in this study may not work well.

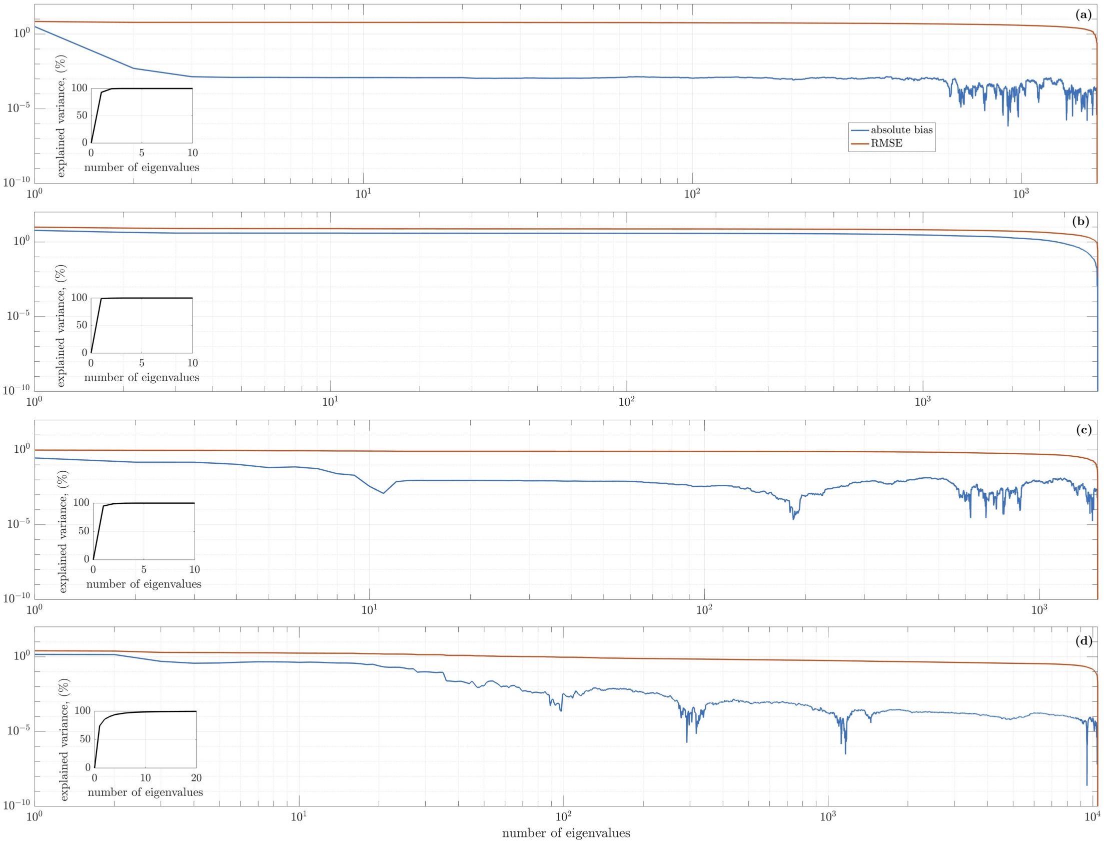

Figure B.1: Relationship between number of eigenvalues J and absolute bias (blue solid lines) and RMSE (red solid lines) for 3 types of model inputs: (a) field measurements, (b) GCM data product, (c) borehole measurements in Sec. 3.6, and (d) numerical model corroboration (presented in Fig. 3.12). Inset plots show the explained variance vs. number of eigenvalues (black solid lines).

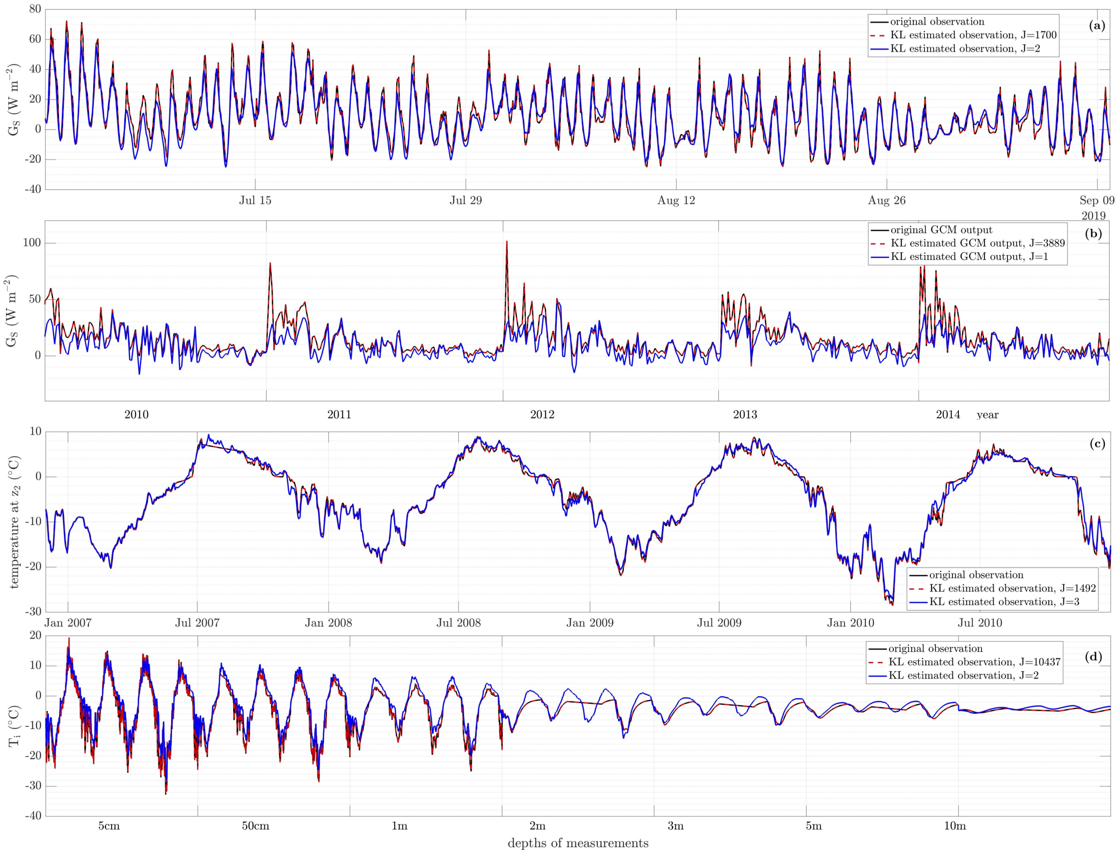

Figure B.**2:** Comparison between the original observation (black solid lines) and the KL expansion estimation with $J = 2$ or 3 (blue solid lines) and $J = N$ (red dashed lines) for 3 types of model inputs: (a) field measurements, (b) GCM data product, and (c) borehole measurements in Sec. 3.6, and (d) numerical model corroboration (Fig. 3.12)

APPENDIX C

CMIP6 GCM Grid Cells Used for Data Retrieving

The following table provides information for grid cells used for data retrieving. The first column is model names, the last column is the GCM resolution in the format M×N, where M is the number of grid cells along longitudinal directions, and N along latitudinal grid cells. The numbers in the second to fourth columns indicate the exact GCM grid cell for each borehole. Specifically, m is the longitudinal number counted from 0°towards the east, and n is the latitudinal number counted from 90°S towards the north. The percentage number in the bracket represents the portion of land coverage of the grid cell (100% if it is not shown). Fig. C.1 presents an example of output grid cells from CESM2 for sensible heat flux *H* and the counting directions for m and n.

Table C.1: GCM grid cells used for data retrieving

	MS	VD	BV	Resolutions
CESM2	m: 55, n: 171 (90.14%)	m: 56, n: 171	m: 45, n: 169 (96.02%)	288×192
HadGEM3-GC31-LL	m: 37, n: 128	m: 37, n: 129 (94.71%)	m: 29, n: 127 (93.91%)	192×144
IPSL-CM6A-LR	m: 29, n: 127	m: 29, n: 127	m: 23, n: 125	144×143
MPI-ESM1-2-LR	m: 37, n: 86	m: 38, n: 86	m: 30, n: 85	192×96
UKESM1-0-LL	m: 37, n: 128	m: 37, n: 129 (94.71%)	m: 29, n: 127 (93.91%)	192×44
ACCESS-ESM1-5	m: 38, n: 129	m: 38, n: 129	m: 30, n: 127	192×145
EC-Earth3	m: 97, n: 228	m: 99, n: 229	m: 79, n: 225	512×256
MIROC-ES2L	m: 26, n: 58	m: 26, n: 58	m: 20, n: 56	128×64
NorESM2-LM	m: 29, n: 85 (99.28%)	m: 29, n: 86	m: 23, n: 84	144×96

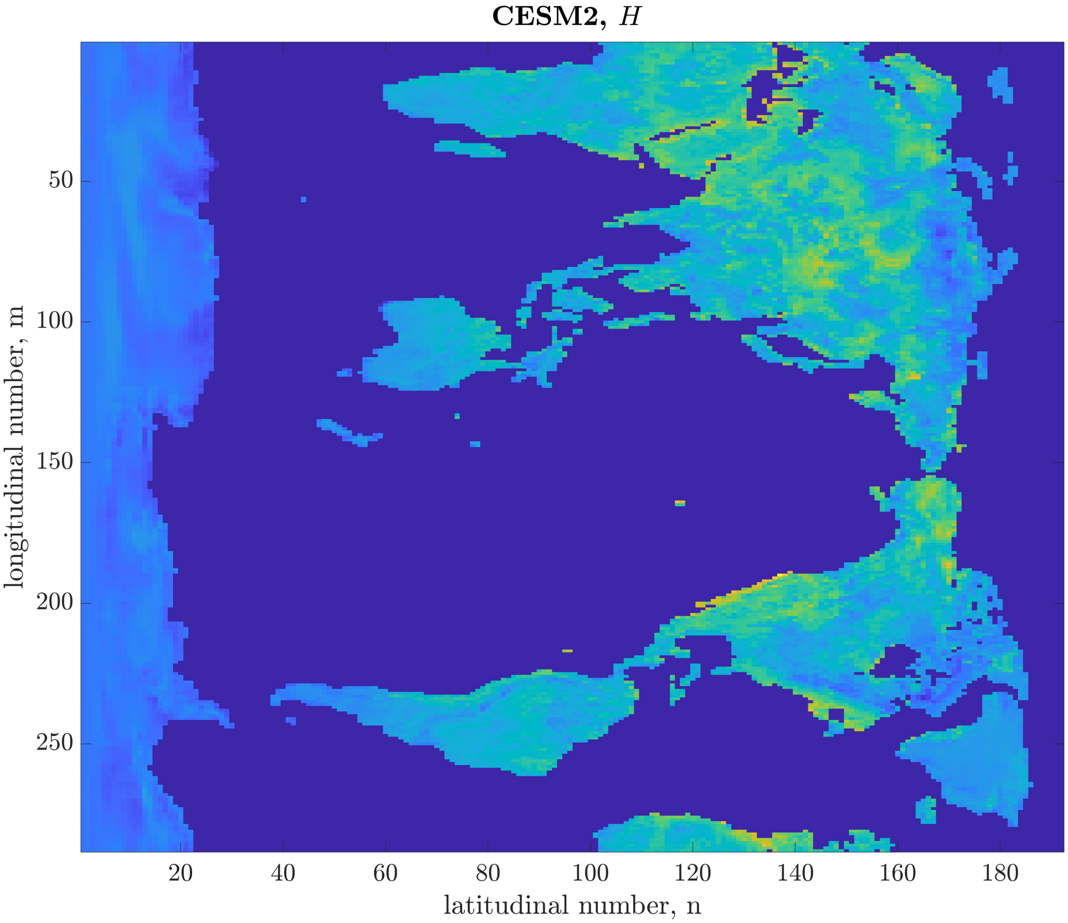

Figure C.1: An example of sensible heat flux (H) output grid cells from CESM2.

BIBLIOGRAPHY

BIBLIOGRAPHY

Alexandrov, G. A., V. A. Ginzburg, G. E. Insarov, and A. A. Romanovskaya (2021), Cmip6 model projections leave no room for permafrost to persist in western siberia under the ssp5-8.5 scenario, *Climatic Change*, *169*(3), 1–11.

Alizamir, M., O. Kisi, A. N. Ahmed, C. Mert, C. M. Fai, S. Kim, N. W. Kim, and A. El-Shafie (2020), Advanced machine learning model for better prediction accuracy of soil temperature at different depths, *PLoS One*, *15*(4), e0231,055.

Babacan, S. D., R. Molina, and A. K. Katsaggelos (2009), Bayesian compressive sensing using laplace priors, *IEEE Transactions on image processing*, *19*(1), 53–63.

Bartelt, P., and M. Lehning (2002), A physical snowpack model for the swiss avalanche warning: Part i: numerical model, *Cold Regions Science and Technology*, *35*(3), 123–145.

Beck, P. S., and S. J. Goetz (2011), Satellite observations of high northern latitude vegetation productivity changes between 1982 and 2008: ecological variability and regional differences, *Environmental Research Letters*, *6*(4), 045,501, doi:10.1088/1748-9326/6/4/045501.

Berner, L. T., et al. (2020), Summer warming explains widespread but not uniform greening in the Arctic tundra biome, *Nature Communications*, *11*(1), 1–12, doi:10.1038/s41467-020-18479-5.

Berveiller, M., B. Sudret, and M. Lemaire (2006), Stochastic finite element: a non intrusive approach by regression, *European Journal of Computational Mechanics/Revue Européenne de Mécanique Numérique*, *15*(1-3), 81–92.

Bieniek, P. A., U. S. Bhatt, J. E. Walsh, R. Lader, B. Griffith, J. K. Roach, and R. L. Thoman (2018), Assessment of alaska rain-on-snow events using dynamical downscaling, *Journal of Applied Meteorology and Climatology*, *57*(8), 1847–1863.

Biskaborn, B. K., J.-P. Lanckman, H. Lantuit, K. Elger, D. Streletskiy, W. Cable, and V. E. Romanovsky (2015), The new database of the Global Terrestrial Network for Permafrost (GTN-P), *Earth System Science Data*, *7*(2), 245–259, doi:10.5194/essd-7-245-2015.

Biskaborn, B. K., et al. (2019), Permafrost is warming at a global scale, *Nature Communications*, *10*(1), 1–11, doi:10.1038/s41467-018-08240-4.

Bjorkman, A. D., et al. (2018), Plant functional trait change across a warming tundra biome, *Nature*, *562*(7725), 57–62, doi:10.1038/s41586-018-0563-7.

Blatman, G., and B. Sudret (2008), Sparse polynomial chaos expansions and adaptive stochastic finite elements using a regression approach, *Comptes Rendus Mécanique*, *336*(6), 518–523.

Blatman, G., and B. Sudret (2011), Adaptive sparse polynomial chaos expansion based on least angle regression, *Journal of computational Physics*, *230*(6), 2345–2367.

Blok, D., M. M. Heijmans, G. SCHAEPMAN-STRUB, A. Kononov, T. Maximov, and F. Berendse (2010), Shrub expansion may reduce summer permafrost thaw in siberian tundra, *Global Change Biology*, *16*(4), 1296–1305.

Bond-Lamberty, B., C. Wang, and S. Gower (2002), Aboveground and belowground biomass and sapwood area allometric equations for six boreal tree species of northern Manitoba, *Canadian Journal of Forest Research*, *32*(8), 1441–1450, doi:10.1139/x02-063.

Boucher, O., et al. (2020), Presentation and evaluation of the ipsl-cm6a-lr climate model, *Journal of Advances in Modeling Earth Systems*, *12*(7), e2019MS002,010.

Briffa, K. R., V. V. Shishov, T. M. Melvin, E. A. Vaganov, H. Grudd, R. M. Hantemirov, M. Eronen, and M. M. Naurzbaev (2008), Trends in recent temperature and radial tree growth spanning 2000 years across northwest Eurasia, *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1501), 2269–2282, doi:10.1098/rstb.2007.2199.

Briffa, K. R., T. M. Melvin, T. J. Osborn, R. M. Hantemirov, A. V. Kirdyanov, V. S. Mazepa, S. G. Shiyatov, and J. Esper (2013), Reassessing the evidence for tree-growth and inferred temperature change during the Common Era in Yamalia, northwest Siberia, *Quaternary Science Reviews*, *72*, 83–107, doi:10.1016/j.quascirev.2013.04.008.

Brooks, R. H., and A. T. Corey (1966), Properties of porous media affecting fluid flow, *Journal of the irrigation and drainage division*, *92*(2), 61–88, doi:10.1061/JRCEA4.0000425.

Brown, J., K. M. Hinkel, and F. Nelson (2000), The circumpolar active layer monitoring (CALM) program: research designs and initial results, *Polar Geography*, *24*(3), 166–258, doi:10.1080/10889370009377698.

Brown, J., O. Ferrians, J. A. Heginbottom, and E. Melnikov (2002), Circum-Arctic Map of Permafrost and Ground-Ice Conditions, Version 2, doi:10.7265/skbg-kf16, accessed on 21 February 2021.

Burke, E. J., Y. Zhang, and G. Krinner (2020), Evaluating permafrost physics in the coupled model intercomparison project 6 (cmip6) models and their sensitivity to climate change, *The Cryosphere*, *14*(9), 3155–3174.

Campbell Scientific (2016), Model HFP01 soil heat flux plate instruction manual., Revision 10/16. Campbell Scientific, Inc.

Campbell Scientific (2018), CS650 and CS655 Water Content Reflectometers., Revision 04/18. Campbell Scientific, Inc.

Chapin III, F. S., et al. (2005), Role of land-surface changes in Arctic summer warming, *Science*, *310*(5748), 657–660, doi:10.1126/science.1117368.

Chen, J., F. P. Brissette, and R. Leconte (2011), Uncertainty of downscaling method in quantifying the impact of climate change on hydrology, *Journal of Hydrology*, *401*(3-4), 190–202.

Cobos, D. R., and J. M. Baker (2003), In situ measurement of soil heat flux with the gradient method, *Vadose Zone Journal*, *2*(4), 589–594, doi:10.2136/vzj2003.5890.

Condit, R., et al. (2000), Spatial patterns in the distribution of tropical tree species, *Science*, *288*(5470), 1414–1418, doi:10.1126/science.288.5470.1414.

Constable, A. J., et al. (2022), Cross-chapter paper 6: Polar regions, in *IPCC AR WGII*, Cambridge University Press.

Danabasoglu, G., et al. (2020), The community earth system model version 2 (cesm2), *Journal of Advances in Modeling Earth Systems*, *12*(2), e2019MS001,916.

Davies, J. H. (2013), Global map of solid Earth surface heat flow, *Geochemistry, Geophysics, Geosystems*, *14*(10), 4608–4622, doi:10.1002/ggge.20271.

Dearborn, K. D., and R. K. Danby (2018), Topographic influences on ring widths of trees and shrubs across alpine treelines in southwest Yukon, *Arctic, Antarctic, and Alpine Research*, *50*(1), e1495,445, doi:10.1080/15230430.2018.1495445.

Demchev, D., M. Y. Kulakov, A. Makshtas, I. Makhotina, K. Fil'Chuk, and I. Frolov (2020), Verification of era-interim and era5 reanalyses data on surface air temperature in the arctic, *Russian Meteorology and Hydrology*, *45*(11), 771–777.

Devi, N., F. Hagedorn, P. Moiseev, H. Bugmann, S. Shiyatov, V. Mazepa, and A. Rigling (2008), Expanding forests and changing growth forms of Siberian larch at the Polar Urals treeline during the 20th century, *Global Change Biology*, *14*(7), 1581–1591, doi:10.1111/j.1365-2486.2008.01583.x.

Devi, N. M., V. V. Kukarskih, A. A. Galimova, V. S. Mazepa, and A. A. Grigoriev (2020), Climate change evidence in tree growth and stand productivity at the upper treeline ecotone in the Polar Ural Mountains, *Forest Ecosystems*, *7*(1), 1–16, doi:10.1186/s40663-020-0216-9.

Döscher, R., et al. (2022), The ec-earth3 earth system model for the coupled model intercomparison project 6, *Geoscientific Model Development*, *15*(7), 2973–3020.

Dwelle, M. C., J. Kim, K. Sargsyan, and V. Y. Ivanov (2019), Streamflow, stomata, and soil pits: Sources of inference for complex models with fast, robust uncertainty quantification, *Advances in Water Resources*, *125*, 13–31, doi:10.1016/j.advwatres.2019.01.002.

El Sharif, H., W. Zhou, V. Ivanov, A. Sheshukov, V. Mazepa, and J. Wang (2019), Surface energy budgets of arctic tundra during growing season, *Journal of Geophysical Research: Atmospheres*, *124*(13), 6999–7017, doi:10.1029/2019JD030650.

Ellison, A. M. (2004), Bayesian inference in ecology, *Ecology Letters*, *7*(6), 509–520, doi:10.1111/j.1461-0248.2004.00603.x.

Elmendorf, S. C., et al. (2012), Plot-scale evidence of tundra vegetation change and links to recent summer warming, *Nature Climate Change*, *2*(6), 453–457, doi:10.1038/nclimate1465.

Environmental Systems Research Institute (2011), ArcGIS Desktop: Release 10, Redlands, CA.

Esper, J., and F. H. Schweingruber (2004), Large-scale treeline changes recorded in Siberia, *Geophysical Research Letters*, *31*(6), doi:10.1029/2003GL019178.

ESRI (2021), Curvature Function, available at: `https://desktop.arcgis.com/en/arcmap/10.7/manage-data/raster-and-images/curvature-function.htm`, accessed on 26 May 2021.

Ester, M., H.-P. Kriegel, J. Sander, X. Xu, et al. (1996), A density-based algorithm for discovering clusters in large spatial databases with noise., in *kdd*, vol. 96, pp. 226–231.

Eyring, V., S. Bony, G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer, and K. E. Taylor (2016), Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, *Geoscientific Model Development*, *9*(5), 1937–1958, doi:10.5194/gmd-9-1937-2016.

Farouki, O. T. (1981), Thermal properties of soils, *Tech. rep.*, Cold Regions Research and Engineering Lab Hanover NH.

Fatichi, S., V. Y. Ivanov, and E. Caporali (2011), Simulation of future climate scenarios with a weather generator, *Advances in Water Resources*, *34*(4), 448–467.

Feng, Y., N. Cui, W. Hao, L. Gao, and D. Gong (2019), Estimation of soil temperature from meteorological data using different machine learning models, *Geoderma*, *338*, 67–77.

Foley, J. A., M. H. Costa, C. Delire, N. Ramankutty, and P. Snyder (2003), Green surprise? How terrestrial ecosystems could affect earth's climate, *Frontiers in Ecology and the Environment*, *1*(1), 38–44, doi:10.1890/1540-9295(2003)001[0038:GSHTEC]2.0.CO;2.

Forbes, B. C., M. M. Fauria, and P. Zetterberg (2010), Russian Arctic warming and 'greening' are closely tracked by tundra shrub willows, *Global Change Biology*, *16*(5), 1542–1554, doi:10.1111/j.1365-2486.2009.02047.x.

Ford, J. D., G. McDowell, and J. Jones (2014), The state of climate change adaptation in the arctic, *Environmental Research Letters*, *9*(10), 104,005.

Ford, J. D., G. McDowell, and T. Pearce (2015), The adaptation challenge in the arctic, *Nature Climate Change*, *5*(12), 1046–1053.

Frost, G. V., and H. E. Epstein (2014), Tall shrub and tree expansion in Siberian tundra ecotones since the 1960s, *Global Change Biology*, *20*(4), 1264–1277, doi:10.1111/gcb.12406.

Frost, G. V., H. E. Epstein, and D. A. Walker (2014), Regional and landscape-scale variability of Landsat-observed vegetation dynamics in northwest Siberian tundra, *Environmental Research Letters*, *9*(2), 025,004, doi:10.1088/1748-9326/9/2/025004.

Frost, G. V., H. E. Epstein, D. A. Walker, G. Matyshak, and K. Ermokhina (2018), Seasonal and long-term changes to active-layer temperatures after tall shrubland expansion and succession in arctic tundra, *Ecosystems*, *21*(3), 507–520.

Fuchs, M., and C. Tanner (1968a), Calibration and field test of soil heat flux plates, *Soil Science Society of America Journal*, *32*(3), 326–328, doi:10.2136/sssaj1968.03615995003200030021x.

Fuchs, M., and C. Tanner (1968b), Surface temperature measurements of bare soils, *Journal of Applied Meteorology (1962-1982)*, *7*(2), 303–305.

Gao, Z., R. Horton, and H. Liu (2010), Impact of wave phase difference between soil surface heat flux and soil surface temperature on soil surface energy balance closure, *Journal of Geophysical Research: Atmospheres*, *115*(D16), doi:10.1029/2009JD013278.

Giorgi, F., and L. O. Mearns (2002), Calculation of average, uncertainty range, and reliability of regional climate changes from aogcm simulations via the "reliability ensemble averaging"(rea) method, *Journal of climate*, *15*(10), 1141–1158.

Graham, R. M., S. R. Hudson, and M. Maturilli (2019), Improved performance of era5 in arctic gateway relative to four global atmospheric reanalyses, *Geophysical Research Letters*, *46*(11), 6138–6147.

Grigorieva, A., and P. Moiseev (2018), Peculiarities and determinants of regeneration of Siberian Larch on the upper limit of its growth in the Urals, *Contemporary Problems of Ecology*, *11*(1), 13–25, doi:10.1134/S1995425518010031.

Hagedorn, F., M. A. Dawes, M. O. Bubnov, N. M. Devi, A. A. Grigoriev, V. S. Mazepa, Z. Y. Nagimov, S. G. Shiyatov, and P. A. Moiseev (2020), Latitudinal decline in stand biomass and productivity at the elevational treeline in the Ural mountains despite a common thermal growth limit, *Journal of Biogeography*, *47*(8), 1827–1842, doi:10.1111/jbi.13867.

Hagedorn, F., et al. (2014), Treeline advances along the Urals mountain range–driven by improved winter conditions?, *Global Change Biology*, *20*(11), 3530–3543, doi:10.1111/gcb.12613.

Hajima, T., et al. (2020), Development of the miroc-es2l earth system model and the evaluation of biogeochemical processes and feedbacks, *Geoscientific Model Development*, *13*(5), 2197–2244.

Halliwell, D. H., and W. R. Rouse (1987), Soil heat flux in permafrost: characteristics and accuracy of measurement, *Journal of Climatology*, *7*(6), 571–584, doi:10.1002/joc.3370070605.

Harlan, R. (1973), Analysis of coupled heat-fluid transport in partially frozen soil, *Water Resources Research*, *9*(5), 1314–1323, doi:10.1029/WR009i005p01314.

Harris, C., W. Haeberli, D. Vonder Mühll, and L. King (2001), Permafrost monitoring in the high mountains of europe: the pace project in its global context, *Permafrost and periglacial processes*, *12*(1), 3–11.

Harris, C., et al. (2009), Permafrost and climate in europe: Monitoring and modelling thermal, geomorphological and geotechnical responses, *Earth-Science Reviews*, *92*(3-4), 117–171.

Harsch, M. A., and M. Y. Bader (2011), Treeline form–a potential key to understanding treeline dynamics, *Global Ecology and Biogeography*, *20*(4), 582–596, doi:10.1111/j.1466-8238.2010.00622.x.

Harsch, M. A., P. E. Hulme, M. S. McGlone, and R. P. Duncan (2009), Are treelines advancing? A global meta-analysis of treeline response to climate warming, *Ecology Letters*, *12*(10), 1040–1049, doi:10.1111/j.1461-0248.2009.01355.x.

Hastings, W. K. (1970), Monte Carlo sampling methods using Markov chains and their applications, doi:10.1093/biomet/57.1.97.

Hengl, T., et al. (2017), SoilGrids250m: Global gridded soil information based on machine learning, *PLoS One*, *12*(2), e0169,748, doi:10.1371/journal.pone.0169748.

Hersbach, H., et al. (2020), The ERA5 global reanalysis, *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049, doi:10.1002/qj.3803.

Heusinkveld, B. G., A. Jacobs, A. Holtslag, and S. Berkowicz (2004), Surface energy balance closure in an arid region: role of soil heat flux, *Agricultural and Forest Meteorology*, *122*(1-2), 21–37, doi:10.1016/j.agrformet.2003.09.005.

Hiemstra, C. A., G. E. Liston, and W. A. Reiners (2002), Snow redistribution by wind and interactions with vegetation at upper treeline in the Medicine Bow Mountains, Wyoming, USA, *Arctic, Antarctic, and Alpine Research*, *34*(3), 262–273, doi:10.1080/15230430.2002.12003493.

Hinkel, K., F. Paetzold, F. Nelson, and J. Bockheim (2001), Patterns of soil temperature and moisture in the active layer and upper permafrost at Barrow, Alaska: 1993–1999, *Global and Planetary Change*, *29*(3-4), 293–309, doi:10.1016/S0921-8181(01)00096-0.

Hinzman, L., D. Kane, R. Gieck, and K. Everett (1991), Hydrologic and thermal properties of the active layer in the Alaskan Arctic, *Cold Regions Science and Technology*, *19*(2), 95–110, doi:10.1016/0165-232X(91)90001-W.

Hinzman, L. D., et al. (2005), Evidence and implications of recent climate change in northern Alaska and other arctic regions, *Climatic Change*, *72*(3), 251–298, doi:10.1007/s10584-005-5352-2.

Hjort, J., O. Karjalainen, J. Aalto, S. Westermann, V. E. Romanovsky, F. E. Nelson, B. Etzelmüller, and M. Luoto (2018), Degrading permafrost puts arctic infrastructure at risk by mid-century, *Nature communications*, *9*(1), 1–9.

Hjort, J., D. Streletskiy, G. Doré, Q. Wu, K. Bjella, and M. Luoto (2022), Impacts of permafrost degradation on infrastructure, *Nature Reviews Earth & Environment*, *3*(1), 24–38, doi:10.1038/s43017-021-00247-8.

Hoch, G., and C. Körner (2009), Growth and carbon relations of tree line forming conifers at constant vs. variable low temperatures, *Journal of Ecology*, *97*(1), 57–66, doi:10.1111/j.1365-2745.2008.01447.x.

Holtmeier, F.-K., and G. Broll (2005), Sensitivity and response of northern hemisphere altitudinal and polar treelines to environmental change at landscape and local scales, *Global Ecology and Biogeography*, *14*(5), 395–410, doi:10.1111/j.1466-822X.2005.00168.x.

Holtmeier, F. K., and G. E. Broll (2007), Treeline advance-driving processes and adverse factors, *Landscape*, *1*, 1–33, doi:10.3097/LO.200701.

Horton, R., and P. Wierenga (1983), Estimating the soil heat flux from observations of soil temperature near the surface, *Soil Science Society of America Journal*, *47*(1), 14–20, doi:10.2136/sssaj1983.03615995004700010003x.

Hou, Z., and Y. Rubin (2005), On minimum relative entropy concepts and prior compatibility issues in vadose zone inverse and forward modeling, *Water Resources Research*, *41*(12).

ITIS (2021), Integrated Taxonomic Information System, CC0, doi:10.5066/F7KH0KBK, available at: `www.itis.gov`, accessed on 30 September 2021.

Ivanov, V. Y., et al. (2021), Breaking down the computational barriers to real-time urban flood forecasting, *Geophysical Research Letters*, *48*(20), e2021GL093,585.

Jame, Y.-W., and D. I. Norum (1980), Heat and mass transfer in a freezing unsaturated porous medium, *Water Resources Research*, *16*(4), 811–819.

Jiang, Q., W. Li, Z. Fan, X. He, W. Sun, S. Chen, J. Wen, J. Gao, and J. Wang (2021), Evaluation of the era5 reanalysis precipitation dataset over chinese mainland, *Journal of hydrology*, *595*, 125,660.

Jin, X.-Y., H.-J. Jin, G. Iwahana, S. S. Marchenko, D.-L. Luo, X.-Y. Li, and S.-H. Liang (2021), Impacts of climate-induced permafrost degradation on vegetation: A review, *Advances in Climate Change Research*, *12*(1), 29–47.

Johansen, O. (1977), Thermal conductivity of soils, *Tech. rep.*, Cold Regions Research and Engineering Lab Hanover NH.

Jones, S. B., J. M. Wraith, and D. Or (2002), Time domain reflectometry measurement principles and applications, *Hydrological Processes*, *16*(1), 141–153.

Jorgenson, M. T., C. H. Racine, J. C. Walters, and T. E. Osterkamp (2001), Permafrost degradation and ecological changes associated with a warmingclimate in central alaska, *Climatic change*, *48*(4), 551–579.

Jorgenson, M. T., et al. (2013), Reorganization of vegetation, hydrology and soil carbon after permafrost degradation across heterogeneous boreal landscapes, *Environmental Research Letters*, *8*(3), 035,017.

Kajimoto, T., Y. Matsuura, A. Osawa, A. P. Abaimov, O. A. Zyryanova, A. P. Isaev, D. P. Yefremov, S. Mori, and T. Koike (2006), Size–mass allometry and biomass allocation of two larch species growing on the continuous permafrost region in Siberia, *Forest Ecology and Management*, *222*(1-3), 314–325, doi:10.1016/j.foreco.2005.10.031.

Kane, D., L. Hinzman, C. Benson, and K. Everett (1989), Hydrology of Imnavait Creek, an arctic watershed, *Ecography*, *12*(3), 262–269, doi:10.1111/j.1600-0587.1989.tb00845.x.

Kang, S., S. Kim, S. Oh, and D. Lee (2000), Predicting spatial and temporal patterns of soil temperature based on topography, surface cover and air temperature, *Forest Ecology and Management*, *136*(1-3), 173–184.

Kaplan, J. O., and M. New (2006), Arctic climate change with a 2°C global warming: Timing, climate patterns and vegetation change, *Climatic Change*, *79*(3), 213–241, doi:10.1007/s10584-006-9113-7.

Karhunen, K. (1947), *Ueber lineare Methoden in der Wahrscheinlichkeitsrechnung*, Soumalainen Tiedeakatemia.

Kataeva, M. N., N. V. Alexeeva-Popova, I. V. Drozdova, and A. I. Beljaeva (2004), Chemical composition of soils and plant species in the Polar Urals as influenced by rock type, *Geoderma*, *122*(2-4), 257–268, doi:10.1016/j.geoderma.2004.01.012.

Kaufmann, M. R., and C. A. Troendle (1981), The relationship of leaf area and foliage biomass to sapwood conducting area in four subalpine forest tree species, *Forest Science*, *27*(3), 477–482, doi:10.1093/forestscience/27.3.477.

Kharuk, V., K. Ranson, S. Im, and M. Naurzbaev (2006), Forest-tundra larch forests and climatic trends, *Russian Journal of Ecology*, *37*(5), 291–298, doi: 10.1134/S1067413606050018.

Kharuk, V. I., S. T. Im, and M. L. Dvinskaya (2010a), Forest–tundra ecotone response to climate change in the Western Sayan Mountains, Siberia, *Scandinavian Journal of Forest Research*, *25*(3), 224–233, doi:10.1080/02827581003766959.

Kharuk, V. I., K. J. Ranson, S. T. Im, and A. S. Vdovin (2010b), Spatial distribution and temporal dynamics of high-elevation forest stands in southern Siberia, *Global Ecology and Biogeography*, *19*(6), 822–830, doi:10.1111/j.1466-8238.2010.00555.x.

Kim, J., V. Y. Ivanov, and S. Fatichi (2016), Climate change and uncertainty assessment over a hydroclimatic transect of michigan, *Stochastic Environmental Research and Risk Assessment*, *30*(3), 923–944.

Kimball, B., R. Jackson, F. Nakayama, S. Idso, and R. Reginato (1976), Soil-heat flux determination: Temperature gradient method with computed thermal conductivities, *Soil Science Society of America Journal*, *40*(1), 25–28, doi: 10.2136/sssaj1976.03615995004000010011x.

Knutti, R., J. Sedláček, B. M. Sanderson, R. Lorenz, E. M. Fischer, and V. Eyring (2017), A climate model projection weighting scheme accounting for performance and interdependence, *Geophysical Research Letters*, *44*(4), 1909–1918.

Koenigk, T., L. Brodeau, R. G. Graversen, J. Karlsson, G. Svensson, M. Tjernström, U. Willén, and K. Wyser (2013), Arctic climate change in 21st century CMIP5 simulations with EC-Earth, *Climate Dynamics*, *40*(11), 2719–2743, doi:10.1007/s00382-012-1505-y.

Konstantinov, P., M. Zhelezniak, N. Basharin, I. Misailov, and V. Andreeva (2020), Establishment of permafrost thermal monitoring sites in east siberia, *Land*, *9*(12), 476.

Koopmans, R. W. R., and R. Miller (1966), Soil freezing and soil water characteristic curves, *Soil Science Society of America Journal*, *30*(6), 680–685.

Körner, C. (2003), *Alpine Plant Life: Functional Plant Ecology of High Mountain Ecosystems 2*, Berlin: Springer, `https://doi.org/10.1007/978-3-642-18970-8`.

Körner, C., and J. Paulsen (2004), A world-wide study of high altitude treeline temperatures, *Journal of Biogeography*, *31*(5), 713–732, doi:10.1111/j.1365-2699.2003.01043.x.

Krebs, P., M. Stocker, G. B. Pezzatti, and M. Conedera (2015), An alternative approach to transverse and profile terrain curvature, *International Journal of Geographical Information Science*, *29*(4), 643–666, doi:10.1080/13658816.2014.995102.

Kullman, L. (2007), Tree line population monitoring of Pinus sylvestris in the Swedish Scandes, 1973–2005: implications for tree line theory and climate change ecology, *Journal of Ecology*, *95*(1), 41–52, doi:10.1111/j.1365-2745.2006.01190.x.

Lafleur, P. M., and E. R. Humphreys (2018), Tundra shrub effects on growing season energy and carbon dioxide exchange, *Environmental Research Letters*, *13*(5), 055,001, doi:10.1088/1748-9326/aab863.

Lafleur, P. M., T. J. Griffis, and W. R. Rouse (2001), Interannual variability in net ecosystem CO2 exchange at the arctic treeline, *Arctic, Antarctic, and Alpine Research*, *33*(2), 149–157, doi:10.1080/15230430.2001.12003417.

Lawrence, D. M., and A. G. Slater (2005), A projection of severe near-surface permafrost degradation during the 21st century, *Geophysical Research Letters*, *32*(24).

Lawrence, D. M., and A. G. Slater (2008), Incorporating organic soil into a global climate model, *Climate Dynamics*, *30*(2), 145–160, doi:10.1007/s00382-007-0278-1.

Le Maître, O., and O. M. Knio (2010), *Spectral methods for uncertainty quantification: with applications to computational fluid dynamics*, Springer Science & Business Media.

Leibman, M. O., A. V. Khomutov, A. A. Gubarkov, Y. A. Dvornikov, and D. R. Mullanurov (2015), The research station" vaskiny dachi", central yamal, west siberia, russia–a review of 25 years of permafrost studies, *Fennia-International Journal of Geography*, *193*(1), 3–30.

Leuning, R., E. van Gorsel, W. J. Massman, and P. R. Isaac (2012), Reflections on the surface energy imbalance problem, *Agricultural and Forest Meteorology*, *156*, 65–74, doi:10.1016/j.agrformet.2011.12.002.

Liljedahl, A. K., et al. (2016), Pan-arctic ice-wedge degradation in warming permafrost and its influence on tundra hydrology, *Nature Geoscience*, *9*(4), 312–318.

Lloyd, A. H., and C. L. Fastie (2002), Spatial and temporal variability in the growth and climate response of treeline trees in Alaska, *Climatic Change*, *52*(4), 481–509, doi:10.1023/A:1014278819094.

Lloyd, A. H., T. S. Rupp, C. L. Fastie, and A. M. Starfield (2002), Patterns and dynamics of treeline advance on the Seward Peninsula, Alaska, *Journal of Geophysical Research: Atmospheres*, *107*(D2), ALT–2, doi:10.1029/2001JD000852.

Lloyd, A. H., A. G. Bunn, and L. Berner (2011), A latitudinal gradient in tree growth response to climate warming in the Siberian taiga, *Global Change Biology*, *17*(5), 1935–1945, doi:10.1111/j.1365-2486.2010.02360.x.

Loeve, M. (1948), Fonctions aleatoires du second ordre, supplement to p, *Lévy, Processus Stochastiques et Mouvement Brownien, Gauthier-Villars, Paris.*

Lunardini, V. (1981), *Heat Transfer in Cold Climates*, Van Nostrand Reinhold Company.

Lunardini, V. (1996), Climatic warming and the degradation of warm permafrost, *Permafrost and Periglacial Processes*, *7*(4), 311–320.

Luo, D., H. Jin, S. S. Marchenko, and V. E. Romanovsky (2018), Difference between near-surface air, land surface and ground surface temperatures and their influences on the frozen ground on the qinghai-tibet plateau, *Geoderma*, *312*, 74–85.

MacDonald, G., K. Kremenetski, and D. Beilman (2008), Climate change and the northern Russian treeline zone, *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1501), 2283–2299, doi:10.1098/rstb.2007.2200.

MacDougall, A. H., C. A. Avis, and A. J. Weaver (2012), Significant contribution to climate warming from the permafrost carbon feedback, *Nature Geoscience*, *5*(10), 719–721.

Macias-Fauria, M., B. C. Forbes, P. Zetterberg, and T. Kumpula (2012), Eurasian Arctic greening reveals teleconnections and the potential for structurally novel ecosystems, *Nature Climate Change*, *2*(8), 613–618, doi:10.1038/nclimate1558.

MacQueen, J. (1967), Some methods for classification and analysis of multivariate observations, in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability (Oakland, CA, USA)*, pp. 281–297.

Malinverno, A. (2002), Parsimonious Bayesian Markov chain Monte Carlo inversion in a nonlinear geophysical problem, *Geophysical Journal International*, *151*(3), 675–688, doi:10.1046/j.1365-246X.2002.01847.x.

Malkova, G., et al. (2022), Spatial and Temporal Variability of Permafrost in the Western Part of the Russian Arctic, *Energies*, *15*(7), 2311, doi:10.3390/en15072311.

Martin, A. C., E. S. Jeffers, G. Petrokofsky, I. Myers-Smith, and M. Macias-Fauria (2017), Shrub growth and expansion in the Arctic tundra: an assessment of controlling factors using an evidence-based approach, *Environmental Research Letters*, *12*(8), 085,007, doi:10.1088/1748-9326/aa7989.

Masson-Delmotte, V., et al. (2021), Climate change 2021: the physical science basis, *Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change*, *2*.

Mathisen, I. E., A. Mikheeva, O. V. Tutubalina, S. Aune, and A. Hofgaard (2014), Fifty years of tree line change in the Khibiny Mountains, Russia: advantages of combined remote sensing and dendroecological approaches, *Applied Vegetation Science*, *17*(1), 6–16, doi:10.1111/avsc.12038.

Matzner, E., and W. Borken (2008), Do freeze-thaw events enhance C and N losses from soils of different ecosystems? A review, *European Journal of Soil Science*, *59*(2), 274–284, doi:10.1111/j.1365-2389.2007.00992.x.

Mauritsen, T., et al. (2019), Developments in the mpi-m earth system model version 1.2 (mpi-esm1. 2) and its response to increasing co2, *Journal of Advances in Modeling Earth Systems*, *11*(4), 998–1038.

Mayr, S., P. Schmid, J. Laur, S. Rosner, K. Charra-Vaskou, B. Dämon, and U. G. Hacke (2014), Uptake of water via branches helps timberline conifers refill embolized xylem in late winter, *Plant Physiology*, *164*(4), 1731–1740, doi:10.1104/pp.114.236646.

Mazepa, V., and N. Devi (2007), Development of multistemmed life forms of Siberian larch as an indicator of climate change in the timberline ecotone of the Polar Urals, *Russian Journal of Ecology*, *38*(6), 440–443, doi:10.1134/S1067413607060112.

Mazepa, V., S. Shiyatov, and N. Devi (2011), Climate-driven change of the stand age structure in the polar Ural Mountains, *Climate Change: Geophysical Foundations and Ecological Effects*, pp. 377–402, doi:10.5772/915.

Mazepa, V. S. (2005), Stand density in the last millennium at the upper tree-line ecotone in the Polar Ural Mountains, *Canadian Journal of Forest Research*, *35*(9), 2082–2091, doi:10.1139/x05-111.

Mazhitova, G., G. Malkova, O. Chestnykh, and D. Zamolodchikov (2004), Active-layer spatial and temporal variability at european russian circumpolar-active-layer-monitoring (calm) sites, *Permafrost and Periglacial Processes*, *15*(2), 123–139.

McGuire, A. D., et al. (2018), Dependence of the evolution of carbon dynamics in the northern permafrost region on the trajectory of climate change, *Proceedings of the National Academy of Sciences*, *115*(15), 3882–3887.

Meester, R., and R. Roy (1996), *Continuum Percolation*, vol. 119, Cambridge University Press, doi:10.1017/CBO9780511895357.

Melnikov, E., M. Leibman, N. Moskalenko, and A. Vasiliev (2004), Active-layer monitoring in the cryolithozone of West Siberia, *Polar Geography*, *28*(4), 267–285, doi:10.1080/789610206.

Miller, R. D. (1980), Freezing phenomena in soils, *Applications of soil physics*.

Moore, I. D., R. Grayson, and A. Ladson (1991), Digital terrain modeling: a review of hydrological, geomorphological, and biological applications, *Hydrological Processes*, *5*(1), 3–30, doi:10.1002/hyp.3360050103.

Morel-Seytoux, H. J., P. D. Meyer, M. Nachabe, J. Tourna, M. T. Van Genuchten, and R. J. Lenhard (1996), Parameter equivalence for the brooks-corey and van genuchten soil characteristics: Preserving the effective capillary drive, *Water Resources Research*, *32*(5), 1251–1258.

Mu, S., and B. Ladanyi (1987), Modelling of coupled heat, moisture and stress field in freezing soil, *Cold Regions Science and Technology*, *14*(3), 237–246, doi:10.1016/0165-232X(87)90016-4.

Muñoz Sabater, J. (2019), ERA5-Land hourly data from 1981 to present, doi:10.24381/cds.e2161bac, accessed on 3 August 2022.

Myers-Smith, I. H., and D. S. Hik (2013), Shrub canopies influence soil temperatures but not nutrient dynamics: an experimental test of tundra snow–shrub interactions, *Ecology and Evolution*, *3*(11), 3683–3700.

Myers-Smith, I. H., et al. (2011), Shrub expansion in tundra ecosystems: dynamics, impacts and research priorities, *Environmental Research Letters*, *6*(4), 045,509, doi:10.1088/1748-9326/6/4/045509.

Myers-Smith, I. H., et al. (2015), Climate sensitivity of shrub growth across the tundra biome, *Nature Climate Change*, *5*(9), 887–891, doi:10.1038/nclimate2697.

Myers-Smith, I. H., et al. (2020), Complexity revealed in the greening of the Arctic, *Nature Climate Change*, *10*(2), 106–117, doi:10.1038/s41558-019-0688-1.

Nicolsky, D., and V. E. Romanovsky (2018), Modeling long-term permafrost degradation, *Journal of Geophysical Research: Earth Surface*, *123*(8), 1756–1771.

Nicolsky, D., V. Romanovsky, V. Alexeev, and D. Lawrence (2007), Improved modeling of permafrost dynamics in a gcm land-surface scheme, *Geophysical research letters*, *34*(8).

Nicolsky, D. J., V. E. Romanovsky, S. K. Panda, S. S. Marchenko, and R. R. Muskett (2017), Applicability of the ecosystem type approach to model permafrost dynamics across the alaska north slope, *Journal of Geophysical Research: Earth Surface*, *122*(1), 50–75.

Ochsner, T. E., and J. M. Baker (2008), In situ monitoring of soil thermal properties and heat flux during freezing and thawing, *Soil Science Society of America Journal*, *72*(4), 1025–1032, doi:10.2136/sssaj2007.0283.

Ochsner, T. E., T. J. Sauer, and R. Horton (2006), Field tests of the soil heat flux plate method and some alternatives, *Agronomy Journal*, *98*(4), 1005–1014, doi: 10.2134/agronj2005.0249.

Osterkamp, T., M. Jorgenson, E. Schuur, Y. Shur, M. Kanevskiy, J. Vogel, and V. Tumskoy (2009), Physical and ecological changes associated with warming permafrost and thermokarst in interior alaska, *Permafrost and Periglacial Processes*, *20*(3), 235–256.

Painter, S., J. Moulton, and C. Wilson (2013), Modeling challenges for predicting hydrologic response to degrading permafrost, *Hydrogeology Journal*, *21*(1), 221–224, doi:10.1007/s10040-012-0917-4.

Painter, S. L., and S. Karra (2014), Constitutive model for unfrozen water content in subfreezing unsaturated soils, *Vadose Zone Journal*, *13*(4), 1–8, doi:10.2136/vzj2013.04.0071.

Painter, S. L., E. T. Coon, A. L. Atchley, M. Berndt, R. Garimella, J. D. Moulton, D. Svyatskiy, and C. J. Wilson (2016), Integrated surface/subsurface permafrost thermal hydrology: Model formulation and proof-of-concept simulations, *Water Resources Research*, *52*(8), 6062–6077, doi:10.1002/2015WR018427.

Paulsen, J., and C. Körner (2014), A climate-based model to predict potential treeline position around the globe, *Alpine Botany*, *124*(1), 1–12, doi:10.1007/s00035-014-0124-0.

Pazdrowski, W., T. Jelonek, A. Tomczak, I. Stypula, and S. Splawa-Neyman (2007), Proportion of sapwood and heartwood and selected biometric features in larch trees (Larix decidua Mill.), *Wood Research*, *54*(4), 1–16.

Pearson, R. G., S. J. Phillips, M. M. Loranty, P. S. Beck, T. Damoulas, S. J. Knight, and S. J. Goetz (2013), Shifts in Arctic vegetation and associated feedbacks under climate change, *Nature Climate Change*, *3*(7), 673–677, doi:10.1038/nclimate1858.

Plotkin, J. B., J. Chave, and P. S. Ashton (2002), Cluster analysis of spatial patterns in Malaysian tree species, *The American Naturalist*, *160*(5), 629–644, doi:10.1086/342823.

Porter, C., et al. (2018), ArcticDEM, doi:10.7910/DVN/OHHUKH, `https://doi.org/10.7910/DVN/OHHUKH`.

Raäisaänen, J. (2007), How reliable are climate models?, *Tellus A: Dynamic Meteorology and Oceanography*, *59*(1), 2–29.

Ripley, B. D. (1976), The second-order analysis of stationary point processes, *Journal of Applied Probability*, *13*(2), 255–266, doi:10.2307/3212829.

Riseborough, D., N. Shiklomanov, B. Etzelmüller, S. Gruber, and S. Marchenko (2008), Recent advances in permafrost modelling, *Permafrost and Periglacial Processes*, *19*(2), 137–156.

Romanovsky, V., and T. Osterkamp (2000), Effects of unfrozen water on heat and mass transport processes in the active layer and permafrost, *Permafrost and Periglacial Processes*, *11*(3), 219–239.

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

Romanovsky, V. E., et al. (2010b), Thermal state of permafrost in russia, *Permafrost and Periglacial Processes*, *21*(2), 136–155.

Sargsyan, K., C. Safta, H. N. Najm, B. J. Debusschere, D. Ricciuto, and P. Thornton (2014), Dimensionality reduction for complex models via bayesian compressive sensing, *International Journal for Uncertainty Quantification*, *4*(1).

Sargsyan, K., H. Najm, and R. Ghanem (2015), On the statistical calibration of physical models, *International Journal of Chemical Kinetics*, *47*(4), 246–276.

Sato, H., and H. Kobayashi (2018), Topography controls the abundance of siberian larch forest, *Journal of Geophysical Research: Biogeosciences*, *123*(1), 106–116, doi:10.1002/2017JG004096.

Schaefer, K., H. Lantuit, V. E. Romanovsky, E. A. Schuur, and R. Witt (2014), The impact of the permafrost carbon feedback on global climate, *Environmental Research Letters*, *9*(8), 085,003.

Schuur, E. A., et al. (2013), Expert assessment of vulnerability of permafrost carbon to climate change, *Climatic Change*, *119*(2), 359–374.

Schuur, E. A., et al. (2015), Climate change and the permafrost carbon feedback, *Nature*, *520*(7546), 171–179.

Seland, Ø., et al. (2020), Overview of the norwegian earth system model (noresm2) and key climate response of cmip6 deck, historical, and scenario simulations, *Geoscientific Model Development*, *13*(12), 6165–6200.

Sellar, A. A., et al. (2019), Ukesm1: Description and evaluation of the uk earth system model, *Journal of Advances in Modeling Earth Systems*, *11*(12), 4513–4558.

Shary, P. A., L. S. Sharaya, and A. V. Mitusov (2005), The problem of scale-specific and scale-free approaches in geomorphometry, *Geografia Fisica e Dinamica Quaternaria*, *28*(1), 81–101.

Shati, F., S. Prakash, H. Norouzi, and R. Blake (2018), Assessment of differences between near-surface air and soil temperatures for reliable detection of high-latitude freeze and thaw states, *Cold Regions Science and Technology*, *145*, 86–92.

Sheshukov, A., and A. Egorov (2002), Frozen barrier evolution in saturated porous media, *Advances in Water Resources*, *25*(6), 591–599, doi:10.1016/S0309-1708(02)00033-7.

Sheshukov, A. Y., and J. L. Nieber (2011), One-dimensional freezing of nonheaving unsaturated soils: Model formulation and similarity solution, *Water Resources Research*, *47*(11), doi:10.1029/2011WR010512.

Shi, X., A.-X. Zhu, J. Burt, W. Choi, R. Wang, T. Pei, B. Li, C. Qin, et al. (2007), An experiment using a circular neighborhood to calculate slope gradient from a DEM, *Photogrammetric Engineering & Remote Sensing*, *73*(2), 143–154, doi:10.14358/PERS.73.2.143.

Shiklomanov, N. I., D. A. Streletskiy, T. B. Swales, and V. A. Kokorev (2017), Climate change and stability of urban infrastructure in russian permafrost regions: prognostic assessment based on gcm climate projections, *Geographical Review*, *107*(1), 125–142.

Shiryaev, A. G., P. A. Moiseev, U. Peintner, N. M. Devi, V. V. Kukarskih, and V. V. Elsakov (2019), Arctic greening caused by warming contributes to compositional changes of mycobiota at the Polar Urals, *Forests*, *10*(12), 1112, doi:10.3390/f10121112.

Shiyatov, S., and V. Mazepa (2011), Climate-driven dynamics of the forest-tundra vegetation in the Polar Ural Mountains, *Contemporary Problems of Ecology*, *4*(7), 758–768, doi:10.1134/S1995425511070071.

Shiyatov, S., and V. Mazepa (2015), Contemporary expansion of Siberian larch into the mountain tundra of the Polar Urals, *Russian Journal of Ecology*, *46*(6), 495–502, doi:10.1134/S1067413615060168.

Shiyatov, S., M. Terent'Ev, and V. Fomin (2005), Spatiotemporal dynamics of forest-tundra communities in the polar urals, *Russian Journal of Ecology*, *36*(2), 69–75, doi:10.1007/s11184-005-0051-9.

Shiyatov, S., M. Terent'Ev, V. Fomin, and N. E. Zimmermann (2007), Altitudinal and horizontal shifts of the upper boundaries of open and closed forests in the Polar Urals in the 20th century, *Russian Journal of Ecology*, *38*(4), 223–227, doi:10.1134/S1067413607040017.

Sivia, D., and J. Skilling (2006), *Data analysis: a Bayesian tutorial*, OUP Oxford.

Smith, M., and D. Riseborough (1996), Permafrost monitoring and detection of climate change, *Permafrost and periglacial processes*, *7*(4), 301–309.

Smith, S., V. Romanovsky, A. Lewkowicz, C. R. Burn, M. Allard, G. Clow, K. Yoshikawa, and J. Throop (2010), Thermal state of permafrost in north america: a contribution to the international polar year, *Permafrost and Periglacial Processes*, *21*(2), 117–135.

Smith, S. L., M. M. Burgess, D. Riseborough, and F. Mark Nixon (2005), Recent trends from canadian permafrost thermal monitoring network sites, *Permafrost and periglacial processes*, *16*(1), 19–30.

Smolyak, S. A. (1963), Quadrature and interpolation formulas for tensor products of certain classes of functions, in *Doklady Akademii Nauk*, vol. 148, pp. 1042–1045, Russian Academy of Sciences.

Sraj, I., K. T. Mandli, O. M. Knio, C. N. Dawson, and I. Hoteit (2014), Uncertainty quantification and inference of Manning's friction coefficients using DART buoy data during the Tōhoku tsunami, *Ocean Modelling*, *83*, 82–97, doi:10.1016/j.ocemod.2014.09.001.

Streletskiy, D., O. Anisimov, and A. Vasiliev (2015), Permafrost degradation, in *Snow and ice-related hazards, risks, and disasters*, pp. 303–344, Elsevier.

Streletskiy, D. A., N. I. Shiklomanov, and F. E. Nelson (2012), Permafrost, infrastructure, and climate change: a gis-based landscape approach to geotechnical modeling, *Arctic, Antarctic, and Alpine Research*, *44*(3), 368–380.

Sturm, M., C. Racine, and K. Tape (2001), Increasing shrub abundance in the Arctic, *Nature*, *411*(6837), 546–547, doi:10.1038/35079180.

Sullivan, P. F., S. B. Ellison, R. W. McNown, A. H. Brownlee, and B. Sveinbjörnsson (2015), Evidence of soil nutrient availability as the proximate constraint on growth of treeline trees in northwest Alaska, *Ecology*, *96*(3), 716–727, doi:10.1890/14-0626.1.

Tape, K., M. Sturm, and C. Racine (2006), The evidence for shrub expansion in Northern Alaska and the Pan-Arctic, *Global Change Biology*, *12*(4), 686–702, doi:10.1111/j.1365-2486.2006.01128.x.

Tape, K. D., D. D. Gustine, R. W. Ruess, L. G. Adams, and J. A. Clark (2016), Range expansion of moose in Arctic Alaska linked to warming and increased shrub habitat, *PloS One*, *11*(4), e0152,636, doi:10.1371/journal.pone.0152636.

Tarek, M., F. P. Brissette, and R. Arsenault (2020), Evaluation of the era5 reanalysis as a potential reference dataset for hydrological modelling over north america, *Hydrology and Earth System Sciences*, *24*(5), 2527–2544.

Tarolli, P., G. Sofia, and G. Dalla Fontana (2012), Geomorphic features extraction from high-resolution topography: landslide crowns and bank erosion, *Natural Hazards*, *61*(1), 65–83, doi:10.1007/s11069-010-9695-2.

Taylor, G. S., and J. N. Luthin (1978), A model for coupled heat and moisture transfer during soil freezing, *Canadian geotechnical journal*, *15*(4), 548–555, doi:10.1139/t78-058.

Tebaldi, C., R. L. Smith, D. Nychka, and L. O. Mearns (2005), Quantifying uncertainty in projections of regional climate change: A bayesian approach to the analysis of multimodel ensembles, *Journal of Climate*, *18*(10), 1524–1540.

Teufel, B., and L. Sushama (2019), Abrupt changes across the arctic permafrost region endanger northern development, *Nature Climate Change*, *9*(11), 858–862.

Tily, R., and C. Brace (2006), A study of natural neighbour interpolation and its application to automotive engine test data, *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, *220*(7), 1003–1017, doi:10.1243/09544070D11004.

Tokarska, K. B., M. B. Stolpe, S. Sippel, E. M. Fischer, C. J. Smith, F. Lehner, and R. Knutti (2020), Past warming trend constrains future warming in cmip6 models, *Science Advances*, *6*(12), eaaz9549.

Tsilingiridis, G., and K. Papakostas (2014), Investigating the relationship between air and ground temperature variations in shallow depths in northern greece, *Energy*, *73*, 1007–1016.

Tyler, N. J., et al. (2007), Saami reindeer pastoralism under climate change: applying a generalized framework for vulnerability studies to a sub-arctic social–ecological system, *Global Environmental Change*, *17*(2), 191–206.

Van Bogaert, R., K. Haneca, J. Hoogesteger, C. Jonasson, M. De Dapper, and T. V. Callaghan (2011), A century of tree line changes in sub-Arctic Sweden shows local and regional variability and only a minor influence of 20th century climate warming, *Journal of Biogeography*, *38*(5), 907–921, doi:10.1111/j.1365-2699.2010.02453.x.

Vasiliev, A., M. Leibman, and N. Moskalenko (2008), Active layer monitoring in West Siberia under the CALM II Program, in *Proc. Ninth Int. Conf. Permafrost*, vol. 2, pp. 1815–21.

Vasiliev, A. A., D. S. Drozdov, A. G. Gravis, G. V. Malkova, K. E. Nyland, and D. A. Streletskiy (2020), Permafrost degradation in the western russian arctic, *Environmental Research Letters*, *15*(4), 045,001.

Vavrus, S. J., M. M. Holland, A. Jahn, D. A. Bailey, and B. A. Blazey (2012), Twenty-first-century Arctic climate change in CCSM4, *Journal of Climate*, *25*(8), 2696–2710, doi:10.1175/JCLI-D-11-00220.1.

Vertessy, R., R. Benyon, S. O'sullivan, and P. Gribben (1995), Relationships between stem diameter, sapwood area, leaf area and transpiration in a young mountain ash forest, *Tree Physiology*, *15*(9), 559–567, doi:10.1093/treephys/15.9.559.

Walker, D. A., et al. (2005), The Circumpolar Arctic vegetation map, *Journal of Vegetation Science*, *16*(3), 267–282, doi:10.1111/j.1654-1103.2005.tb02365.x.

Walker, M. D., et al. (2006), Plant community responses to experimental warming across the tundra biome, *Proceedings of the National Academy of Sciences*, *103*(5), 1342–1346, doi:10.1073/pnas.0503198103.

Walvoord, M. A., and B. L. Kurylyk (2016), Hydrologic impacts of thawing permafrost—A review, *Vadose Zone Journal*, *15*(6), doi:10.2136/vzj2016.01.0010.

Wang, J., and R. Bras (1999), Ground heat flux estimated from surface soil temperature, *Journal of Hydrology*, *216*(3-4), 214–226, doi:10.1016/S0022-1694(99)00008-6.

Wang, Q., R. Horton, and M. Shao (2002), Horizontal infiltration method for determining brooks-corey model parameters, *Soil Science Society of America Journal*, *66*(6), 1733–1739.

Wang, Z.-H., and E. Bou-Zeid (2012), A novel approach for the estimation of soil ground heat flux, *Agricultural and Forest Meteorology*, *154*, 214–221, doi:10.1016/j.agrformet.2011.12.001.

Way, R., and C. Lapalme (2021), Does tall vegetation warm or cool the ground surface? constraining the ground thermal impacts of upright vegetation in northern environments, *Environmental Research Letters*, *16*(5), 054,077.

Wilcox, E. J., D. Keim, T. de Jong, B. Walker, O. Sonnentag, A. E. Sniderhan, P. Mann, and P. Marsh (2019), Tundra shrub expansion may amplify permafrost thaw by advancing snowmelt timing, *Arctic Science*, *5*(4), 202–217.

Williams, K., et al. (2018), The met office global coupled model 3.0 and 3.1 (gc3. 0 and gc3. 1) configurations, *Journal of Advances in Modeling Earth Systems*, *10*(2), 357–380.

Wilmking, M., G. P. Juday, V. A. Barber, and H. S. Zald (2004), Recent climate warming forces contrasting growth responses of white spruce at treeline in Alaska through temperature thresholds, *Global Change Biology*, *10*(10), 1724–1736, doi:10.1111/j.1365-2486.2004.00826.x.

Wu, B., S. P. Oncley, H. Yuan, and F. Chen (2020), Ground heat flux determination based on near-surface soil hydro-thermodynamics, *Journal of Hydrology*, *591*, 125,578, doi:10.1016/j.jhydrol.2020.125578.

Xu, D., V. Y. Ivanov, J. Kim, and S. Fatichi (2019), On the use of observations in assessment of multi-model climate ensemble, *Stochastic Environmental Research and Risk Assessment*, *33*(11), 1923–1937.

Zeng, H., G. Jia, and B. C. Forbes (2013), Shifts in Arctic phenology in response to climate and anthropogenic factors as detected from multiple satellite time series, *Environmental Research Letters*, *8*(3), 035,036, doi:10.1088/1748-9326/8/3/035036.

Zevenbergen, L. W., and C. R. Thorne (1987), Quantitative analysis of land surface topography, *Earth Surface Processes and Landforms*, *12*(1), 47–56, doi:10.1002/esp.3290120107.

Zhang, W., P. A. Miller, B. Smith, R. Wania, T. Koenigk, and R. Döscher (2013), Tundra shrubification and tree-line advance amplify arctic climate warming: results from an individual-based dynamic vegetation model, *Environmental Research Letters*, *8*(3), 034,023, doi:10.1088/1748-9326/8/3/034023.

Zhang, X., S. F. Sun, and Y. Xue (2007), Development and testing of a frozen soil parameterization for cold region studies, *Journal of Hydrometeorology*, *8*(4), 690–701, doi:10.1175/JHM605.1.

Zhang, Y., W. Chen, and D. W. Riseborough (2008), Transient projections of permafrost distribution in canada during the 21st century under scenarios of climate change, *Global and Planetary Change*, *60*(3-4), 443–456.

Zhou, W., et al. (2021), Spatiotemporal dynamics of encroaching tall vegetation in timberline ecotone of the polar urals region, russia, *Environmental Research Letters*, *17*(1), 014,017, doi:10.1088/1748-9326/ac3694.

Ziehn, T., M. A. Chamberlain, R. M. Law, A. Lenton, R. W. Bodman, M. Dix, L. Stevens, Y.-P. Wang, and J. Srbinovsky (2020), The australian earth system model: Access-esm1. 5, *Journal of Southern Hemisphere Earth Systems Science*, *70*(1), 193–214.