# Simulating the Present-Day and Future Distribution of Permafrost in the UVic Earth System Climate Model

by

Christopher Alexander Avis
B.Sc., University of Victoria, 2005
M.Sc., University of Victoria, 2007

A Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the School of Earth and Ocean Sciences

# Simulating the Present-Day and Future Distribution of Permafrost in the UVic Earth System Climate Model

by

Christopher Alexander Avis
B.Sc., University of Victoria, 2005
M.Sc., University of Victoria, 2007

Supervisory Committee

Dr. Andrew Weaver, Supervisor
(School of Earth and Ocean Sciences, University of Victoria)

Dr. Katrin Meissner, Member
(Climate Change Research Centre, University of New South Wales)

Dr. Vivek Arora, Member
(Canadian Centre for Climate Modelling and Analysis)

Dr. Trevor Lantz, Outside Member
(School of Environmental Studies, University of Victoria)

**Supervisory Committee**

---

Dr. Andrew Weaver, Supervisor
(School of Earth and Ocean Sciences, University of Victoria)

---

Dr. Katrin Meissner, Member
(Climate Change Research Centre, University of New South Wales)

---

Dr. Vivek Arora, Member
(Canadian Centre for Climate Modelling and Analysis)

---

Dr. Trevor Lantz, Outside Member
(School of Environmental Studies, University of Victoria)

## ABSTRACT

Warming over the past century has been greatest in high-latitudes over land and a number of environmental indicators suggest that the Arctic climate system is in the process of a major transition. Given the magnitude of observed and projected changes in the Arctic, it is essential that a better understanding of the characteristics of the Arctic climate system be achieved. In this work, I report on modifications to the UVic Earth System Climate model to allow it to represent regions of perennially-frozen ground, or permafrost. I examine the model's representation of the Arctic climate during the $20^{th}$ Century and show that it capably represents the distribution and thermal state of permafrost in the present-day climate system. I use Representative Concentration Pathways to examine a range of possible future permafrost states to the year 2500. A suite of sensitivity experiments is used to better understand controls on permafrost. I demonstrate the potential for radical environmental changes in the Arctic over the $21^{st}$ Century including continued warming, enhanced precipitation and a reduction of between 29 and 54 % of the present-day permafrost area by 2100. Model projections show that widespread loss of high-latitude wetlands may accompany the loss of near surface permafrost.

# Contents

# List of Tables

# List of Figures

# ACKNOWLEDGEMENTS

I would like to thank my graduate supervisor, Dr. Andrew Weaver, for agreeing to take me on as a graduate student and for his invaluable mentorship as well as my committee for their thoughtful input into my work. Michael Eby and Ed Wiebe provided invaluable assistance with questions about the model and technical support. I also benefitted greatly from discussions with Rita Wania, Damon Matthews, Kirsten Zickfeld, Jed Kaplan, Jeremy Fyke and Joe Melton. This work was supported by an NSERC CGS research award and graduate support from Dr. Andrew Weaver and the University of Victoria and I am extremely grateful for this financial support.

I would also like to thank all of the graduate students and summer students who have been members of the UVic Climate Lab during my time here for their insight and companionship. Finally, this research would have been impossible were it not for the support of my family and my wife Dianna who encouraged me to give it my all.

# DEDICATION

For Dianna

# Chapter 1

# Introduction and Motivation

## 1.1 Context

In its 2007 Summary for Policy Makers (IPCC, 2007), the Intergovernmental Panel on Climate Change (IPCC) stated that *"warming of the climate system is unequivocal, as is now evident from observations of increases in global average air and ocean temperatures, widespread melting of snow and ice, and rising global average sea level"*, noting that *"most of the observed increase in global average temperatures since the mid-20$^{th}$ Century is very likely due to the observed increase in anthropogenic greenhouse gas concentrations"*.

The global average temperature increase from 1906-2005 was 0.74 °C ± 0.18 °C, with the rate of warming over the second half of the 20$^{th}$ Century nearly twice the 100-year average (IPCC, 2007). While the warming of the climate system is widespread, globally, it is nowhere more pronounced than in the Arctic where the regional-average rate of warming over the past century is nearly twice the global average (IPCC, 2007), with temperatures rising most rapidly in winter. The rise of surface air temperature in the Arctic is especially strong due to several region-specific factors (ACIA, 2004; Serreze and Barry, 2011). Melting ice and snow expose less reflective ocean and land surfaces, thereby increasing the absorption of solar radiation by the land surface. A reduction in sea ice also reduces the insulation that limits heat fluxes between Arctic atmosphere and the ocean (Holland et al., 2001), which results in warming of the atmosphere during winter. The lower Arctic atmosphere also tends to be stably stratified, which limits vertical mixing, meaning that surface warming is less readily transmitted throughout the troposphere as opposed to at lower latitudes, where free convection can more readily occur (Gravesen and Wange, 2009). There is also evidence that changes in atmospheric and oceanic circulations may have played some role in increasing heat fluxes into the Arctic, thereby contributing to polar amplification (Serreze and Barry, 2011). Averaged over the year, cloud cover in the Arctic tends to have a warming influence, owing to absorption and downwards re-radiation of long-wave radiation by clouds (Serreze and Barry, 2011). Arctic cloud cover might be expected to increase owing to increased moisture transport into the Arctic and increased evaporation within the Arctic; the enhanced cloud cover would thereby contribute to Arctic amplification, as would the increased abundance of water vapour, through its role as a greenhouse gas. Finally, modelling studies suggest that black carbon aerosols may have contributed to the strong Arctic warming through their lowering of the surface albedo when deposited on highly reflective snow and ice (Serreze and Barry, 2011).

In this work, like some recent literature review papers (ACIA, 2004; Hinzman et al., 2005), I shall use the term Arctic to describe areas of high northern latitudes that are dominated by the influence of snow, ice and permafrost, rather than the area strictly encompassed by the Arctic circle. In addition to the observed warming, other recent observations suggest that the Arctic environment is in the midst of profound changes. Briefly, after ACIA (2004), the most significant observed changes are:

1. **Increasing precipitation**: Precipitation rates in the Arctic have increased by roughly 8 % over the past century, with the increase in precipitation most pronounced in winter months.

2. **Decreasing snow cover**: Despite increased rates of precipitation, snow cover has decreased in most areas, with an overall decrease of about 10 % over the past 30 years (ACIA (2004)). Satellite observations over the period 1966 to 2005 reveal that snow cover has decreased in every month save for November and December (Lemke et al., 2007). A comprehensive picture of changes in snow depth is not yet available, though depths appear to have decreased over much of Canada and Eastern Russia, but have increased elsewhere in Russia, associated with increases in precipitation (Walsh, 2005). Observations from Barrow, Alaska over the past 60 years also suggest a trend towards an earlier snowfree season (Hinzman et al., 2005).

3. **Melting glaciers and ice sheets**: The rate of loss of ice from glaciers and the Greenland Ice sheet is increasing. Glaciers throughout the Arctic are melting,

with particularly high rates of glacial retreat in Alaska. The Greenland Ice sheet is also experiencing melt; microwave satellite data suggests that the total melt area of the ice sheet increased by 30 % between 1979 and 2008, with particularly extreme melt in 2007 where nearly 50 % of the ice sheet area was affected (Allison et al., 2009). Increased glacial melt, coupled with increased precipitation have also contributed to increased river discharge to the Arctic Ocean in recent years.

4. **Reduction in sea ice**: End of summer sea ice extent in the Arctic Ocean has decreased dramatically at a rate of 11.1 ± 3.3 %/decade, far faster than predicted by climate models used in the IPCC Assessment Reports (Allison et al., 2009). In 2007, the sea ice extent dropped dramatically to its all-time minimum, a value that was nearly surpassed in 2010 and 2011. Alarmingly, recent observations of Arctic Ocean ice thickness and volume suggest that thick, multi-year ice was observed to decrease in volume by 57 % and in areal cover by 33 % from 2003 to 2008 (Kwok et al., 2009). Ice present on lake and rivers has been affected too as these bodies of water are increasingly characterized by later winter freeze up and earlier spring breakup.

Climate change predictions from general circulation models (GCMs) indicate that future warming will be greatest in high-latitude regions of the Northern Hemisphere. By the end of the $21^{st}$ Century, it is estimated that the average warming of Arctic will be between 5.9 °C and 3.4 °C [1](Christensen et al., 2007). Many of the unique facets of the Arctic environment and the ecosystems that lie within it are set by the characteristics of the cryosphere. These characteristics, as the above observations suggest, are in the process of rapid change. As a result of the high rates of projected warming, the Arctic region and the tundra and boreal forest ecosystems have been identified as particularly vulnerable to climate change (IPCC, 2007). Furthermore, as shall be discussed, a number of processes operate within the Arctic climate system with the potential to influence the global climate as a whole. As such, understanding and predicting the warming response of the components of the cryosphere that influence this region is a particularly important challenge. This work focuses on understanding and projecting changes in the largest single component of the cryosphere: frozen

---

[1] These results are multi-model ensemble averages from atmosphere-ocean GCM simulations conducted in support of the IPCC Fourth Assessment Report (AR4) (Christensen et al., 2007). The upper and lower limits come from simulations forced by the A2 and B1 emissions scenarios, respectively

ground, which covers approximately 51 % of the exposed land area in the Northern Hemisphere (Lemke et al., 2007) and includes regions of near-surface ground affected by short-term freeze-thaw cycles, seasonally frozen ground and permafrost.

## 1.2 Characteristics of the Ground in Permafrost Regions

At this point, it is important to clarify some of the terminology used in this thesis work. As a result of forces that act on water molecules in the soil, the increasing pressure that moisture is subject to with depth and the presence of dissolved salts, water in the ground begins to freeze at temperatures below the nominal freezing point at the surface (0 °C). As shall be discussed in the next chapter, freezing in the soil also occurs gradually over a range of temperatures, rather than at a single temperature as for unconfined water at the surface. Some amount of unfrozen moisture may be present in the ground even when ground temperatures are as low as -10 °C; as much permafrost is warmer than -5 °C, permafrost may contain reasonable amounts of unfrozen moisture.

A distinction must therefore be made when classifying the ground in permafrost regions. Generally speaking, classification is usually based on the ground being cryotic (below 0 °C) versus non-cryotic (above 0 °C) or frozen (containing ice) versus unfrozen (containing no ice). Adding further confusion, frozen ground (ground containing some ice) generally does not imply that all moisture in the ground is in the form of ice - that is to say, there may be some liquid moisture present in frozen ground. In my thesis, I have chosen to classify ground layers based on their thermal state (ie. based on the cryotic criterion) as this offers an unambiguous definition in terms of a single temperature threshold and a straightforward way of identifying ground layers from model output. As I will also need to discuss phase changes in this thesis, the term *frozen* will be applied strictly to describe frozen moisture (ice) while *unfrozen* will be used solely to describe unfrozen moisture (liquid water) with freezing / thawing used to describe the phase change of water in the ground (rather than the crossing of the 0 °C temperature threshold).

Figure 1.1 shows a representative thermal profile in ground containing permafrost, plotted to show the minimum and maximum ground temperatures at given depths. This figure should not be seen as a snapshot of temperatures at two different times of the year as the maximum and minimum temperatures in different ground layers occur at different times of the year. The depth of zero annual temperature amplitude, (often practically identified as the depth at which the annual temperature change is less than 0.1 °C) lies within the permafrost (Williams and Smith (1989)). Below this level, ground temperatures effectively remain constant with time (provided the ground is in equilibrium with the climate) and gradually increase with depth following the *geothermal gradient*[2].

I adopt the most widely used definition of permafrost: *ground (soil or rock) that remains at or below 0 °C for at least two years*, with the *permafrost table* identifying its top and the *permafrost base* occuring where the geothermal gradient results in ground temperatures again rising above 0 °C (Harris et al., 1988). Thus, while the permafrost table is set by the annual cycle of freezing and thawing, the base of permafrost is determined by mean annual temperatures and the value of the geothermal heat flux (Williams and Smith, 1989). I define the *active layer* as the layer above permafrost that seasonally rises and falls above 0 °C, which differs from the standard definition of the active layer as *the top layer of ground subject to annual thawing and freezing in areas underlain by permafrost* (Harris et al., 1988). This more common definition may incorporate the top most layer of permafrost (above the depth of zero annual temperature amplitude) which experiences some phase changes throughout the year - the *seasonally active permafrost*. The definition of the active layer as a layer that seasonally freezes and thaws means that the permafrost regions are a subset of the larger geographic area that is subject to seasonally frozen ground; in the Northern Hemisphere, about one half of the exposed land mass experiences this seasonal freeze and thaw (Zhang et al., 2003). Where the permafrost is in climatic equilibrium, the active layer extends to the permafrost table, but if permafrost is in the process of degradation, a *talik* or unfrozen layer may exist between the active layer and the permafrost table. Because of the presence of taliks and seasonally active permafrost, the thickness of the active layer is, in general, not synonymous with the depth of the permafrost table.

---

[2] The geothermal gradient is the increase in temperature with depth in the ground, associated with the *geothermal heat flux*: the amount of sensible heat escaping from the interior of the Earth through conduction (Harris et al., 1988)

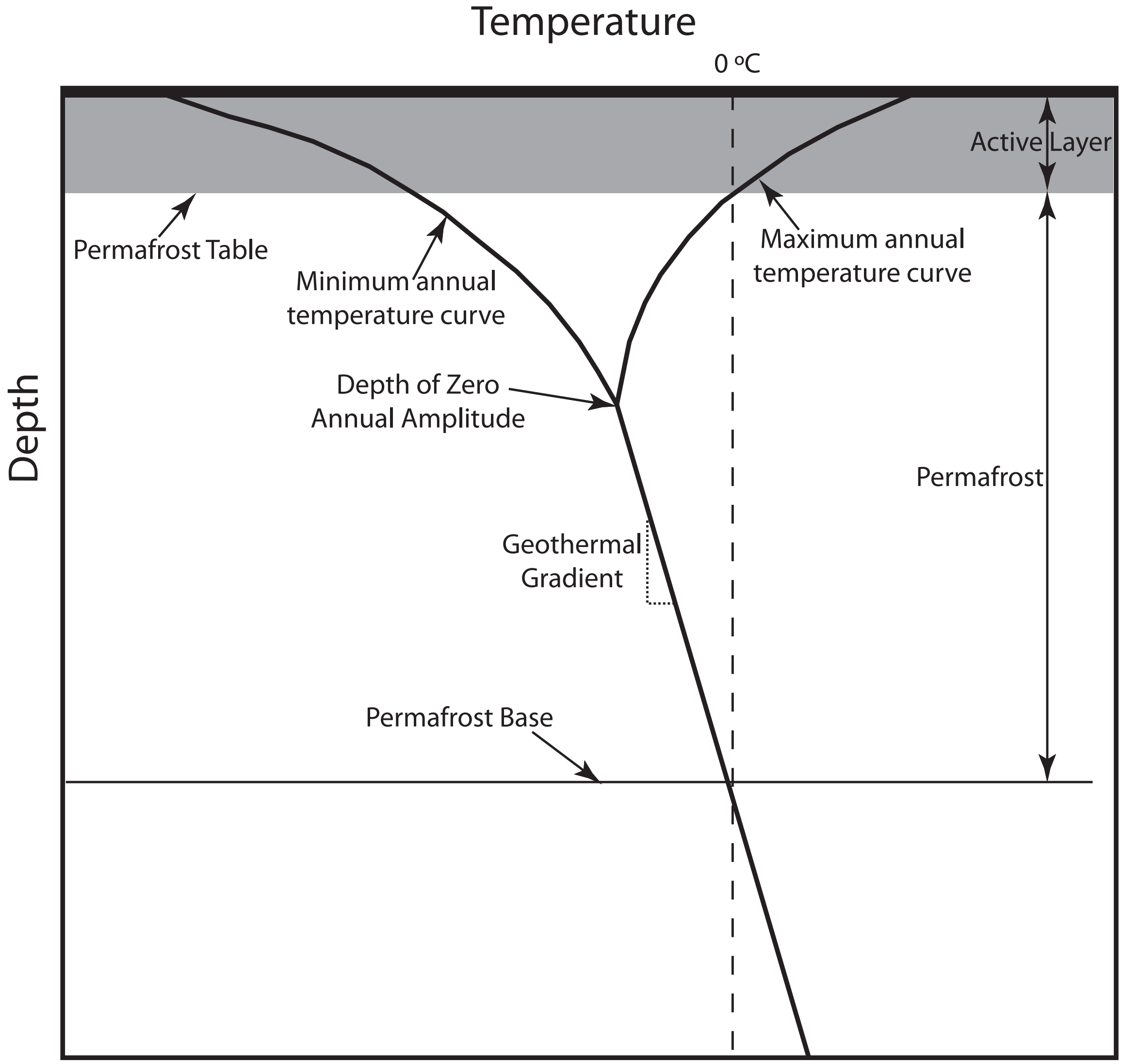


Figure 1.1: Representative annual cycle in ground temperature with depth in permafrost regions. Definitions of permafrost characteristics based on thermal criteria as used in this thesis are also illustrated. The figure was designed after similar figures in Smith and Burgess (2002) and French (2007).

## 1.3 Geographic Controls on Permafrost

The presence and characteristics of permafrost are determined by both large-scale geographic factors and site-specific conditions. At large scales, there are two dominant controls on the existence of permafrost. As the amount of incoming solar radiation received at a site throughout the year generally decreases towards the poles, a site's latitude exerts a strong influence on the existence of permafrost, its temperature and depth. Secondly, altitude also has a strong influence on permafrost characteristics as surface air temperature tends to decrease with elevation. Consequently, permafrost may be present at relatively low latitudes if a site's elevation is sufficiently high. Such permafrost may be restricted to high elevation pockets in mountainous regions (alpine permafrost) or broad areas at high elevation plateaus (montaine permafrost) as in the permafrost in the Himalayan Plateau. Additionally, there are a number of region- and site-specific permafrost controls. These include a site's aspect, vegetation and snow cover, the thermal characteristics (heat capacity and thermal conductivity) of the subsurface, the geothermal heat flux, the presence of water bodies adjacent to a site and the influence of regional climate circulations (such as warm ocean currents, for example) (French, 2007). Land surface characteristics are highly heterogeneous at high latitudes, varying substantially on scales as small as a meter, meaning that characteristics of permafrost regions such as mean annual ground temperature, active layer thickness and permafrost thickness can vary on comparable scales.

The competing effects of large-scale and site-specific controls mean that large-scale mapping of permafrost is extremely challenging. At very high-latitudes and altitudes permafrost is ubiquitous barring exceptional site specific controls (such as large water bodies and regions of extreme geothermal heat flux, as in Icelandic hotsprings). Near its low -latitude and -altitude limits, site-specific ground conditions will determine its presence. Consequently, the distribution of permafrost at large scales is generally expressed in terms of *zonation* which represents an estimate of the fraction of ground within a region that is actually underlain by permafrost. The zonation units adopted by the International Permafrost Association's Northern Hemisphere permafrost map (displayed in Figure 1.2) are continuous (90 % - 100 % of ground underlain by permafrost), discontinuous (50 % - 90 %), sporadic (10 % - 50 %) and isolated (0 - 10 %). Williams and Smith (1989) suggests that the predominantly latitudinal pattern of permafrost zonation might be best envisioned as representing a gradual transition (with increased occurrence of permafrost) from regions of seasonally frozen ground in the temperate regions to perennially frozen ground of the north. This is well illustrated by Figure 1.3, in which an idealized meridional transect demonstrates that as one moves northward from the southern limits of the discontinuous permafrost zone, permafrost gradually becomes uniformly present, while its thickness increases and that of the active layer decreases.

In North America the known thickness of permafrost typically ranges from 15-30 m at its southern limit to thicknesses of over 500 m in the Canadian Arctic; Russian permafrost is typically thicker and ranges in thickness from 300 m near its southern limit to over 600 m along the Siberian coast (French, 2007). The deepest known permafrost occurs at a depth of 1,450 m in Siberia (Davis, 2001), though in general, the maximum depth of frozen ground exceeds 1,000 m only in fairly limited areas (Williams and Smith, 1989). Active layer thicknesses typically range from 1 metre or more in sub-arctic regions to thicknesses on the order of a few 10s of centimetres in polar regions (French, 2007).

While nearly one quarter of the exposed Northern Hemisphere land surface lies within these permafrost zones, the amount of ground actually underlain by permafrost (as judged by the area of permafrost zones multiplied by the estimated fraction of ground underlain by permafrost within such zones) is considerably less and is estimated to range between 12.8 and 17.8 % of the exposed Northern Hemisphere land area (Zhang et al., 2000).

While permafrost is also found in the Southern Hemisphere, it is considerably less extensive and knowledge about its state and distribution is limited. Permafrost is known to occur in high elevation regions in the Andes, the New Zealand Alps and mountain summits in Irian Java and Papua New Guinea; it is found at lower elevations in Patagonia, on the sub-Antarctic islands and underlying exposed land in Antarctica (Boelhouwers and Hall, 2002; Heginbottom, 2002), but there is no hemisphere-wide permafrost map available for the Southern Hemisphere.

Permafrost is also known to be present as subsea permafrost on the shallow continental shelves of polar seas (Davis, 2001). Detailed knowledge of the distribution of subsea permafrost is limited and is determined using a combination of modelling and direct measurement from drilling or acoustic geophysical surveying (French, 2007; Walsh, 2005)) Subsea permafrost is of interest as it is thought to contain large volumes of methane in the form of gas hydrates, in which gas molecules are trapped in the interstices of an ice crystal lattice when temperatures are sufficiently cold and pressures are sufficiently high. Degradation of hydrates could substantially increase the flux of methane to the atmosphere, amplifying climate warming, although modelling studies suggest that catastrophic release of methane hydrates is unlikely this century (Brook et al., 2008). Owing to the uncertainty in the distribution of subsea and Southern Hemisphere permafrost, this work will largely focus on the distribution of terrestrial permafrost in the Northern Hemisphere.

The above discussion of permafrost controls implicitly assumes that permafrost is in climatic equilibrium - that is that the temperature in ground layers is no longer

Figure 1.2: Permafrost Zonation in the Northern Hemisphere, as represented by the International permafrost association permafrost map (Brown et al., 1998). Figure taken from Schuur et al. (2008)

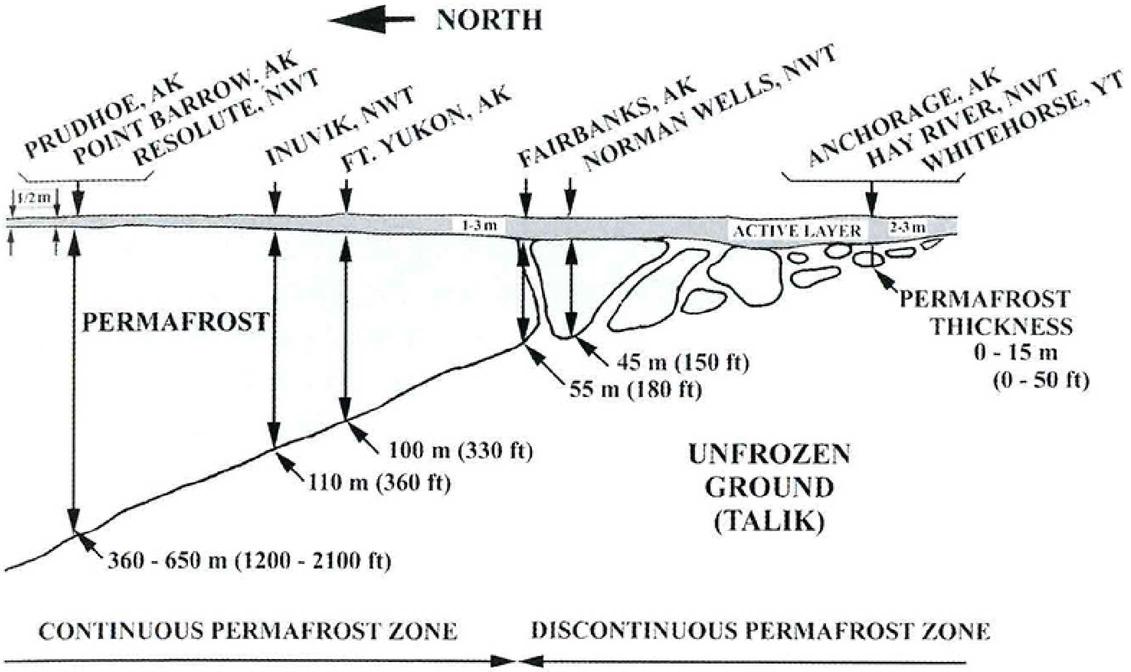


Figure 1.3: Sketch of latitudinal variations in permafrost zonation, thickness and active layer thickness in Alaska and Western Canada. Figure taken from Davis (2001).

changing to equilibrate to the present climate. While this may be a reasonable assertion for near surface layers (eg. to a depth of about 10 m or so) which have an equilibration time on the order of years to decades, given typical ground thermal characteristics, it is almost certainly invalid for very deep ground layers (eg. at depths on the order of 100 m or so) which can have equilibration times on the order of 10,000 years or more (Lunardini, 1981) - or regions where the ground has very low thermal diffusivity. Permafrost in climatic disequilibrium is known as *relict* permafrost and may include deep layers at the base of permafrost, or permafrost that is in disequilibrium in its entirely and remains because of its great depth or the presence of highly insulating near surface ground layers that restrict its rate of degradation. While it is difficult to assess the overall extent of relict permafrost, Williams and Smith (1989) and French (2007) suggest that a large amount of permafrost in the discontinuous permafrost zone may be relict, as evidenced, for example, by unfrozen *residual thaw layers* (or *taliks*) that lie between the base of the active layer and the permafrost table. It is also suggested that the majority of subsea permafrost is relict, having formed during past glacial cycles when continental shelves were exposed to low annual mean temperatures during sea level minima and were subsequently submerged when sea levels rose (Walsh, 2005).

## 1.4 Observed Changes in Permafrost and Seasonally Frozen Ground

Systematic measurements and widespread monitoring of the thermal state of permafrost began in the mid-$20^{th}$ Century at sites in Russia, Alaska and the Tibetan Plateau, while permafrost monitoring in Canada and Europe began in earnest in the 1980s and 1990s (Lemke et al. (2007)). In 1999, the Global Terrestrial Network for Permafrost (GTN-P) was established as an international monitoring program to characterize changes in permafrost thermal characteristics; the network's two prongs are the Circumpolar Active Layer Monitoring Program (CALM) and the Thermal State of Permafrost (TSP) program that focuses on borehole temperature measurements at depth (Romanovsky et al., 2010). Most recently, during the International Polar Year (IPY, 2007-2008), a major effort was undertaken to expand this monitoring network in order to capture data where there was little pre-existing monitoring sites.

Surface air temperatures are increasing over most permafrost sites and the available long-term measurements indicate a general warming trend at most permafrost sites (Walsh, 2005). Figure 1.4, taken from the IPCC's Fourth Assessment Report (Lemke et al., 2007) shows changes in ground thermal characteristics over the latter half of the $20^{th}$ Century in Russia, where some of the longest running measurements have been made. On average, the active layer thickness at permafrost sites (blue dots, top panel) has increased by approximately 20 cm (middle panel), while, outside of permafrost regions (red dots, top panel) the depth to which the ground seasonally freezes has decreased by about 34 cm, on average (bottom panel). Elsewhere, at sites along the Qinghai-Xizang Highway on the Tibetan Plateau, the active layer thickness has increased by up to 1 m, while basal thawing of permafrost has been reported at sites in Alaska and Tibet (Lemke et al., 2007).

Borehole temperature measurements paint a similar picture, as shown in Figure 1.5, produced as part of the IPY's analysis of the state of permafrost. Most borehole temperature measurements show a steady increase in temperature with the rate of temperature increase generally greater in colder sites, and lower in sites close to nominal freezing point of water (Romanovsky et al., 2010). The slower rate of temperature increase close to 0 °C is indicative of energy absorption as ice present in the ground thaws. While observations at the majority of sites point to warming of the ground, a few locations show no long-term temperature trend, or even a slight cooling trend (Lemke et al., 2007). These latter observations are not necessarily inconsistent with increased Arctic air temperatures as ground temperatures may respond not only to air temperature changes, but also to changes in factors such as vegetation cover and snow depth which may alter the ground's insulating cover (Davis, 2001). The snapshot of permafrost during the IPY also shows that mean annual ground temperatures at most sites in the discontinuous permafrost zone were greater than -2 °C, while sites in the continuous zone had temperatures ranging between -15 °C and -1 °C (Romanovsky et al., 2010). Identifying the status of permafrost in the discontinuous zone as "warm permafrost" reflects the observation of Williams and Smith (1989) that most discontinuous permafrost is either relict or in such delicate balance that climatic or other environmental changes can have drastic disequilibrium effects.

Though harder to quantify than temperature trends measured at fixed sites, observations also suggest that the total area of ground underlain by permafrost is decreasing. Degradation and upwards movement of the permafrost lower limit on the Tibetan plateau has been reported (Allison et al., 2009) and the areal extent of isolated permafrost islands along the Qingai-Xizang Highway has decreased by approximately

36 % over the past 30 years (Lemke et al., 2007). In North America, the southern boundary of discontinuous permafrost continues to move northward, as it has since the end of the Little Ice Age (Allison et al., 2009; Halsey et al., 1995). At coastal sites along the Arctic Ocean, mean annual shoreline erosion rates in permafrost regions varying between 0.7 to 3.2 m/yr have been reported in Alaska and Russia, with higher erosion rates occurring along ice-rich coasts (Lemke et al., 2007). Estimates based on air temperature measurements, also suggest that the maximum extent of seasonally frozen ground has decreased by about 7% in the NH from 1901 to 2002 (Lemke et al., 2007).

## 1.5 Mechanisms and Consequences of Permafrost Degradation

Four key mechanisms are thought to contribute to permafrost degradation, the first two of which are gradual mechanisms and the latter two, rapid or catastrophic mechanisms.

1. **Active Layer Thicknening**: Deepening of the active layer can be caused by increased summer heating, increased snow cover (which increases insulation of the ground in winter) or increased infiltration of water into the ground (which can increase the ground's thermal conductivity in summer). Thickening of the active layer is indicative of increased heat flow into the ground, which can raise the temperature of the top of permafrost above 0 °C, leading to its degradation.

2. **Talik formation**: Taliks are unfrozen layers in permafrost regions. Talik development will occur between the base of the active layer and the permafrost table when summer ground heating progresses to the point that the entire summer-thawed layer no longer re-freezes in winter. Once this has occurred, deeper permafrost degradation occurs by talik expansion. Permafrost thaw occurs both from the permafrost table downwards and from the base upwards. Once heating that was induced by active layer thickening or talik formation penetrates to the base of the permafrost, basal thaw will also begin to occur.

3. **Thermokarst development**: Many permafrost soils contain extensive amounts of ground ice, in excess of the soil's saturated liquid moisture holding capacity. When such ground ice thaws, the ground may gradually or very rapidly subside.

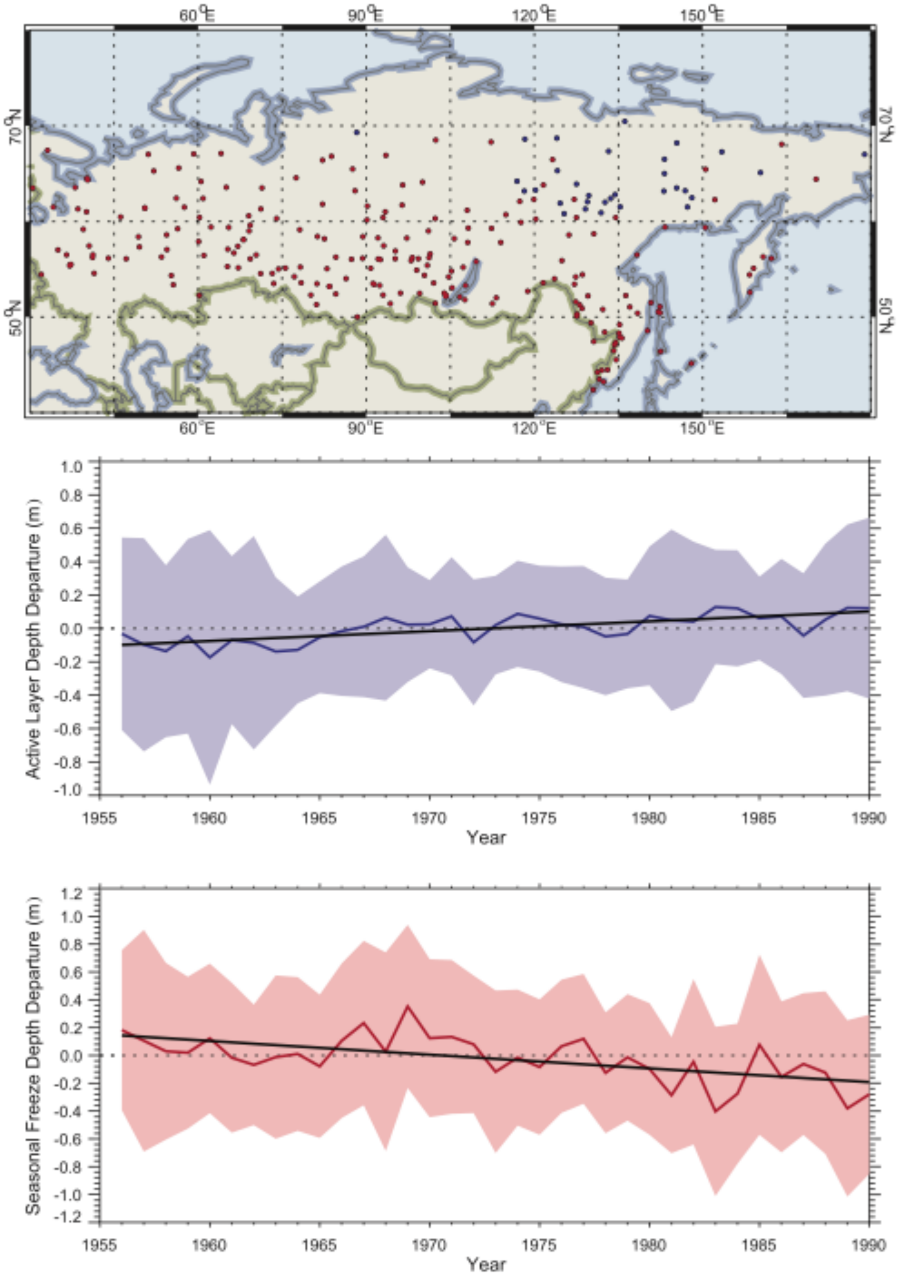


Figure 1.4: Variations in the thickness of the active layer and depth of seasonal freeze from 1956 to 1990. The top panel shows the location of permafrost sites (in blue) and seasonally frozen ground sites (red). Middle panel: changes in Active Layer Thickness. The black line represents the mean active layer change over the period, the blue line represents the mean anomaly at the sites after removing the long term mean at each site and the shaded region represents the 5-95 % confidence interval from the mean for each year. Bottom panel represents the changes in the depth of seasonal freeze with plot lines defined as in the middle panel. Figure taken from Lemke et al. (2007).

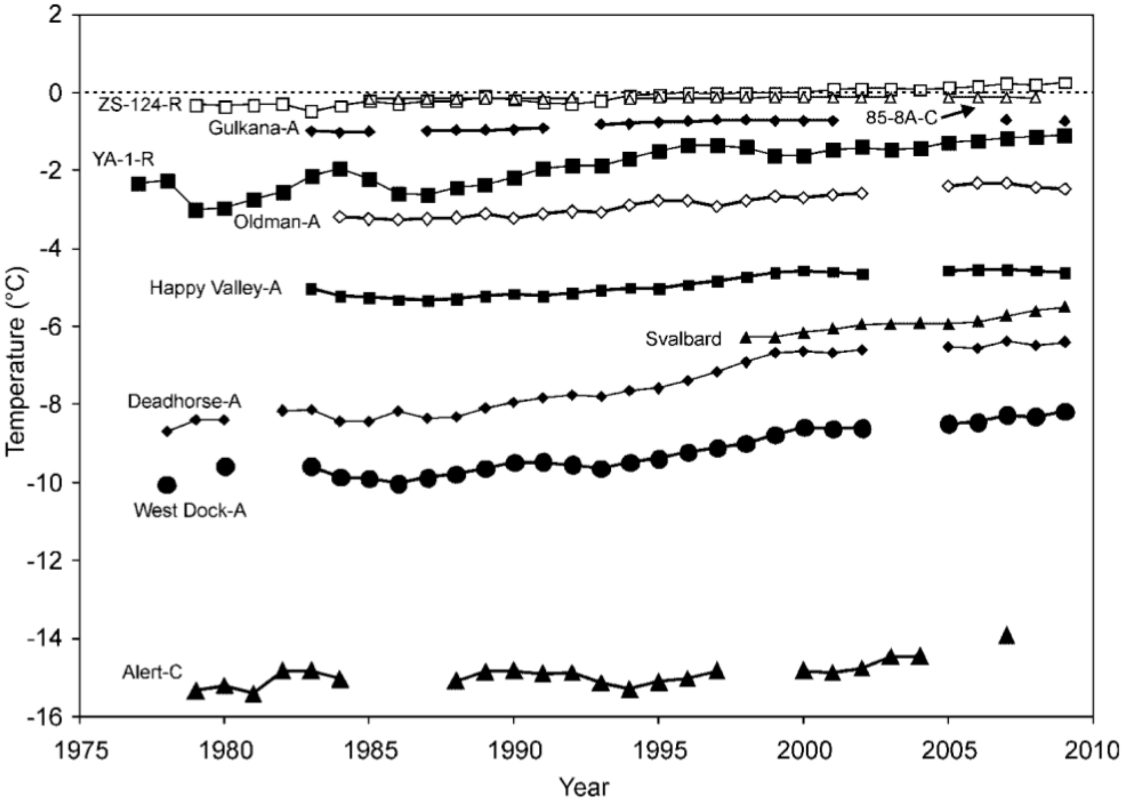


Figure 1.5: Time series showing changes in ground temperature at depths between 10 - 20 m from borehole measurements in the circumpolar permafrost region. Figure taken from Romanovsky et al. (2010).

Owing to spatial variations in thaw rates and ground ice content, the resulting terrain is uneven and called *thermokarst* by analogy with karst topography produced by differential dissolution of limestone-bearing terrain. Thermokarst development can cause particularly rapid degradation of permafrost when thaw triggers mass movements on slopes (Walsh, 2005). It has been suggested (Schuur et al., 2008) that when permafrost thaw is ongoing in a region, two major episodes of thermokarst development might occur with the first associated with degradation of Little Ice Age-aged ice and the latter with degradation of Pleistocene-aged ice.

4. **River and Coastal Erosion**: Rising sea levels, increased river discharge and increased storm activity can directly erode permafrost bearing terrain, at the same time increases sediment (and carbon) delivery to water courses (Walsh, 2005).

Each of these mechanisms might also be enhanced by other environmental changes such as fire, or human-induced land cover changes, both of which could rapidly alter the ground's thermal regime. In principle, a given permafrost site might experience degradation to varying degrees by each of these processes and it is challenging to directly rank the processes in order of importance. Nonetheless, Schuur et al. (2008) have prepared a diagram (reproduced in Figure 1.6) illustrating the hypothesized relative importance of each process acting over a large region as permafrost degradation proceeds.

Degradation of permafrost will have a number of impacts that will be felt both locally in the Arctic and in the global climate system as a whole through climate feedbacks. Briefly, the key hypothesized impacts from permafrost degradation that have been identified in the literature are as follows:

1. **Impacts on high-latitude hydrology**: Permafrost and seasonally frozen ground act as an semi-impermeable barriers to water flow through the soil, confining infiltration to the active layer while subsurface flow is restricted to taliks (French, 2007). So strong is this effect that the dominant control on hydrology at the watershed level in high-latitude regions is the presence or absence of permafrost and the thickness of the active layer (Hinzman et al., 2003). The existence of permafrost produces unique hydrological conditions. For instance, in the high Arctic, low precipitation rates present desert-like conditions, yet patchy

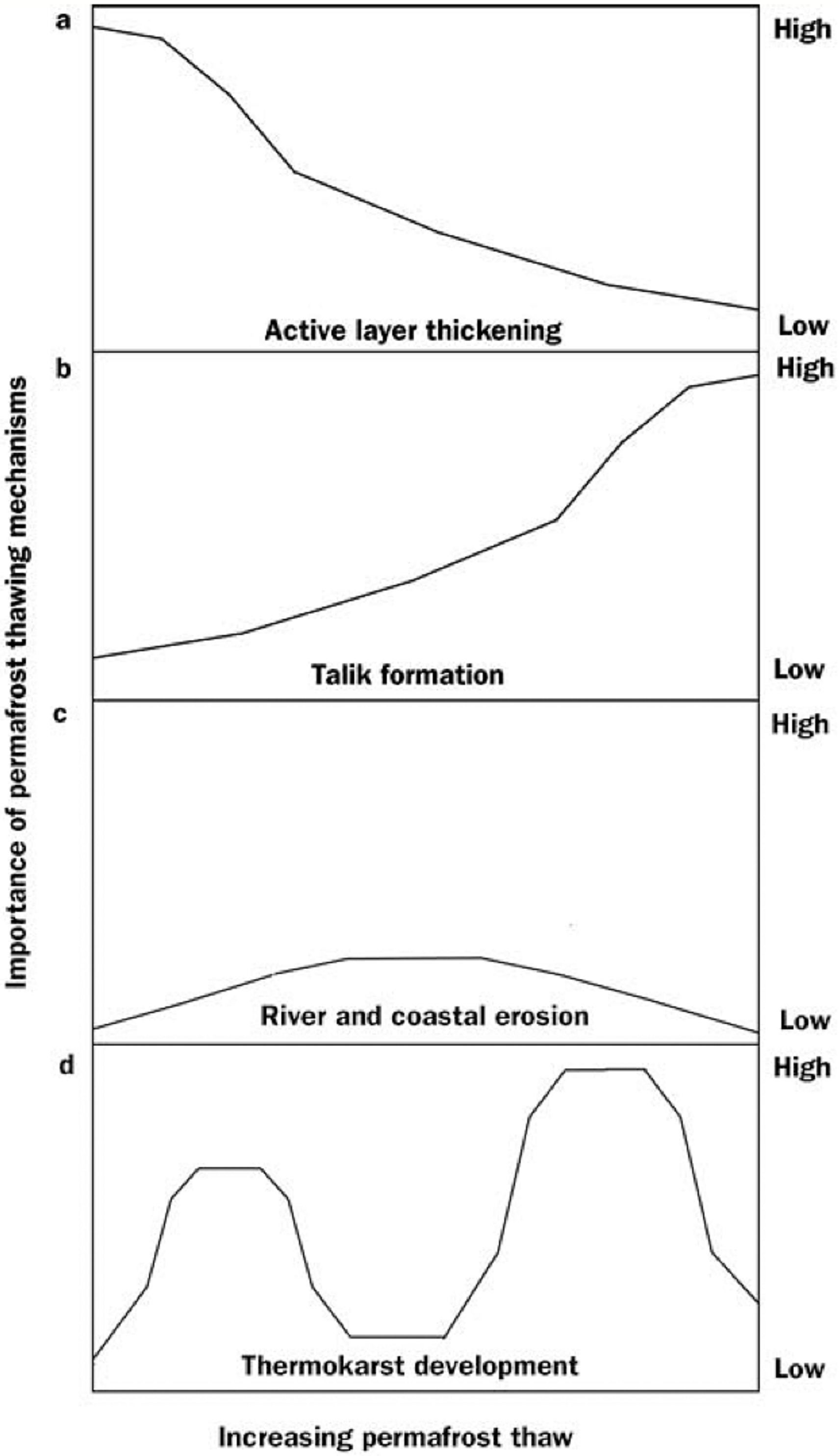


Figure 1.6: Key permafrost thaw mechanisms and their hypothesized relative importance as thaw progresses over a large region. Figure taken from Schuur et al. (2008)

wetlands are present in the summer months owing to the maintenance of a high water table associated with shallow active layers (Walsh, 2005). Conversely, in permafrost-free regions of the subarctic with permafrost-free soils, surface soils can be very dry because infiltration is not restricted (Hinzman et al., 2005). As well, a recent survey of the first-order controls of the abundance of arctic lakes found that the presence of some form of permafrost increased the lake densities and area fractions by approximately 100-170 % relative to permafrost-free terrain (Smith et al., 2007).

Permafrost thaw's effect on soil moisture may vary as thaw progresses. Initial thaw, may increase near surface soil moisture storage through deepening of the active layer (Lawrence and Slater, 2005) and thermokarst development which can lead to the formation of small water bodies such as thaw lakes which form in thermokarst depressions. As permafrost fully degrades, interactions between surface and sub-surface ground water will increase, with the nature of changes to near surface soil water depending on the hydraulic gradient between the surface and subsurface (Hinzman et al., 2005), resulting in changes in the volume of wetlands, ponds and lakes (Walsh, 2005). Satellite studies show a link between permafrost degradation and changes in Arctic lakes. From 1973-2004, lake area increased by 12 % and lake number rose by 4 % in areas of continuous permafrost; in discontinuous permafrost zones, total lake area decreased by 9 % and 13 % (Smith et al., 2005). It has been suggested that thermokarst processes may be responsible for the increase in lake area and number as colder continuous permafrost begins to thaw, while a reduction in lake area and number has occurred in discontinuous terrain in a more advanced state of thaw, likely due to drainage through taliks. Thermokarst formation may also promote the erosion of drainage channels where ice contents are high (Walsh, 2005), further altering local hydrology. Finally, it should be noted that the high-latitude soil moisture budget will also be affected by changes in precipitation and evapotranspiration, both of which are expected to increase over the $21^{st}$ Century.

2. **Impacts on Arctic biota**: Plants and animals are adapted to the Arctic environment. In permafrost regions, where the permafrost table is close to the surface, it presents a barrier to plant roots; consequently, many plants have shallow root systems and tend to grow slowly due to cold temperatures (Davis, 2001). Other characteristics of this environment such as high-latitude wet-

lands provide unique habitats for plant and animal species (Walsh, 2005). It has been suggested that projected warmer Arctic temperatures should favour taller, denser vegetation in the Arctic (ACIA, 2004) leading to northwards migration of the shrubline and the treeline, at the same time producing a further amplification of warming through a reduction of the surface albedo. Paleoecological data suggests that the treeline has co-varied with temperature over long time scales, advancing northwards during warmer periods; observations in Alaska over the $20^{th}$ Century, also point towards a general advancement of the treeline (Hinzman et al., 2005). The future distribution of vegetation will also depend on the distribution of soil moisture, which will be in part influenced by changes in permafrost. If widespread soil drying occurs in a region, this may lead to desertification, rather than a greening of the Arctic. On the other hand, vegetation productivity may be stimulated by increased atmospheric carbon dioxide concentrations and the increased availability of nutrients released from the microbial breakdown of soil organic matter.

3. **Impacts on stability and land use**: As permafrost warms and its frozen soil moisture content decreases, the ground's structural integrity will decrease, leading to a reduction in slope stability and ground bearing capacity and an increase in the rate of slow ground creep (slow deformation of the ground) (Williams and Smith, 1989). Ground subsidence and the formation of thermokarst terrain is also expected and will be especially pronounced in areas with large amounts of excess ice; areas on the order of 10s or 100s of square kilometers have been reported to be underlain by sediments containing 50 % or more by volume of excess ice (Williams and Smith (1989)) and ground subsidence may be particular extreme in such regions. Such landscape instabilities will pose a threat to infrastructure such as buildings, roads and pipelines. Changing soil permeability may also undermine the integrity of containment structures that rely on the impermeability of frozen ground; as well, land usage may shift depend on the nature of site specific soil moisture changes (Lemke et al., 2007).

4. **Climate Feedbacks**: In a comprehensive review, McGuire et al. (2006) identified a number of potential positive and negative feedbacks to the global climate system that may operate within the Arctic. A full description of the feedbacks identified by the authors is beyond the scope of this work, but a diagrammatic summary of the major terrestrial feedbacks identified is presented in Figure 1.7.

> The feedbacks involve complicated biogeochemical, ecological and hydrological interactions, many of which are only tentatively understood. Projecting the magnitude of such feedbacks is challenging, but the authors suggest that positive feedbacks will dominate on the basis of available evidence. Permafrost degradation is directly linked to positive feedbacks through increased sensible heat fluxes to the atmosphere as a result in a reduction of the ground heat sink after ground ice has thawed. As well, if permafrost leads to an overall decrease in soil moisture through drainage, then sensible heat fluxes may increase as a consequence of a reduction in latent heat fluxes when evapotranspiration rates decrease. Permafrost degradation also indirectly influences other ecosystem feedbacks through its control on the physiological and structural characteristics of vegetation and on microbial activity.

One of the most widely discussed Arctic feedbacks that is strongly influenced by permafrost degradation is the permafrost carbon feedback (PFC). Decomposition of carbon by soil microbes can occur in permafrost soils, though this is generally restricted to thin lenses of unfrozen water that exist in soils below 0 °C. As permafrost thaws, the soil's unfrozen moisture content increases dramatically, activating the soil carbon pool in these regions. Large amounts of carbon are stored in Arctic soils owing to an imbalance between photosynthesis and microbial decomposition as low temperatures decrease rates of microbial decomposition more than photosynthesis (Weller, 2005). The carbon ultimately originates from plant photosynthesis at the surface, so carbon densities tend to be greatest near the top of the soil column, but mixing associated with soil freeze-thaw processes (*cryoturbation*) can distribute the carbon throughout the soil column. Schuur et al. (2008) identify permafrost regions as containing one of the largest *vulnerable carbon pools*[3], conservatively estimating that 1672 Pg of soil carbon (over double the amount currently in the atmosphere) is present in the northern circumpolar permafrost zone, with 1024 Pg C within 3 meters of the surface. Loss of carbon from the permafrost pool will lead to a positive feedback on warming, though the intensity of the feedback will depend on the form in which carbon enters the atmosphere[4].

---

[3] Such pools are characterized by their potential to rapidly release carbon to the atmosphere under warming scenarios

[4] The feedback may be partially offset by increased carbon uptake from enhanced land and oceanic primary productivity in the Arctic

Under oxygen-rich conditions, aerobic bacteria decompose soil carbon to $CO_2$, but such organisms cannot function in oxygen-poor conditions characteristic of environments such as wetlands. Here, anaerobic bacteria produce methane and a range of other partially oxidized organic compounds (Brady and Weil, 2001). $CH_4$ is oxidized to $CO_2$ in the atmosphere through reaction with the OH radical and has an atmospheric residence time of approximately 12 years. Despite this relatively short lifespan, $CH_4$ has a radiative efficiency an order of magnitude greater than $CO_2$ on a per molecule basis and thus, a global warming potential 25 times that of $CO_2$, on a century timescale (Forster et al., 2007). The dominant natural source of methane comes from wetlands, accounting for between 20 % - 40 % of total global emissions depending on the estimate (Denman et al., 2007). Global wetland methane emissions might increase in warming scenarios due to an expansion of existing wetlands and from increased emissions from existing wetlands owing to the temperature sensitivity of methane production rates. Another methane source that may be affected by permafrost thaw is the seepage of $CH_4$ from geological reservoirs. Methane is known to accumulate in hydrocarbon reservoirs such as coal beds with permafrost and glaciers acting as a cap that restricts the loss of gas from these reservoirs to the atmosphere. A recent study by Anthony et al. (2010) points to abundant gas seeps along the margins of thawing permafrost and retreating glaciers, suggesting that this methane source may strongly increase accompanying permafrost thaw. The overall feedback arising from carbon cycle changes in regions of permafrost thaw is difficult to assess a priori as it will depend on hydrological changes (and how these impact decomposition of soil carbon into $CO_2$ and $CH_4$) and the stimulation of primary productivity from warming Arctic temperatures.

## 1.6 Permafrost in Climate Models

The characteristics and magnitudes of climate impacts and feedbacks resulting from thawing permafrost are only tentatively understood and remain under-studied. Recent major literature reviews have identified such uncertainties as an important knowledge gap that could be filled, in part, by climate models (see Allison et al. (2009); Rowland et al. (2010); Walsh (2005), for example). Permafrost models span a wide range of complexity in terms of the numerical approach used to identify permafrost and the spatial scale over which permafrost extends. As a consequence of the substantial importance of latent heat effects in permafrost regions, analytical models

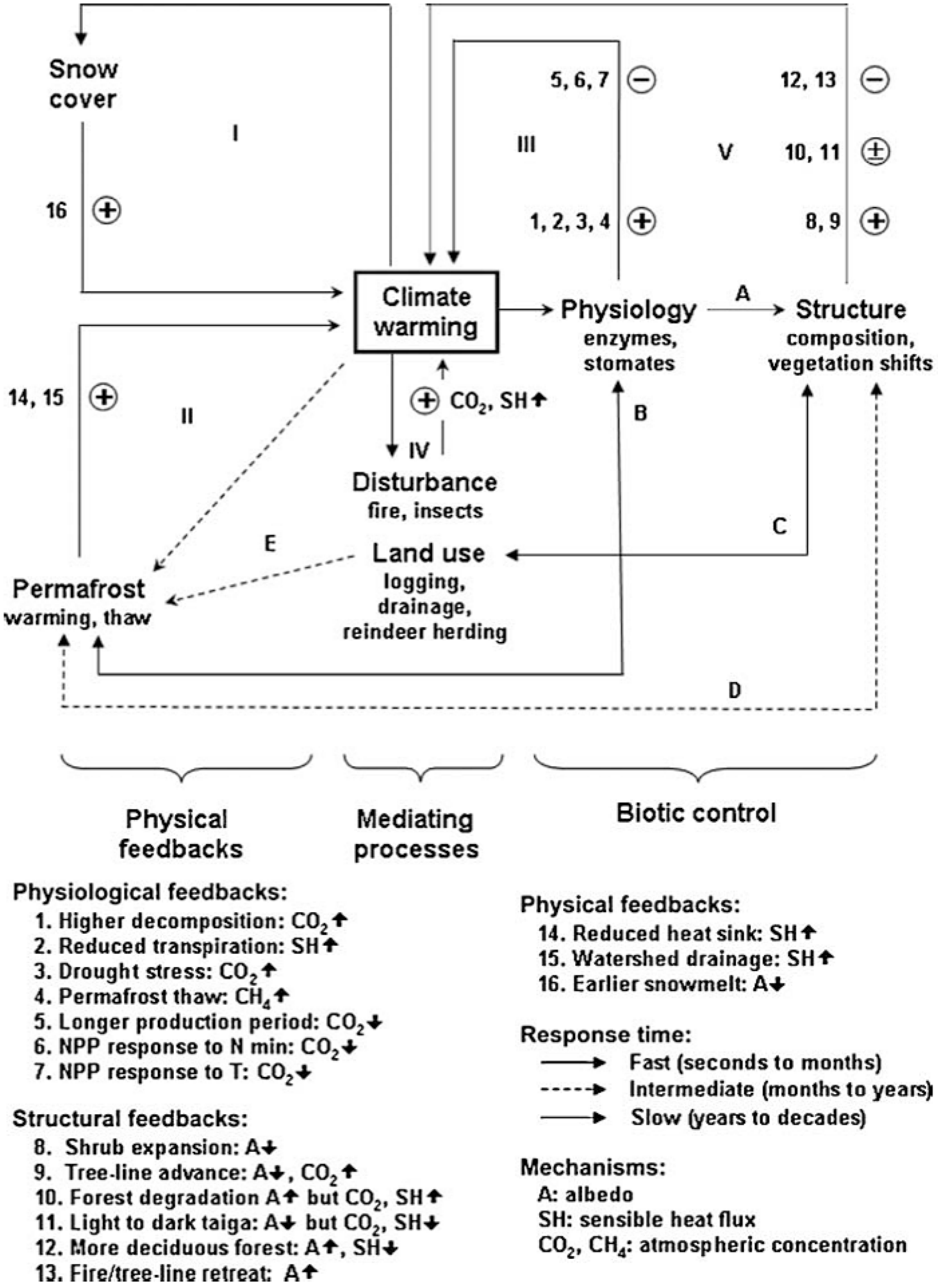


Figure 1.7: Arctic Terrestrial Ecosystem Feedbacks from McGuire et al. (2006).

based on straightforward solutions to the 1-D conductive heat flow equation cannot be used. One modelling approach involves the use of analytic equations (such as the Stefan equation or the TTOP model) that relates the presence and characteristics of permafrost to estimated ground thermal characteristics and atmospheric temperature (through freezing and thawing indices, for example) (Riseborough et al., 2008). Another approach is to relate the occurrence of permafrost to measurable climatic and topographic factors, in order to produce estimates of the probability of occurrence of permafrost over large regions, under the assumption that permafrost is in climatic equilibrium. Finally, the presence of permafrost may be directly assessed by models that solve the heat equation (accounting for the effects of freezing and thawing) numerically, accounting for energy fluxes between the ground and the atmosphere.

Recent approaches to modelling the response of permafrost to transient climate change have involved either driving stand-alone regional- or global-scale permafrost models with output from general circulation models (GCMs) or the modification of climate model land surface schemes to allow for simulation of permafrost dynamics. The former approach allows for the simulation of permafrost at high horizontal and subsurface resolutions and also for representation of an increased number of land surface processes. This approach cannot, however, represent feedbacks between the ground thermal regime and the global climate. Such feedbacks can, conceivably be captured by including a representation of permafrost within coupled climate models, but the computational requirements of such models limit the resolution and complexity of the representation of permafrost.

Since biological, hydrological and climatological aspects of the Arctic are all intricately coupled (Hinzman et al., 2005), it is doubtful that a fully accurate representation of a change in one component of the Arctic (ie. just changes in the ground thermal regime), can be accounted for by simply studying the one component in isolation. Here again, models will be a useful tool to better understand the full complexity of climate change in the Arctic.

## 1.7 Thesis Outline

Climate models are ideal tools with which to study interactions between permafrost and other aspects of the Arctic environment, but the representation of frozen ground in coupled climate models is in its relative infancy. A number of modelling groups are working towards including and improving a representation of permafrost within their models including the National Centre for Atmospheric Research (NCAR) with the Community Land Model (CLM; Lawrence and Slater (2005); Lawrence et al. (2011)), the United Kingdom Met Office with the Joint UK Land Environment Simulator (Dankers et al., 2011), the Laboratoire des Sciences du Climat et de l'Environment with the ORganizing Carbon and Hydrology in Dyanic EcosystEms (ORCHIDEE) mode (Koven et al., 2011). In my thesis, I report on modifications made to the UVic Earth System Climate Model (UVic ESCM), an intermediate complexity climate model, to allow it to achieve a representation of permafrost and seasonally frozen ground.

Three major issues arise when attempting to accurately simulate permafrost at large scales in such models. Firstly, as mentioned above, the need for a computationally efficient model necessarily limits the spatial resolution of the model, potentially reducing the accuracy of the simulation. A related issue is the high degree of spatial heterogeneity of land surface characteristics that occurs at very small scales that is challenging to represent in a coarse resolution model. Finally, data on ground characteristics such as soil characteristics and depth to bedrock that is necessary input data to models is limited at such large scales, necessitating simplifying assumptions. In my thesis, I attempt to assess how these factors might affect the simulated permafrost distribution in the model.

The aim of my thesis is to address the following questions:

1. How do changes to the land surface scheme of the UVic ESCM, imposed to allow the model to simulate frozen ground, affect the ground climate of the model?
2. How well does the UVic ESCM simulate regions of permafrost and seasonally frozen ground in the present climate?
3. How does the present day distribution of permafrost in the model change when land surface and ground characteristics, and model vertical resolution are altered?
4. What is the projected rate of permafrost degradation following climate change scenarios and how do rates of degradation depend on variation of land and ground model characteristics?
5. What is the nature and extent of major environmental changes in the Arctic accompanying permafrost degradation as assessed by the ESCM?

Briefly, the structure of my thesis is as follows:

- **Chapter 2** presents a description of the standard configuration of version 3.0 of the UVic ESCM (UVic 3.0) which was the starting point for my model development work. I describe the characteristics and physics of the MOSES land surface scheme and TRIFFID dynamic vegetation model as implemented in the ESCM and outline the major changes that I made to MOSES to enable the ESCM to achieve a realistic representation of soil freeze/thaw processes. As one of the major modifications made to MOSES was the reintroduction of multiple soil layers (versus the single soil layer in the original MOSES implementation), the new version of the ESCM is designated UVic 3.0M. A secondary version of UVic 3.0M (UVic 3.0E) is also introduced wherein the land surface scheme is decoupled from the other model components and forced by ERA-40 reanalysis data.

- **Chapter 3** focuses on the simulation of permafrost and seasonally frozen ground in the latter half of the $20^{th}$ Century. I compare the representation of these regions in UVic 3.0E and UVic 3.0M to assess how the atmospheric model's regional temperature and precipitation biases might affect the simulated representation of permafrost. I also conduct a suite of sensitivity experiments using UVic 3.0E to understand how the representation of permafrost depends on land surface and ground characteristics.

- **Chapter 4** examines the future response of the terrestrial cryosphere using UVic 3.0M. I identify a range of possible cryosphere responses, following a suite of four Representative Concentration Pathways that are designed to span a likely range of anthropogenic radiative forcing pathways. I also test the sensitivity of the rate of permafrost degradation to different configurations of the land surface scheme.

- **Chapter 5** introduces a simple wetland identification scheme that is capable of representing wetlands at a global scale based on top-layer soil moisture characteristics. I examine the response of high-latitude wetlands in the Northern Hemisphere in UVic 3.0M to permafrost degradation following the four Representative Concentration Pathways.

- **Chapter 6** summarizes the key findings of my thesis research. I also discuss future research avenues suggested by my research.

# Chapter 2

# Model Description

This work is based on version 3.0 of the UVic Earth System Climate Model (UVic ESCM 3.0), an intermediate complexity climate model with a horizontal resolution of 3.6° x 1.8° [1]. I begin with a brief description of the land surface scheme in UVic 3.0 before describing the modified version used in my thesis work. The fully coupled model is described in detail in Weaver et al. (2001) and Meissner et al. (2003). It consists of a 3-dimensional ocean general circulation model coupled to a reduced complexity energy-moisture balance atmospheric model, a dynamic-thermodynamic sea ice model, a land surface scheme and a dynamic vegetation model. It also includes representations of the terrestrial and oceanic carbon cycles.

The land surface scheme (LSS) is a simplified single soil layer version of the Hadley Centre's Met Office Surface Exchange Scheme (MOSES) (Cox et al., 1999). Vegetation in MOSES is updated through a dynamic vegetation model: Top-Down Representation of Interactive Foliage and Flora Including Dynamics (TRIFFID) which simulates competition between five plant functional types (PFTs): Broadleaf Trees, Needleleaf Trees, $C_3$ and $C_4$ Grasses and Shrubs.

## 2.1 Land Surface Scheme - UVic 3.0

MOSES is a tiled land surface scheme in which each model grid cell consists of a mosaic of surface types or *tiles*. As implemented in the ESCM, the tiles consist of the five TRIFFID PFTs as well as bare soil. Fluxes of energy, water and carbon between the atmosphere and the vegetation/soil are calculated separately for each tile and then

---

[1] Zonal x meridional

aggregated over the grid cell based on the fractional coverage of each tile[2]. Each grid cell contains a single 1 m thick soil layer which is un-tiled such that soil temperature, moisture and carbon content are updated based on the aggregated fluxes from the overlying surface cover tiles. Soil thermal conductivity and heat capacity are constant and do not depend on soil moisture content.

### The Surface Energy Budget

The surface energy budget on each tile is governed by three separate temperatures: the surface air temperature ($\mathrm{T}_a$), the soil (ground) temperature, ($\mathrm{T}_g$) and the surface or *skin* temperature ($\mathbf{T}^*$). $\mathrm{T}_a$ and $\mathrm{T}_g$ are the same for each tile within a grid cell. $\mathbf{T^*}$ is defined as the temperature that balances the surface energy budget, which depends on fluxes of radiation ($\mathbf{R}_n$), sensible heat (**SH**) and latent heat (**LH**) between the surface and atmosphere and the heat flux between surface and ground (**GH**):

$$\boldsymbol{R_n} - \boldsymbol{SH} - \boldsymbol{LH} - \boldsymbol{GH} = 0 \tag{2.1}$$

### Net Radiation

The net radiation, $\mathbf{R}_n$, is the difference between incoming radiation: shortwave radiation absorbed by the surface, $\mathrm{SW}_{\downarrow,abs}$ (calculated in the atmospheric module, using LSS surface albedo) plus longwave radiation from the atmosphere, $\mathrm{LW}_{\downarrow}$ (assumed to be completely absorbed by the surface) and outgoing longwave radiation emitted by the surface. The surface is treated as a blackbody, emitting at the skin temperature, $\mathbf{T}^*$.

$$\boldsymbol{R_n} = SW_{\downarrow,abs} + LW_{\downarrow} - \sigma\boldsymbol{T^{*4}} \tag{2.2}$$

### Sensible Heat Flux

The sensible heat flux between the atmosphere and the land surface depends on the gradient in thermal energy between near surface air (at temperature $\mathrm{T}_a$) and the air immediately above the surface (at $\mathbf{T}^*$) as well as on an aerodynamic resistance, $\mathbf{r}_a$:

---

[2] In this document, tile-specific quantities are represented in **bold**. Gridcell aggregated quantities and variables that are constant across a grid cell are represented using normal type. Constants and variables are defined in Appendix A and B

$$\boldsymbol{SH} = \frac{\rho_a c_p(\boldsymbol{T^*} - T_a)}{\boldsymbol{r_a}} \tag{2.3}$$

Here, $\rho_a$ is the density of air and is given by the ideal gas law:

$$\rho_a = \frac{P}{RT_a} \tag{2.4}$$

The aerodynamic resistance decreases as wind speed increases and as vegetation height in a tile increases, reflecting the increased turbulent fluxes under such conditions.

### Latent Heat Flux

Latent heat fluxes cool the land surface as moisture evaporates, transpires or sublimates from the land. Rates of evapotranspiration (**E**) are calculated following the Penman-Monteith formulation. **E** is governed firstly by the air's humidity deficit - the degree to which it is unsaturated. This is determined as the difference between the air's specific humidity and its specific humidity at saturation. Moisture transport occurs at the interface between the air and the soil/vegetation, so the saturated specific humidity ($\mathbf{q}_{sat}$) is diagnosed at the skin temperature. So long as this saturated specific humidity is greater than the actual specific humidity in the air ($\mathrm{q}_a$, diagnosed in the atmosphere model), **E** will occur at a rate proportional to the difference between these two quantities.

Secondly, like surface-to-air sensible heat fluxes, latent heat fluxes are influenced by the degree of turbulent mixing in the atmosphere and hence the aerodynamic resistance, $\mathbf{r}_a$. Since vegetation and soil resist the loss of moisture, a second canopy / soil resistance factor must also be included (the *surface resistance*, $\mathbf{r}_s$). As these resistances operate in series, the net resistance is the sum of the aerodynamic and surface resistances. The rate of **E** is then given by:

$$\boldsymbol{E} = \frac{\rho_{air}(\boldsymbol{q_{sat}}(\boldsymbol{T^*}) - q_a)}{\boldsymbol{r_a} + \boldsymbol{r_s}} \tag{2.5}$$

The latent heat flux associated with the moisture flux is then:

$$\boldsymbol{LH} = \frac{\rho_{air}(\boldsymbol{q_{sat}}(\boldsymbol{T^*}) - q_a)(L_c)}{\boldsymbol{r_a} + \boldsymbol{r_s}} \tag{2.6}$$

where $\mathrm{L}_c$ is the latent heat of vaporization. The surface resistance over bare soil tiles has a relatively simple formulation, depending on a soil moisture availability factor, $\beta$:

$$\boldsymbol{r_{s,soil}} = min\left(10^6, \frac{100}{\beta}\right) \tag{2.7}$$

$\beta$ is a measurement of soil moisture availability to plants and depends on the volumetric liquid soil moisture concentration[3], $\theta_l$ relative to the concentration at the critical ($\theta_c$) and wilting ($\theta_w$) points. $\beta$ is calculated as:

$$\beta(\theta_l) = \begin{cases} 0 & \theta_l \leq \theta_w \\ \frac{\theta_l - \theta_w}{\theta_c - \theta_w} & \theta_w < \theta_l < \theta_c \\ 1 & \theta_c \leq \theta_l \end{cases} \tag{2.8}$$

Transpiration occurs through the stomata of plants. The size of stomatal openings (and, hence, the surface resistance on PFT tiles) depends on the rate of net photosynthesis within leaves which, in turn, depends on the humidity deficit, skin temperature, atmospheric $CO_2$ concentration and the soil moisture availability factor. The calculation of $\mathbf{r}_{s,PFT}$ is described in detail in Cox et al. (1999).

### Ground Heat Flux

The ground heat flux is modelled via Fourier's law, with the flux directly proportional to the temperature gradient between the surface and the soil through the thermal conductivity of the soil ($\lambda_{soil}$). With $\mathbf{T^*}$ and $T_g$ separated by a distance, $\Delta z/2$ (half the thickness of the soil layer), the ground heat flux is given by:

$$\boldsymbol{GH} = \lambda_{soil} \frac{\boldsymbol{T}^* - T_g}{\frac{\Delta z}{2}} \tag{2.9}$$

The surface energy budget equation can then be expressed in terms of T*:

$$SW_{\downarrow,abs} + LW_{\downarrow} - \sigma \boldsymbol{T^{*4}} - \frac{\rho_a c_p(\boldsymbol{T}^* - T_a)}{\boldsymbol{r_a}} - \frac{\rho_a(\boldsymbol{q_{sat}}(\boldsymbol{T*}) - q_a)(L_c)}{\boldsymbol{r_a} + \boldsymbol{r_c}} - \lambda_{soil} \frac{\boldsymbol{T}^* - T_g}{\frac{\Delta z}{2}} = 0 \tag{2.10}$$

A Newton-Raphson method is used to determine the component fluxes by solving

---

[3] As UVic 3.0 allows moisture to exist only in its liquid form, here the volumetric concentration of liquid moisture is the same as the volumetric concentration of total moisture - $\theta_l = \theta$; refer to Appendix A for a discussion of moisture units used in this thesis

for $\mathbf{T}^*$, using $\mathbf{T}^* = \mathrm{T}_a$ as an initial guess. Since this method generally does not produce an exact solution, any imbalance in the energy budget is added to ground heat flux.

### 2.1.1 Snow Processes

When snow is present in a grid cell, it is assumed to be evenly distributed with the same areal density / depth within each tile. The presence of snow impacts the land surface scheme in three ways. Firstly, snow increases the grid cell albedo. Secondly, when lying snow is present, moisture fluxes occur via sublimation ($\mathbf{S}_{snow}$), with moisture withdrawn from the snow pack. The latent heat term in the surface energy budget is then modified to reflect this change:

$$\boldsymbol{LH} = \frac{(\rho_a(\boldsymbol{q_{sat}}(\boldsymbol{T^*}) - q_a)(L_c + L_f)}{\boldsymbol{r_a} + \boldsymbol{r_c}} \tag{2.11}$$

where $\mathrm{L}_f$ is the latent heat of fusion.

Finally, the presence of snow results in the absorption of energy when snowmelt occurs. Snowmelt occurs at the base of the snow pack; when the temperature of the soil layer rises above the freezing point, it is reset to 0 °C and the excess energy is used to melt snow.

### 2.1.2 Soil Moisture and Temperature

In UVic ESCM 3.0, soil moisture remains liquid at all temperatures. Soil moisture ($\mathrm{M}_u$)[4] is augmented by rainfall (RF) and snowmelt (SM), and is diminished due to evapotranspiration and the water flux through the bottom of the soil layer (W). Thus:

$$\frac{dM_u}{dt} = RF + SM - E - W \tag{2.12}$$

The bottom boundary condition for moisture is that of free drainage where moisture is influenced only by the force of gravity, with drainage equal to the hydraulic conductivity, K (i.e. W = K). The hydraulic conductivity varies according to the liquid moisture content relative to the saturation, following the work of Clapp and Hornberger (1978).

---

[4] Here, $\mathrm{M}_u$ is the total gridcell liquid soil moisture content as opposed to the volumetric liquid soil moisture content introduced earlier. Soil moisture units are clarified in Appendix A

$$K = K_s \left(\frac{\theta_l}{\theta_s}\right)^{2b+3} \tag{2.13}$$

This function may be qualitatively explained as follows: water moves through unsaturated soil through small water-filled pores, avoiding larger spaces that are filled with air (Hornberger et al., 1998). As the liquid moisture content of the soil increases so too does the number of water filled pores. This provides a larger surface area for drainage and a higher conductivity.

Soil temperature is updated via the ground heat flux, or as a result of snow melt. The bottom boundary condition for temperature is that of zero heat flux.

### 2.1.3 Vegetation Dynamics

The biosphere, as modelled by TRIFFID consists of soil carbon as well as tiles of the five PFTs, each occupying a certain fraction of the gridcell ($\mathbf{f}$) with vegetation carbon density $\boldsymbol{\mathcal{C}}$ partitioned into stocks of roots, leaves and woody tissues based on PFT-specific allometric constraints. The physical state of vegetation is characterized by its height (itself a function of woody tissue carbon content) and its leaf area index (LAI).

A photosynthesis-stomatal conductance model (described in Cox et al. (1999)) determines the net primary productivity(NPP, $\mathbf{N}$) on each tile, a portion of which ($\boldsymbol{\Lambda}$) is used to increase the areal fraction of vegetation and the remainder of which is used to update the vegetation carbon density. $\boldsymbol{\Lambda}$ is a function of vegetation LAI such that at low LAI, most of the NPP is used for vegetation growth (increasing $\boldsymbol{\mathcal{C}}$), while at high LAIs, most of the NPP is used for expansion (increasing $\mathbf{f}$). Local litterfall ($\boldsymbol{\Gamma_{local}}$) from roots, leaves and woody tissue reduces $\boldsymbol{\mathcal{C}}$. The vegetation carbon density of a given PFT is then updated as follows:

$$\frac{d\boldsymbol{\mathcal{C}}}{dt} = (1 - \boldsymbol{\Lambda})\boldsymbol{N} - \boldsymbol{\Gamma_{local}} \tag{2.14}$$

The vegetation areal coverage is updated per the following equation:

$$\boldsymbol{\mathcal{C}}\frac{d\boldsymbol{f}}{dt} = \boldsymbol{\Lambda N f \alpha} - \boldsymbol{\Gamma_{largescale}} \tag{2.15}$$

$\boldsymbol{\alpha}$ is a factor that accounts for the fact that vegetation expansion is limited based on competition from other PFTs in a gridcell. Competition follows a dominance hierarchy wherein shrubs dominate over grasses and trees dominate over both shrubs and grasses. To say that one PFT dominates over another implies that the dominated PFT is unable to expand into fractions of the gridcell occupied by the dominant PFT. The two tree species co-compete with the outcome of competition determined by a factor that depends on the relative height of the broadleaf and needleleaf trees such that the taller PFT outcompetes the shorter PFT. $C_3$ and $C_4$ grasses similarly co-compete. The portion of the NPP allocated for expansion that does not result in an increase in $\mathbf{f}$ owing to competition from other PFTs (1-$\boldsymbol{\alpha}$) acts as an additional source of litter fall. A third source of litter fall comes from large scale disturbances such as forest fires. The latter two litterfall terms for a source of litterfall from large scale disturbances ($\boldsymbol{\Gamma_{largescale}}$).

Soil carbon ($\mathcal{C}_s$) is increased based on the gridcell aggregated litterfall rate ($\Gamma$), comprising the sum of the three PFT-specific litterfall components outlined above and is reduced by soil respiration:

$$\frac{d\mathcal{C}_s}{dt} = \Gamma - R_{soil} \tag{2.16}$$

The soil respiration rate depends on the ground temperature (through a function, f($T$)), liquid moisture content (f($\theta_l$)) and the soil carbon content:

$$R_{soil} = \kappa_s \mathcal{C}_s f_{\theta_l} f_T \tag{2.17}$$

The moisture and temperature functions have the following form:

$$f(T) = q_{10}^{0.1(T_g - 298K)} \tag{2.18}$$

$$f(\theta_l) = \begin{cases} 1 - 0.8\left(\frac{\theta_l - 0.5(\theta_s+\theta_w)}{\theta_s}\right) & \theta_l > 0.5(\theta_s+\theta_w) \\ 0.2 + 0.8\left(\frac{\theta_l - \theta_w}{0.5(\theta_s+\theta_w)-\theta_w}\right) & \theta_w < \theta_l <= 0.5(\theta_s+\theta_w) \\ 0.2 & \theta_l <= \theta_w \end{cases} \tag{2.19}$$

Equation 2.18 reflects the observation that warmer soil temperatures increase the rate of microbial activity and hence the rate of soil carbon decomposition. $q_{10}$ is a factor that accounts of the rate of respiration increase for each 10 K increase in soil temperature. It is here approximated as a constant, 2, which corresponds to its typical value in the range of temperatures from 5 °C - 20 °C (McGuire et al., 1992). The moisture function represents the fact that respiration is limited in dry soils due to a lack of moisture as well as in very wet soils which lack sufficient oxygen to support microbial activity (Bonan, 2008).

## 2.2 Land Surface Scheme - UVic 3.0M

In order for a realistic representation of permafrost and seasonal frost to be included in the UVic ESCM, a new version of the model with a revised land surface scheme was developed. Revisions to the LSS included the addition of multiple sub-surface layers, the implementation of new soil characteristics, an alteration of soil thermal characteristics based on soil moisture content, a representation of soil-freeze thaw processes, and a revision of the surface energy budget. The revised version of the model will be referred to as UVic 3.0M (for *multilayer*). A description of each of the key modifications to the ESCM now follows. As much of the development work consisted of reversing simplifications made to the MOSES code when it was originally implemented in the UVic ESCM, some of the description in this section follows Cox et al. (1999).

### Subsurface layers

Alexeev et al. (2007) argue that to correctly evaluate the transient response of permafrost on centennial timescales, the subsurface depth of a LSS should extend to $\sim$ 200 - 300 m. To this end, the total thickness of subsurface layers in MOSES has been increased from 1 m to 250 m with the number of layers increased from 1 to 14. Layer thicknesses increase exponentially from 0.1 m in the top layer to 104 m in the bottom layer. The number of layers and their thicknesses were chosen somewhat pragmatically. It was found that decreasing the thickness of the top soil layer much below 0.1 m lead to numerical instabilities in soil temperature when running the model with its standard configuration with a 1-hour timestep for the land surface scheme. The total number of layers was set at 14 to avoid excessive additional computational cost in the model.

The top eight layers (to a depth of 10 m) are soil; given the large uncertainties on the depth-to-bedrock at a global scale (Richter and Markewitz, 1995), it is felt that this is a reasonable order-of-magnitude value for soil depth. The bottom six layers are bedrock. These layers are taken to be hydrologically inactive, meaning that moisture that drains from the bottommost soil layer is immediately transferred to the river runoff scheme. Soil and bedrock layers are able to exchange heat fluxes, however.

### Soil Characteristics

MOSES allows for three different classes of soils based on texture with a gridcell's soil type determined as fine, medium or coarse-grained based on the soil texture database of Wilson and Henderson-Sellers (1985). The relative abundance of sand, silt and clay sized particles differs amongst the different textural classes and soil thermal and hydrological characteristics in MOSES were determined from empirical functions of soil particle abundance. UVic 3.0 soil characteristics are those corresponding to the medium soil textural class in MOSES. Recently, Dharssi et al. (2009) identified an error in how the empirical functions used to determine MOSES' soil hydrological characteristics were applied[5]. To address this error, and to introduce some heterogeneity into the soil distribution in the model, a grain size distribution for each land grid cell was determined by interpolating data from the $1^o$ x $1^o$ International Satellite Land Surface Climatology Project, Initiative II (ISLSCP-II, (Scholes and Brown de Colstoun, 2011)) soil dataset to the UVic model grid. These grain size distributions are used to determine the hydrological and thermal parameters of mineral soils in the model. Mineral hydrological parameters are determined following the work of Cosby et al. (1984):

$$b_m = 0.157\%_{Clay} - 0.003\%_{Sand} + 3.10 \tag{2.20}$$

$$\Psi_{s,m} = 10^{-0.0095\%_{Sand}+0.0063\%_{Silt}-0.46} \tag{2.21}$$

$$K_{s,m} = 10^{0.0126\%_{Sand}-0.0064\%_{Clay}-2.75} \tag{2.22}$$

$$\theta_{s,m} = -0.00142\%_{Sand} - 0.00037\%_{Clay} + 0.505 \tag{2.23}$$

where $b_m$ is the Clapp-Hornberger coefficient, $\Psi_{s,m}$ is the saturated soil matric potential (or, suction), $K_{s,m}$ is the hydraulic conductivity of saturated soil and $\theta_{s,m}$ is the porosity (ie. the volumetric soil moisture concentration at saturation). Soil volumetric heat capacity for dry soil ($C_{dry,m}$) is given by Farouki (1981), while the dry soil thermal ($\lambda_{dry,m}$) conductivity is based on Johansen (1975):

---

[5] The error appears to be that an exponential term in equations 2.21 and 2.22 had base $e$ rather than base 10

$$C_{dry,m} = \left( \frac{2.128\%_{Sand} + 2.385\%_{Clay}}{\%_{Sand} + \%_{Clay}} \right) (1 - \theta_{s,m}) \cdot 10^6 \tag{2.24}$$

$$\lambda_{dry,m} = \frac{364.5(1 - \theta_{s,m}) + 64.7}{2700 - 2556.9(1 - \theta_{s,m})} \tag{2.25}$$

Table 2.1 compares the mineral soil characteristics in UVic 3.0M to UVic 3.0.

| **Parameter** | **Symbol** | **UVic 3.0 value** | **Range of UVic 3.0M Values** |
|---|---|---|---|
| Clapp-Hornberger exponent | $b_m$ | 6.63 | 2.8-11.4 |
| Sat. soil matric potential | $\Psi_{s,m}$ | 0.0453 m | 0.0467-0.6068 m |
| Sat. hydraulic conductivity | $K_{s,m}$ | 0.0005 kg $m^{-2}s^{-1}$* | 0.00149-0.02790 kg $m^{-2}s^{-1}$ |
| Porosity | $\theta_{s,m}$ | 0.458 $m^3m^{-3}$ | 0.370-0.475 $m^3m^{-3}$ |
| Dry volumetric heat capacity | $C_{dry,m}$ | $0.33 \cdot 10^6$ J $m^{-3}K^{-1}$ * | 0.91-1.47·1 $^6$ J $m^{-3}K^{-1}$ |
| Dry thermal conductivity | $\lambda_{dry,m}$ | 0.75 W $m^{-1}K^{-1}$* | 0.187-0.270 W $m^{-1}K^{-1}$ |

Table 2.1: Mineral soil hydrological and thermal parameters in UVic 3.0 and UVic 3.0M. *Denotes parameters in UVic 3.0 that were adjusted from their original MOSES values. MOSES values for these parameters are as follows: $K_{s,m} = 0.00472$ kg $m^{-2}s^{-1}$, $c_{dry,m} = 1.19 \cdot 10^5$ J $m^{-3}K^{-1}$, $\lambda_{dry,m} = 0.227$ W $m^{-1}K^{-1}$

| **Parameter** | **Symbol** | **Value for organic soils** |
|---|---|---|
| Clapp-Hornberger exponent | $b_o$ | 2.7 |
| Saturated soil matric potential | $\Psi_{s,o}$ | 0.0103 m |
| Saturated hydraulic conductivity | $K_{s,o}$ | 0.0200 kg $m^{-2}s^{-1}$ |
| Porosity | $\theta_{s,o}$ | 0.900 $m^3$ $m^{-3}$ |
| Dry volumetric heat capacity | $C_{dry,o}$ | $0.25 \cdot 10^6$ J $m^{-3}K^{-1}$ |
| Dry thermal conductivity | $\lambda_{dry,o}$ | 0.05 W $m^{-1}K^{-1}$ |

Table 2.2: Organic soil thermal and hydrological parameters in UVic 3.0M. Values are taken from Lawrence and Slater (2008), with the exception of the value for hydraulic conductivity which was adjusted by the author.

### Organic soil

The organic component of soil has significantly different thermal and hydrological characteristics as compared to typical mineral soils (see Table 2.2). In particular, organic-rich soils tend to have higher porosity and hydraulic conductivity and a lower thermal conductivity relative to mineral soils. These characteristics are particularly important in many permafrost regions where organic material in soil is often very abundant. As such, a representation of organic soil was included in UVic 3.0M to capture these important characteristics.

Inclusion of organic soil in the model requires some measurement of soil carbon content. Soil carbon areal densities ($\sigma_c$ - kg C/m$^2$) from the ISLSCP-II soils dataset (Scholes and Brown de Colstoun, 2011) were used in this work as opposed to the model's simulated soil carbon pool. This was done as it was found that, despite the fact the the model's global vegetation carbon stocks agree well with observations, UVic 3.0 had a tendency to underestimate soil carbon content relative to observations, with the tendency particularly pronounced in the Northern Hemisphere mid-to-high latitudes. It is noted as well that specifying soil carbon contents from observations as opposed to based on simulated soil carbon stocks means that feedbacks associated with interactions between soil carbon contents and soil thermal and hydrological characterstics are not represented. The approach to representing organic soils taken here follows that of Lawrence and Slater (2008) with only minor modifications. Briefly, this approach is as follows:

1. Areal carbon densities ($\sigma_c$, (kg C m$^{-2}$)) were obtained from the ISLSCP-II soils dataset and interpolated to the UVic model grid (Figure 2.1). These data give a representation of the total soil carbon content within the upper 1.5 m of the soil; the dataset has no information on soil carbon below these depths, so it is assumed that soils deeper than 1.5 m are strictly mineral in composition.

2. The carbon density in each model layer ($\rho_c$, kg C m$^{-3}$) in each soil layer was determined. Zinke et al. (1986) produced cumulative carbon profiles (CCPs) for tropical/temperate and boreal/polar soils which Lawrence and Slater (2008) used in their work. As these distinct CCPs (Figure 2.2) are relatively similar, for the purpose of this work, the profiles were averaged and the averaged profile was used to partition the soil carbon into different soil layers. The carbon density ($\rho_c$) in a given layer is then determined from the gridcell areal carbon density and the values of the cumulative carbon profile at the top and bottom of the layer:

$$\rho_c = (CCP(z_{bottom}) - CCP(z_{top})) \cdot \sigma_c \tag{2.26}$$

3. The carbon density in each layer is restricted so that it never exceeds the density

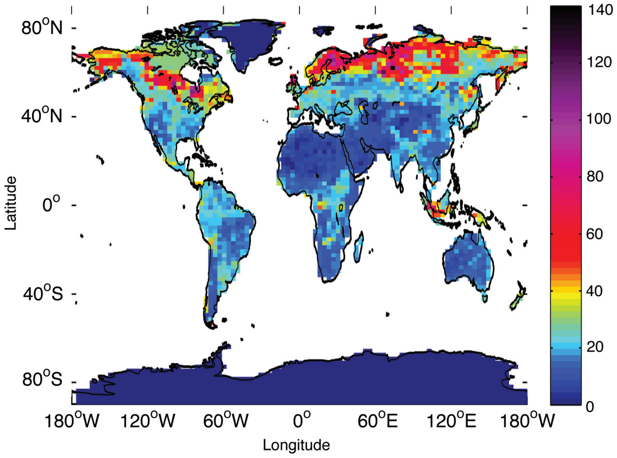


Figure 2.1: Map of the areal density of soil carbon. Data were taken from the ISLSCP-II (Scholes and Brown de Colstoun, 2011) soils dataset and interpolated to the UVic model grid. The data represent the density of carbon (kg C $m^{-2}$) based on carbon content in the top 1.5 m of soil.

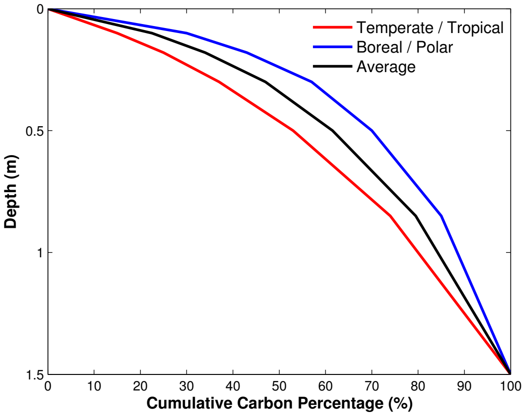


Figure 2.2: Cumulative carbon distribution functions for temperate/tropical and boreal/polar soils based on the work of Zinke et al. (1986). Cumulative carbon profiles represent the percentage of the total soil carbon stock that lies above a given depth. The black curve represents an average of the other curves and is used as the single, global cumulative carbon profile in this work.

of peat ($\rho_p = 130$ kg $m^{-3}$). If the carbon density in a particular layer exceeds this value, the excess carbon is added to the next layer down. In this way, the areal carbon density of a particular cell is conserved

4. The fraction of organic soil ($f_o$) and mineral soil ($f_m$) in a given layer is defined based on the density of carbon in the layer relative to peat:

$$f_o = \frac{\rho_c}{\rho_p} \qquad f_m = 1 - \frac{\rho_c}{\rho_p} \tag{2.27}$$

5. The overall thermal and hydrological parameters in each soil layer are then determined based on the relative abundance of organic and mineral material in the layer:

$$b = f_o \cdot b_o + f_m \cdot b_m \qquad \Psi_s = f_o \cdot \Psi_{s,o} + f_m \cdot \Psi_{s,m} \tag{2.28}$$

$$K_s = f_o \cdot K_{s,o} + f_m \cdot K_{s,m} \qquad \theta_s = f_o \cdot \theta_{s,o} + f_m \cdot \theta_{s,m} \tag{2.29}$$

$$C_{dry} = f_o \cdot C_{dry,o} + f_m \cdot C_{dry,m} \qquad \lambda_{dry} = f_o \cdot \lambda_{dry,o} + f_m \cdot \lambda_{dry,m} \tag{2.30}$$

My implementation of organic soils also differs from Lawrence and Slater (2008) in that I set the organic soil hydraulic conductivity to 0.02 kg $\text{m}^{-2}$ $\text{s}^{-1}$ (intermediate between values for hemic and fibric peat), a value an order of magnitude lower than the value of 0.1 kg $\text{m}^{-2}$ $\text{s}^{-1}$ for fibric peat used in the earlier study. I adjusted this parameter to achieve a better representation of global wetland areas (as discussed in Chapter 5) as it was found that the higher hydraulic conductivity value resulted in an underestimate of wetlands in certain regions as near surface soils remained too dry due to enhanced drainage.

### Soil moisture fluxes

Liquid water moves through the soil layers under the influence not only of gravity but also capillary, and adsorptive forces which act to retain water within soil (Ward and Robinson, 2000). Just as an energy potential term can be associated with the force of gravity, the soil *matric potential* ($\Psi$) is introduced as a potential associated with the forces associated with soil's affinity for water. The combined gravitational potential energy and matric potential is known as the soil water *potential*. As with gravitational potential energy, the energy associated with soil matric potential has no absolute zero-level. Rather, it is defined with reference to a well-defined body of water whose potential is set to zero. Generally this reference body of water is taken to be an unconfined body of pure water at the same temperature and elevation as the soil water. In this case, the matric potential represents the energy per kilogram of water that is required to move water from the reference reservoir to the soil sample.

Matric potential is often considered to be akin to a “negative pressure” acting to retain moisture in the soil as an increase in air pressure (equal in magnitude to the matric potential) can over come the matric potential forces and allow water to drain from the soil. Consequently, matric potential is often measured in Pa. The matric potential can equivalently be expressed as a metre-equivalent of water (as in MOSES) through a conversion factor that accounts for the pressure produced by a

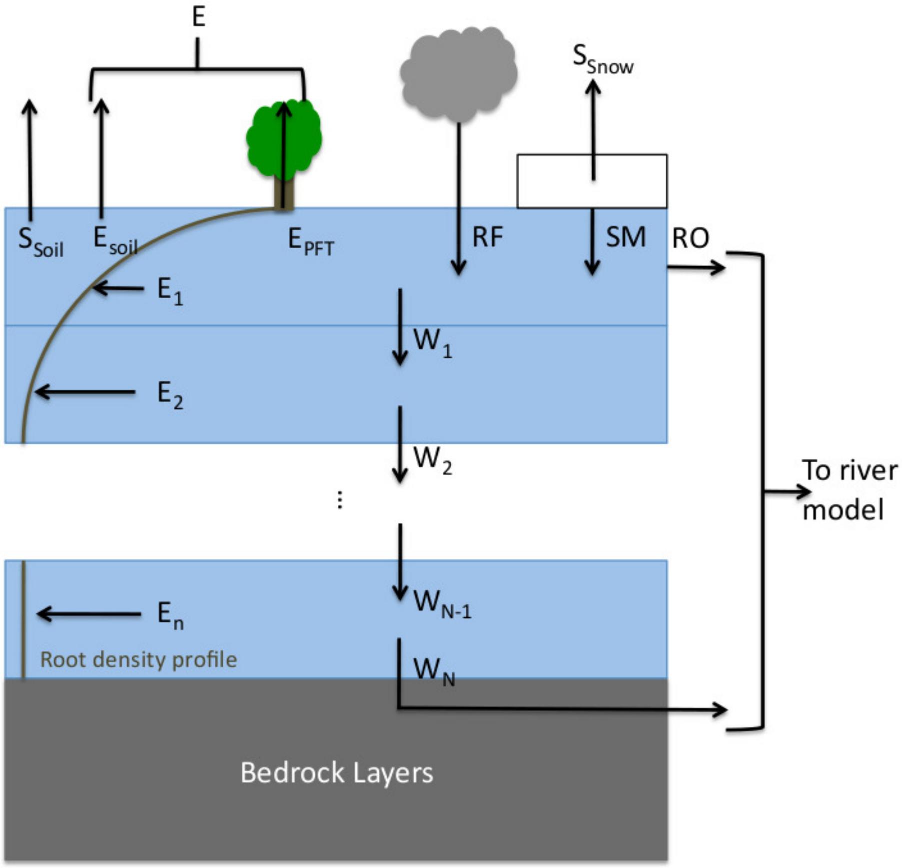


Figure 2.3: Soil moisture fluxes in UVic 3.0M

column of water with a given height. Generally, matric potential is expressed as a negative-definite quantity. In the absence of gravitational forces, soil moisture will tend to move from regions of high (less negative) matric potential to regions of low (more negative) matric potential as a result of the stronger attractive forces in the low matric potential region; the moisture flux is in the direction of decreasing matric potential just as the motion of a falling object under gravity moves in the direction of decreasing gravitational potential energy, or height. Here, both matric potential and the gravitational potential (measured by the depth below the surface, z) are treated as positive-definite quantities for convenience with a compensating negative sign inserted into the appropriate equations.

The magnitude of the matric potential depends on the amount of water present in a soil sample. As the water content increases, each individual water molecule becomes progressively less strongly bound by the attractive forces within the soil sample. Thus, the matric potential is a decreasing function of the amount of water present in the soil. Here, the matric potential depends on the liquid soil moisture content relative to its saturation value ($\Phi_s$), following Clapp and Hornberger (1978):

$$\Psi = \Psi_s \left(\frac{\theta_l}{\theta_s}\right)^{-b} \tag{2.31}$$

In principle, water will flow horizontally if there is a horizontal gradient in matric potential. However, the dominant soil moisture processes which alter moisture budgets and hence, the matric potential, occur at the surface (infiltration and evaporation) and bottom of the soil layer (interactions with ground water). As a result, it is generally found that the largest soil water matric potential gradients (and associated soil moisture movements) are in the vertical direction (Williams and Smith, 1989). Following this observation, it is assumed in UVic 3.0M that soil moisture moves exclusively in the vertical direction (as in most other land surface schemes) until it reaches the bedrock layer (at a nominal depth of 10 m) wherein it transported out of the gridcell as runoff. Water fluxes (W, positive downwards) follow Darcy's law in which fluxes are proportional to the gradient in soil water potential through the hydraulic conductivity:

$$W = K\left(\frac{d(\Psi + z)}{dz}\right) = K\left(\frac{d\Psi}{dz} + 1\right) \tag{2.32}$$

Moisture fluxes at each soil layer's boundaries are first calculated and the liquid moisture content within a layer is then updated following a discretized form of Richard's equation:

$$\frac{dM_{u,n}}{dt} = W_{n-1} - W_n - E_n \tag{2.33}$$

with $E_n$ as the fraction of the total evapotranspiration extracted from the $n^{th}$ soil layer. Soil moisture fluxes are limited such that the liquid soil moisture within a layer does not exceed its saturation value and does not fall below zero. Rainfall and snowmelt replenish the top layer soil moisture content. In the event that the rate of rainfall + snowmelt exceeds the surface hydraulic conductivity, excess moisture is diagnosed as runoff (RO) and sent to the river routing scheme. As before, the boundary condition at the base of the soil layers is that of free drainage. Frozen soil moisture in the top soil layer is also lost to sublimation ($S_{soil}$). Figure 2.3 gives a summary of the fluxes of moisture within the soil and between the soil, the surface and the atmosphere.

### Soil moisture stress and evapotranspiration

In the multilayer scheme, plant root characteristics vary through a PFT-dependent rooting depth ($\boldsymbol{d_r}$). Root density ($\boldsymbol{\rho_{root}}$) is assumed to decay exponentially with depth, with the rooting depth as the decay constant:

$$\boldsymbol{\rho_{root}}(z) = \rho_{root,o} e^{-\frac{z}{\boldsymbol{d_r}}} \tag{2.34}$$

Here, $\rho_{root,o}$ is a constant that is the same for all PFTs (and which never comes into play as the factor cancels out in the code). The root density profile affects moisture fluxes in two ways. Firstly, the overall soil moisture availability factor ($\boldsymbol{\beta}$, now PFT-dependent) is determined as an integral over the soil layers (with total soil depth, D) of the soil moisture availability factor at a given depth (B(z)), weighted by the root density:

$$\boldsymbol{\beta_{PFT}} = \frac{\int_0^D e^{-\frac{z}{\boldsymbol{d_r}}} B(z) dz}{\int_0^D e^{-\frac{z}{\boldsymbol{d_r}}} dz} \tag{2.35}$$

B has the same functional form as equation 2.8, and is a function of the volumetric liquid moisture concentration. Since B is constant within a soil layer, the integral may be expressed as a discrete sum over the N soil layers

$$\boldsymbol{\beta} = \frac{\sum_{n=1}^{N}(e^{-z_{top,n}/\boldsymbol{d_r}} - e^{-z_{bot,n}/\boldsymbol{d_r}})\,B_n}{1 - e^{-D/\boldsymbol{d_r}}} \tag{2.36}$$

where $z_{top,n}$ and $z_{bot,n}$ are the upper and lower boundary, respectively of the $n^{th}$ soil layer.

The distribution of roots is also used to partition the total evapotranspiration (E) diagnosed by the surface energy budget code into evapotranspiration within each layer ($E_n$). It is assumed that contribution of a layer to the total evapotranspiration from a PFT scales as the ratio of the moisture availability within a layer ($B_n$) to the total moisture availability function summed over all layers:

$$\boldsymbol{E_{n,PFT}} = \boldsymbol{E_{PFT}}\frac{(e^{-z_{top,n}/\boldsymbol{d_{r,PFT}}} - e^{-z_{bot,n}/\boldsymbol{d_{r,PFT}}})\,B_n}{(1 - e^{-D/\boldsymbol{d_r}})\boldsymbol{\beta_{PFT}}} \tag{2.37}$$

Thus, the total evapotranspiration withdrawn from a given layer is given by the above function summed over all PFTs, weighted by the fraction of gridcells occupied by the PFTS:

$$E_n = \sum_{PFT=1}^{5} \boldsymbol{f_{PFT}}\boldsymbol{E_{PFT}}\frac{(e^{-z_{top,n}/\boldsymbol{d_{r,PFT}}} - e^{-z_{bot,n}/\boldsymbol{d_{r,PFT}}})\,B_n}{(1 - e^{-D/\boldsymbol{d_r}})\boldsymbol{\beta_{PFT}}} \tag{2.38}$$

$E_1$, the evapotranspiration from the top soil layer includes an additional term to account from the evaporation from bare soil. In the event that the evapotranspirative demand exceeds the available liquid moisture in a given layer, the deficit is extracted from the next soil layer down.

### Thermal conductivity and Heat Capacity

A key difference from UVic 3.0 is that the soil thermal conductivity and heat capacity now depend on soil moisture contents, rather than being constants. Thermal conductivity varies linearly between its dry soil value ($\lambda_{dry}$) and its saturation value ($\lambda_{sat}$) based on the total moisture content relative to saturation:

$$\lambda = (\lambda_{sat} - \lambda_{dry})\,\frac{\theta}{\theta_s} + \lambda_{dry} \tag{2.39}$$

When the soil is dry, air fills all the pore spaces (fractional volume, $\theta_s$) with soil particles filling the remaining space (fractional volume, 1-$\theta_s$) and the thermal conductivity is a geometric mean of the compacted soil thermal conductivity ($\lambda_{cs}$)

and the thermal conductivity of air:

$$\lambda_{dry} = \lambda_{cs}^{(1-\theta_s)} \lambda_{air}^{\theta_s} \tag{2.40}$$

If the soil were saturated, water and ice would have respective volumetric concentrations given by:

$$\theta_l^s = \theta_s \, \frac{\theta_l}{\theta_l + \theta_i} \qquad \theta_f^s = \theta_s \, \frac{\theta_i}{\theta_l + \theta_i} \tag{2.41}$$

and the thermal conductivity of soil at saturation is given by a similar geometric mean:

$$\lambda_{sat} = \lambda_{cs}^{(1-\theta_s)} \lambda_w^{\theta_l^s} \lambda_i^{\theta_i^s} \tag{2.42}$$

The thermal conductivity of compacted soil need not be specified if the dry soil thermal conductivity is known as equations 2.40 and 2.42 can be used to express the saturated soil thermal conductivity in terms of the dry soil thermal conductivity:

$$\lambda_{sat} = \lambda_{dry} \left( \frac{\lambda_w^{\theta_l^s} \lambda_i^{\theta_i^s}}{\lambda_a^{\theta_s}} \right) \tag{2.43}$$

As the thermal conductivity of water and, in particular, ice (Appendix A) are considerably higher than of air, thermal conductivity represents a strong coupling between the hydrological and thermal states of the soil.

The soil's volumetric heat capacity is determined as the sum of the volumetric heat capacities of the soil particles, water and ice. As the volumetric heat capacity of air is substantially lower than that of the other constituents, it is negligible and not included in this calculation.

$$C_s = C_{dry} + \rho_w \, c_w \, \theta_l + \rho_i \, c_i \, \theta_f \tag{2.44}$$

### Soil heat fluxes

Heat fluxes between soil layers occur by diffusion, in proportion to the temperature gradient between layers. The thermal conductivity that mediates soil heat transport, is that at the boundary between layers ($\lambda_b$) which is determined by interpolating values of $\lambda$ calculated within layers to the layer boundary. Then, the diffusive heat flux through the base of layer n is:

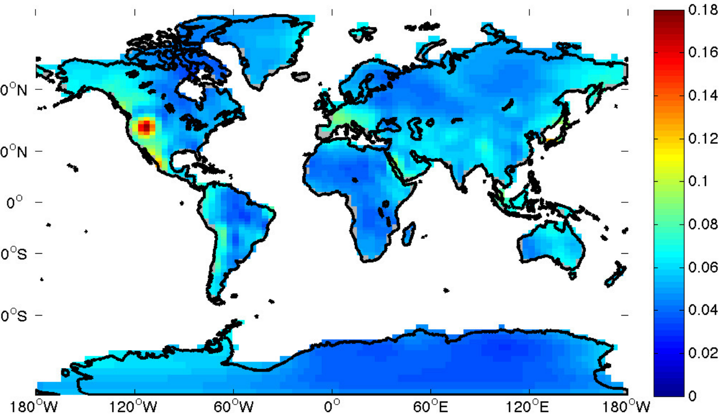


Figure 2.4: Spatial map of geothermal heat flux (W $m^{-2}$) from the work of Hamza et al. (2008).

$$GH_n = -\lambda_b \frac{T_{n+1} - T_n}{\frac{\Delta z_n}{2} + \frac{\Delta z_{n+1}}{2}} \tag{2.45}$$

As with moisture fluxes, heat fluxes are defined to be positive downwards. Since the scheme now accounts for the heat capacity of water in the soil, the flux of heat associated with water percolating through the soil must be accounted for in order for the model to conserve heat. The model first diagnoses the moisture flux that occurs in a given timestep and then subsequently diagnoses the heat transfer that occurs in the timestep. The latter calculation includes a term to account for the flux of heat that was transported by percolating water. This percolation flux for layer n is given by:

$$J_n = \begin{cases} c_w W_{n-1}(T_{n-1} - T_n) & W_{n-1} \geq 0 \,\&\, W_n \geq 0 \\ c_w W_{n-1}(T_{n-1} - T_n) - c_w W_n(T_{n+1} - T_n) & W_{n-1} \geq \ \&\, W_n < 0 \\ 0 & W_{n-1} < 0 \,\&\, W_n \geq 0 \\ -c_w W_n(T_{n+1} - T_n) & W_{n-1} < 0 \,\&\, W_n < 0 \end{cases} \tag{2.46}$$

The temperature of soil layer n is then updated as per the following equation:

$$C_{s,n}\Delta z_n \frac{dT_{g,n}}{dt} = G_{n-1} - G_n + J_n \tag{2.47}$$

Since moisture transport is restricted to soil layers, $J_n$ is zero for bedrock layers. The heat capacity and thermal conductivity of the bedrock layers are constants and consistent with the characteristics of granitic rock (Banks, 2008). The bottom boundary condition for the bottommost bedrock layer is a specified geothermal heat flux, derived from the work of Hamza et al. (2008). A map of this geothermal heat flux is shown in Figure 2.4.

### Phase changes

The thermodynamic point at which phase transitions occur in a systems is diagnosed based on the Gibbs free energy ($\Delta$G)[6]. The stable equilibrium state of a system is that in which $\Delta$G is minimized. When a system contains multiple phases of a substance, each is assigned its own free energy (e.g. $\Delta G_l$ and $\Delta G_s$ in the case of liquid and solid water, respectively). Each phase's free energy changes in response to pressure and temperature changes at different rates. Phase transitions occur when two phases have differing free energies in which case the quantity of the phase with the lower free energy is increased at the expense of the phase with the higher free energy in order to minimize the total free energy of the system. For example, under one atmosphere of pressure, pure water has a Gibbs free energy equal to ice at 0 °C. Below this point, the free energy of ice is lower than that of liquid water and water will freeze to ice as soon as the temperature falls below zero. Conversely, above 0 °C, the free energy of liquid water is lower than that of ice, so ice melts to water once the temperature crosses 0 °C. Thus, 0 °C is diagnosed as the freezing point of water at standard atmospheric pressure.

In soils, a number of factors come into play to both lower the freezing point and allow freezing to occur over a range of temperatures. Firstly, the freezing point of soil water is suppressed due to the fact that soil water contains dissolved salts which reduce its free energy relative to of pure water. However, the concentration of dissolved salts in soil water is seldom high enough that the freezing point is lowered much below -0.1 °C (Williams and Smith, 1989). While freezing begins at this salt-suppressed

---

[6] The Gibbs free energy can be thought of as the energy that a system contains that is available to do work (Davis, 2001)

freezing point, soil forces lower the free energy of the remaining liquid moisture once freezing starts to occur. This allows water to be in equilibrium with ice over a range of soil temperatures and for freezing to occur over this temperature range rather than at a single freezing temperature. The Gibbs free energy of soil water differs from that of unconfined water due to the effects of capillary and adsorptive forces. These forces were also responsible for changing the matric potential of soil moisture away from that of pure water at the same temperature and elevation. Thus, one can make an identification between the change in matric potential of soil water relative to unconfined water and an equivalent change in Gibbs free energy.

In UVic 3.0M, all moisture in a soil layer experiences the same matric potential and the same capillary and absorptive forces. In actuality, water in a given soil sample will experience a range in intensity of soil forces owing to range of different sizes of soil pores. Water molecules in small pore spaces experience strong soil forces, strong matric potential and thus a strong reduction in Gibbs free energy relative to unconfined water, while the reverse is true for molecules in the larger pore spaces (Davis (2001)). As soil temperature falls, the water in the largest pore spaces is the first to have the same ΔG as ice and thus the first to freeze. Thus, thinking of the matric potential in the soil model as a kind of average matric potential experienced by soil water, the matric potential should get progressively larger as the soil temperature continues to fall and liquid water remains in progressively smaller and smaller pores. A relationship can be obtained for the matric potential as a function of temperature through a complicated derivation involving minimizing the Gibbs free energy of the soil-water-ice system at temperatures below the freezing point of unconfined water ($T_m = 273.15$ K) (see Cox et al. (1999); Davis (2001); Williams and Smith (1989) for details):

$$\Psi(T) = -\left(\frac{\rho_{ice}}{\rho_{water}}\frac{L_f}{T_m g}\right)(T - T_m) \tag{2.48}$$

where $L_f$ is the latent heat of fusion, $\rho_w$ is the density of water, $\rho_i$ is the density of ice, $T_m$ is the normal melting point of ice (273.15 K), g is the acceleration due to gravity and T is the soil temperature. This equation also describes the volumetric liquid moisture content as a function of temperature; equating 2.48 and 2.31 yields the following result:

$$\theta_l^{max} = \theta_s\left(-\frac{\epsilon(T - T_m)}{\Psi_s}\right)^{-\frac{1}{b}} \tag{2.49}$$

where

$$\epsilon = \frac{\rho_{ice}}{\rho_{water}} \frac{L_f}{g T_m} = 114.3 m \, K^{-1} \tag{2.50}$$

$\theta_l^{max}$ is the maximum amount of liquid water that *could exist* in soil at temperature T. The *actual* amount of liquid water in the soil may be less than $\theta_l^{max}$ if the total volumetric water content, $\theta$, is less than $\theta_l^{max}$, in which case the soil water freezing has yet to begin. Equating $\theta$ to $\theta_l^{max}$ and solving for temperature gives $T_{max}$, the temperature above which all water is liquid, or, the temperature below which soil water begins to freeze.

$$T_{max} = T_m - \frac{\Psi_s}{\epsilon} \left( \frac{\theta_s}{\theta} \right)^b \tag{2.51}$$

Phase changes in soil occur over a range of soil temperatures, resisting the rate of soil temperature change as thermal energy is transferred to / from latent heat. To account for this, the heat capacity of the soil is modified to account for phase change effects. This *effective heat capacity* is given as:

$$C_{s,eff} = C_s + \rho_w \, c_w \, \theta_l + \rho_i \, c_i \, \theta_f + ((c_w - c_i)T + L_f) \rho_w \frac{\delta \theta_l}{\delta T} \tag{2.52}$$

The additional term in this equation represents two factors, both proportional to the mass of $H_2O$ that changes phases. Firstly, since liquid water has a greater specific heat capacity than ice, when ice melts to liquid water, energy must be absorbed in increasing the heat capacity of the water undergoing the phase transition. Secondly, to effect the phase transition latent heat must be absorbed to break the bonds in ice.

### Soil respiration

Soil respiration is also altered in UVic 3.0M. Rather than having separate soil carbon stocks within each soil layer and specifying functions to account for carbon transfer between layers, a single soil carbon pool is maintained, but is considered to be distributed amongst different soil layers, following the soil cumulative carbon distribution function introduced earlier. The overall soil respiration is then determined based on the relative amount of carbon in each layer and the moisture and temperature conditions in the layers. Thus:

$$R_s = \kappa_s \mathcal{C}_s \sum_{n=1}^{N} \frac{(CCP(z_{bottom,n}) - CCP(z_{top,n}))}{100\%} f(\theta_{u,n}) f(T_n) \tag{2.53}$$

### 2.2.1 Revised Surface Energy Budget

Several minor modifications where made to the surface energy budget code. Firstly, the form of surface-to-atmosphere moisture fluxes is more complicated than in UVic 3.0 where moisture fluxes on a tile took the form of either evaporation *or* sublimation from lying snow. Now, moisture fluxes depend on the state of the land surface as follows:

1. Where snow is present and the snow height **exceeds** the vegetation height (or where any snow is present on a bare soil tile), moisture fluxes consist exclusively of sublimation from the snow pack. Snow sublimation is subject only to an aerodynamic resistance.
2. Where snow is present and the snow height is **less than** the vegetation height, transpiration occurs from vegetation along with sublimation from the snow pack. Evapotranspiration is governend by the soil moisture stress factor based on the liquid soil moisture content of the tile.
3. Where snow is absent on vegetated tiles moisture fluxes occur exclusively by transpiration and evaporation from the soil below the canopy.
4. Where snow is absent on bare soil tiles, moisture fluxes occur by evaporation and by sublimation from bare soil if any frozen moisture is present.

In the latter case, there will generally be some liquid moisture present in the soil alongside the frozen moisture. Here then, both sublimation and evaporation is allowed to occur at the same time. Sublimation is assumed to occur over the fraction of the tile containing frozen soil moisture, while evaporation occurs over the fraction containing liquid soil moisture. Thus, the sublimation and evapotranspiration rates are scaled by the following factors:

$$f_{frozen} = \frac{\theta_f}{\theta_l + \theta_f} \qquad f_{liquid} = \frac{\theta_l}{\theta_l + \theta_f} \tag{2.54}$$

As well, the function used to calculated the bare soil surface resistance to evaporation was updated. Previously, soil surface resistance had the same functional form as the soil moisture stress factor used in assessing evapotranspiration from plants. The new functional form for soil surface resistance is consistent with empirical functions derived from field observations of bare soil evaporation (see Ács (2003) for example). A similar function was assumed for resistance to bare soil sublimation. The functions take the form:

$$\boldsymbol{r_{s,soil,evap}} = min\left(10^6, \left(\frac{\theta_c}{\theta_l}\right)^2\right) \qquad \boldsymbol{r_{s,soil,sub}} = min\left(10^6, \left(\frac{\theta_c}{\theta_f}\right)^2\right) \tag{2.55}$$

The energy budget code was also altered to allow snowmelt to occur when the skin temperature exceeds the freezing point. Recall that $T^*$ is the skin temperature that balances the energy budget. If $T^*$ exceeds the freezing point and snow is present, it is reset to 273 K and the energy budget is re-evaluated at the freezing point. The energy budget at this point will not be balanced without the addition of a new energy term - this new term represents energy that is available to melt snow. Thus in the new scheme, snowmelt occurs both from the top of the snow surface, when $T^*>273$ K and from the bottom of the snow surface if $T_{g,1} > 273$ K.

Finally, the code was modified to account for the insulating effects of snow. In UVic 3.0, the presence of snow has no impact on surface-to-soil heat fluxes as the thermal conductivity governing such fluxes is simply that of the soil, even if snow is present. In UVic 3.0M, the snowpack is treated, not as a separate layer, but as an extension of the top soil layer for the purposes of determining an effective thermal conductivity in the presence of snow. In this procedure, $T_{g,1}$ remains consistently at a depth $\Delta z_1$ / 2 below the surface which may be within the soil layer (first case below), or within the snow layer (second case) depending on the depth of snow. In the first case the effective surface-to-soil thermal conductivity is determined by treating the snow and soil layers that heat must pass through as resistors in series. The effective thermal conductivity for heat transfer between the surface and the first ground layer is then given by:

$$\lambda_{eff} = \begin{cases} \lambda_{soil}\left(\frac{\Delta z_1}{\Delta z_1 + 2\Delta z_{snow}(\lambda_{soil}/\lambda_{snow} - 1)}\right) & \Delta z_{snow} \leq \Delta z_1 \\ \lambda_{snow} & \Delta z_{snow} > 0.5\Delta z_1 \end{cases} \tag{2.56}$$

## 2.3 Offline Model Configuration

I also developed an *offline* version of the land surface scheme in which the LSS was decoupled from the other climate system components and driven by reanalysis data. The development of this version was motivated by a desire to better understand the characteristics of the new implementation of MOSES without the added complexity of climate feedbacks resulting from land surface - atmosphere feedbacks. As well, the creation of an offline version of the model allowed me to examine the impact of climate biases on the state of permafrost and seasonally frozen ground as will be discussed in Chapter 3.

In the offline version, variables that would have been determined by the energy-moisture balance atmospheric model (surface air temperature, relative humidity, rainfall and snowfall) were specified from the ECMWF ERA-40 reanalysis product (Uppala et al., 2005). I denote this version of the model as UVic 3.0E (to distinguish it from the fully coupled version of the model - UVic 3.0M). Monthly mean fields were derived from 6-hourly data over the entire 45 years of the ECMWF reanalysis, which extends from September, 1957 to August, 2002 to produce a mean monthly climate for this period. The data supplied to the model in a particular timestep is then determined by linearly interpolating between the monthly fields that bracket the timestep.

Not all the atmospheric data supplied to the LSS was derived from the ECMWF reanalysis since some atmospheric variables required by the LSS are specified from data fields by default (wind speed) or by model option ($CO_2$ concentration). Other key variables are calculated within the model based on existing data fields (shortwave radiation, since the atmospheric shortwave albedo is specified by a monthly data field), or are a function of data specified from the reanalysis (longwave radiation, a function of surface air temperature). As the LSS operates on an hourly timestep, it is also necessary to represent the diurnal cycle in atmospheric variables. The diurnal cycle in insolation is calculated within the model and the diurnal temperature range is determined from monthly data fields; the determination of both of these variables is unchanged from UVic 3.0 and UVic 3.0M.

Since the land surface scheme is coupled to a dynamic vegetation model, vegetation both responds to and influences atmospheric and soil climate variables. Thus, a change in sub-surface conditions introduced by varying land surface scheme parameters or physics can be amplified or diminished by vegetation feedbacks, complicating the analysis of sensitivity experiments. As well, many vegetation parameters in the UVic ESCM were adjusted to achieve a realistic representation of vegetation given a particular configuration of the LSS. If any of the additional versions of the LSS presented in this chapter were adopted into a new operational version of the model, it is likely that some adjustment would be necessary to the vegetation parameters to accompany such a change. As developing a new set of vegetation parameters for each model configuration variant would require a substantial amount of effort, I decided that the dynamic vegetation model should also be decoupled from the LSS for the experiments conducted in this chapter and the following chapter. Thus, additional monthly data fields of vegetation fraction, height, leaf area index and litterfall were produced from a fully coupled simulation and used to drive the offline LSS. These were developed based on the equilibrium vegetation distribution from the UVic 3.0M under perpetual year 1800 forcing. Some vegetation interactions remain in the offline version of the LSS through the control that vegetation exerts on evapotranspiration. Canopy conductance (and surface resistance for PFTs) remains a function of soil moisture concentration. Thus, through vegetation, rates of evaporation still vary based on changes in soil liquid water content.

In Appendix B, I use the offline configuration of the model to understand how the impacts of the modifications to the land surface scheme outlined in this chapter affect the the thermal and hydrological state of the land surface scheme.

## 2.4 Summary

In this chapter, the UVic Earth System climate model was described. Fluxes of heat, moisture and carbon between the atmosphere and the land surface and sub-surface are governed by the model's land surface scheme, MOSES, while the state of vegetation and the soil carbon pool is determined by the TRIFFID dynamic vegetation module. UVic ESCM version 3.0 was the starting point for my thesis research and the land surface scheme was subsequently modified to allow the model to simulate permafrost and seasonally frozen ground. The key modifications made to UVic 3.0 to produce the new UVic 3.0M version of the model were:

- An increase in the number of subsurface layers from a single 1 m thick layer to 14 layers with a total thickness of 250 m. The top eight layers (to a depth of 10 m) are soil layers, while the bottom six layers are granitic bedrock.

- A revision of soil thermal and hydrological parameters. UVic 3.0M features soil mineral and organic characteristics that vary from gridcell to gridcell based on soil organic and mineral characteristics from the ISLSCP-II soil dataset.
- Variation of soil thermal and hydrological parameters depending on soil moisture content and phase.
- Inclusion of a representation of soil phase changes over a range of temperatures in the soil.
- A more sophisticated treatment of snow processes. This includes simulation of the insulating effect of snow, a better representation of snowmelt and a more sophisticated diagnosis of moisture fluxes in gridcells containing lying snow.

Finally, an offline configuration of the land surface scheme forced by ERA-40 reanalysis data (UVic 3.0E) was introduced.

# Chapter 3

# Permafrost in the Present-Day Climate: Modelled Distribution and Sensitivity Experiments

In this Chapter, I compare the simulated distribution of regions of permafrost, seasonally and intermittently frozen ground in the mid-to-late $20^{th}$ Century against available observations. In the first part of this chapter, the permafrost distribution of two different configurations of UVic ESCM 3.0M are contrasted. The first is a configuration with the land surface scheme fully coupled to the other model components (the *coupled* version - U3.0M), while the second is the offline configuration driven by atmospheric fields generated from ECMWF ERA-40 forcing data (U3.0E)). The objective of this comparison is to identify key biases in the UVic ESCM climatology and to assess how these biases might impact the simulated permafrost distribution. In the second portion of the chapter, I perform a series of sensitivity experiments where land surface scheme characteristics are perturbed and investigate the resulting effect on the permafrost and seasonally frozen ground distribution to gain a sense of the extent of the uncertainty in the model simulations. Version U3.0E of the model is used for this purpose as the lack of land surface-atmospheric feedbacks was found to facilitate analysis.

## 3.1 Model Climate Biases

Like many global climate models, the UVic ESCM displays notable regional biases when its climatology is compared against observational datasets. Here, I identify biases in atmospheric fields in high latitude regions by comparing model output to the ERA-40 reanalysis. The coupled model was first allowed to reach an equilibrium state by spinning the model up for 4,000 years under perpetual year 1800 forcing, at which point time varying forcing fields were specified to drive the model from 1800 to 2005. To facilitate the equilibration of deep permafrost, the ground temperature at the base of the soil profile was evaluated after 1,000 years of spin-up and the deep soil temperatures were initialized by using this value adjusted by the geothermal gradient. This allowed the deep layers of the model to equilibrate relatively rapidly. Greenhouse gas forcings were applied by specifying the atmospheric carbon dioxide concentration (whose radiative effects are calculated internally as a reduction in outgoing longwave radiation) and an additional radiative forcing field (calculated outside the model) to account for the radiative effects of non-$CO_2$ greenhouse gases. Historical crop and pasture regions were specified as fractions of land grid cells; within these fractions, only $C_3$ and $C_4$ grasses were allowed to exist and compete in TRIFFID. The direct effect (i.e. reflective effect) of sulphate aerosols was also specified as a perturbation to the local surface albedo as described by Matthews (2004), based on fields of sulphate optical depths. All of the above forcings were derived from historical data fields supplied with Representative Concentration Pathway datasets (Moss et al. (2010)). A more detailed description of the generation of these forcings and their inclusion in the UVic ESCM is described in the next chapter. Additionally, variations in shortwave radiation associated with orbital parameters and volcanic forcings were accounted for using pre-existing model data fields and code (Matthews (2004); Weaver et al. (2001)). Model output was extracted for 1957 - 2002 (the same interval as the ERA-40 data) and annual and seasonal mean fields for variables of interest were generated for this period.

### 3.1.1 Temperature Biases

The atmospheric component of the coupled model is a 2-dimensional energy-moisture balance model (EMBM). The atmosphere in this case is vertically integrated, meaning that a single surface air temperature (SAT) and surface specific humidity ($q_a$) value is simulated in each grid cell. The total energy and moisture content of the atmosphere is related to these surface values under the assumption that energy and moisture decrease exponentially with height. If the EMBM were three-dimensional, the energy content of the atmosphere (E, J m$^{-3}$) at a given height would be given by the following equation:

$$E(z) = \rho_a(z) c_{pa}(z) T_a(z) \tag{3.1}$$

where $\rho_a$ is the density of air, c$_{pa}$ is the specific heat capacity of air at constant pressure and T$_a$ is the atmospheric temperature in K. The atmospheric moisture content (M, kg m$^{-3}$) at a specific height would depend on the density of air and the specific humidity, q$_a$:

$$M(z) = \rho_a(z) q_a(z) \tag{3.2}$$

The rate of decrease of energy and moisture content with height are determined by energy and moisture scale heights (h$_T$ = 8.4 km and h$_q$ = 1.8 km, respectively) that represent the height at which these quantities would have dropped to 1/e of their surface values[1]. One can easily integrate these function from the surface (z = 0) to the top of the atmosphere (z = $\infty$), to obtain expressions for the total atmospheric energy and moisture content in terms of surface values:

$$E = \rho_a c_{pa} h_T T_a \qquad M = \rho_a h_q q_a \tag{3.3}$$

The simulated SAT evolves according to the following equation:

$$\rho_a c_{pa} h_T \frac{\delta T_a}{\delta t} = Q_{local} + \rho_a c_{pa} h_T \boldsymbol{\nabla} \cdot (\nu \boldsymbol{\nabla_h} T_a) + \rho_a c_{pa} h_T \beta_T \boldsymbol{\nabla} \cdot (\boldsymbol{U_T} T_a) \tag{3.4}$$

The first term, Q$_{local}$, represents local sources and sinks of energy due to absorption of incoming shortwave radiation and outgoing longwave radiation from the surface, loss of long wave from the atmosphere due to radiative cooling, exchange of sensible heat between the atmosphere and surface and latent heat release from precipitation. The second term is a diffusion term that represents the heat transport by large scale atmospheric circulation as well as heat transport by eddies at synoptic and smaller scales. The rate of diffusion is governed by $\nu$, a latitude-dependent, time-invariant

---

[1] $\rho_a$ is taken to be constant at the surface and equal to 1.25 kg m$^{-3}$ while c$_{pa}$ is equal to 1004 J kg$^{-1}$K$^{-1}$ at the surface

diffusion coefficient that was empirically derived to simulate SAT and precipitation rates in good agreement with observations. The last term accounts for heat transport by advection with $\boldsymbol{U_T}$ as the wind field governing temperature advection, derived from the NCEP reanalysis winds and with $\beta_T = 0.3$ as a dimensionless parameter used to control the extent of advective heat transport.

Figures 3.1 and 3.2 show annual mean and seasonal mean surface air temperature differences between the coupled model and ERA-40. While reanalyses are not flawless representations of the climate, they are constrained by observations and the ERA-40 data presented here is treated as a good estimate of the true state of the global climate. Although the ESCM air temperatures generally agree well with observations in the annual and zonal mean (See Figure 10 c) in Weaver et al. (2001), or Figure 6 a) in Meissner et al. (2003), for example), the model has some relatively strong regional temperature biases.

Three regions in particular show strong temperature biases than may impact the simulated permafrost distribution[2]. Firstly, the north-eastern margins of North America and Asia are consistently warmer than observations, especially during the winter months (Figure 3.2, bottom panel) where certain regions may be up to 20 K too warm[3]. Secondly, the high altitude Tibetan plateau region tends to be too cool, especially in the spring and summer (Figure 3.2, first and second panel). Finally, land areas adjacent to the North Atlantic tend to be too cool, especially in winter. All three of these anomalies have persisted since early versions of the model (see Figures in Weaver et al. (2001) and Meissner et al. (2003), for example). An additional set of comparison plots (not shown) were generated with version 3.0 and compared to those in 3.0M, indicating that the newly-included land surface processes in 3.0M had little impact on SAT biases.

The cause of the North Atlantic bias is well understood: in the ESCM, North Atlantic Deep Water formation tends to occur too far to the south with the result that there is more-extensive sea ice simulated in the region than is observed. The resulting higher surface albedo leads to less absorbed insolation and cooler SATs. The other temperature anomalies are likely due to the level of sophistication of the atmosphere model which does not represent some atmospheric dynamics (such as convective motion) and features a relatively simplistic representation of energy transport by the

---

[2] Refer to Figure 3.9 further in this chapter for a map of regions of permafrost and seasonally frozen ground

[3] surface air temperature over the Greenland ice sheet is also warm relative to ERA-40, though nearly all of Greenland is glaciated and there is little exposed, permafrost-bearing land

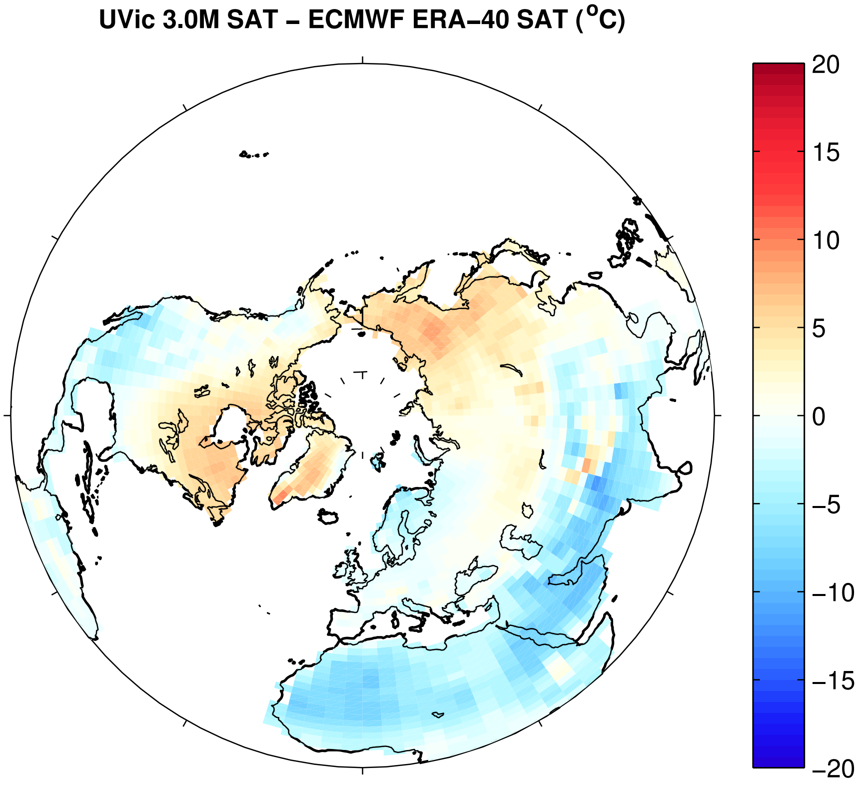


Figure 3.1: Annual mean model surface air temperature bias. The plots show the difference between temperature fields of the coupled UVic 3.0M model and the ECMWF ERA-40 surface air temperature fields used to force the offline version of the model.

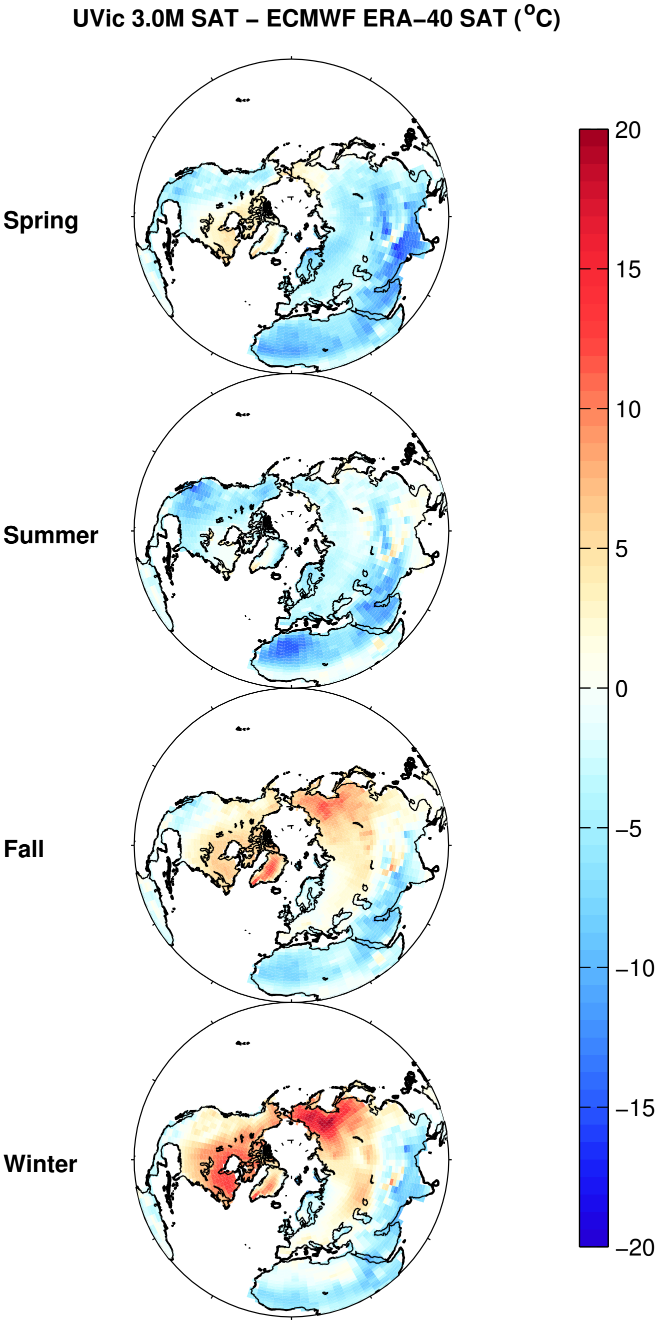


Figure 3.2: Seasonal model temperature biases. The plots show the difference between temperature fields of the coupled UVic 3.0M model and the ECMWF ERA-40 surface air temperature fields used to force the offline version of the model.

large scale circulation. The strongest and most extensive of the temperature anomalies, as far as representation of permafrost in the model is concerned, is the winter warm bias in north-eastern North America and Asia. It has been suggested (Michael Eby, personal communication) that this may be due to the static atmospheric diffusion coefficient in the model.

Meridional heat transport in the model depends both on the meridional temperature gradient and the atmospheric diffusion coefficient. Poleward heat transport in the climate system tends to vary strongly throughout the year with the strongest heat transport occurring in the winter months associated with the strongest meridional temperature gradients and stirring activity of midlatitude storms. While the model does simulate seasonally varying temperature gradients, the diffusion coefficients do not vary through time. It is therefore possible that this approach tends to lead to winter heat transport that is too strong, resulting in warmer SATs at high latitudes. That the warm bias is particularly pronounced in eastern margins is likely due to advective transport in the atmosphere which communicates the influence of the ocean to inland sites. Heat transport is more realistically handled in the 3-dimensional ocean model. Given the prevalent westerly winds in the midlatitudes, sites along the eastern margins are farthest downwind from the ocean and are therefore most significantly affected by issues with atmospheric heat transport.

### 3.1.2 Precipitation Biases

The distribution of permafrost may also depend to some extent on the ESCM's simulated precipitation. Rainfall and snowmelt determine the soil's moisture budget which, in turn, influences its thermal conductivity and heat capacity. As well, evapotranspiration influences the surface energy budget and through that, the ground heat flux. Finally, snowfall influences the ground thermal state by altering lowering the surface thermal conductivity when present and altering the surface energy budget when snowmelt occurs.

The specific humidity in the atmospheric component of the ESCM is governed by the following equation:

$$\rho_a h_q \frac{\delta q}{\delta t} = \rho_w (E + S - P) + \rho_a h_q \boldsymbol{\nabla} \cdot (\kappa \boldsymbol{\nabla} q) + \rho_a h_q \beta_q \boldsymbol{\nabla} \cdot (\boldsymbol{U_q} q) \tag{3.5}$$

As in the temperature equation, the first term represents local sources (evapotranspiration/sublimation) and sinks (precipitation) of moisture and the second and third

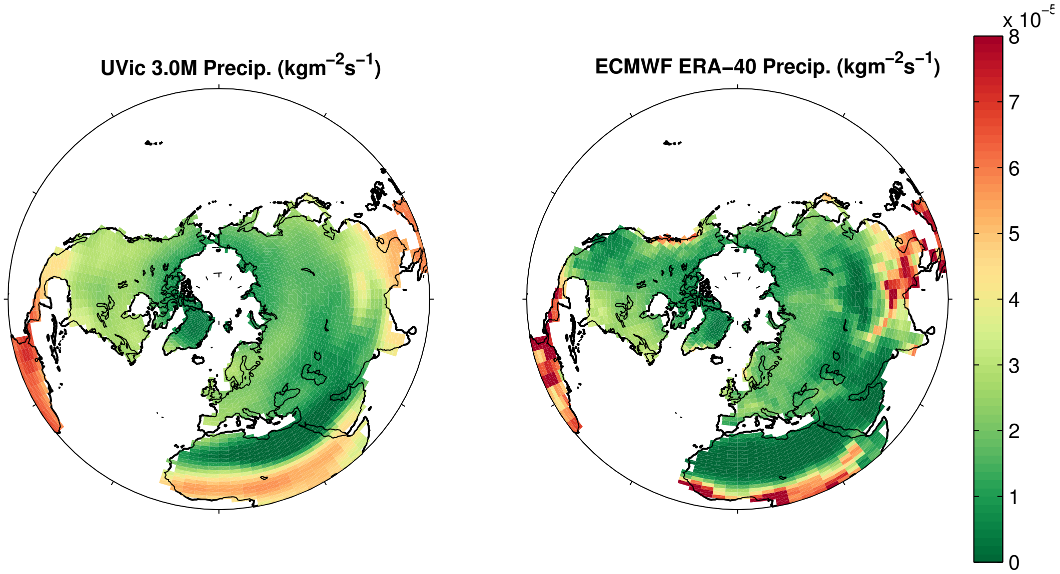


Figure 3.3: Annual mean precipitation rates (kg m$^{-2}$s$^{-1}$) in UVic 3.0M and the ERA-40 reanalysis.

terms account for diffusion and advection of moisture, respectively. $\kappa$ is the latitude-dependent, horizontal moisture diffusivity coefficient, $\beta_q$ $(= 0.5)$ is the advection scaling factor for moisture and $\boldsymbol{U_q}$ is the wind field used for moisture advection.

The precipitation field in the coupled model compares favourably with the reanalysis in terms of the annual mean precipitation rate (Figure 3.3), although the model precipitation field tends to be more diffuse. Restricting the analysis to mid-to-high latitudes, I note that the ESCM tends to have heightened precipitation rates in the interior of continents relative to ERA-40. This is likely a consequence of the simplified representation of moisture transport by diffusion and advection by mean winds in the 2-D atmosphere module. Surface air temperature biases over the ocean may also contribute to this precipitation bias. For instance, SATs over the North Pacific tend to be persistently warm (not shown) by 2-5 K. This may lead to air over the North Pacific acquiring excessive amounts of moisture, which would contribute to a specific humidity bias over continental interiors. Given the generally westerly pattern of the prevailing wind in the midlatitudes, this may explain why precipitation rates in the continental interior of North America differ most strongly from the reanalysis. Figure 3.4 indicates that these anomalously high precipitation rates are particularly pronounced in winter and to a lesser extent in fall and spring. However, in the summer

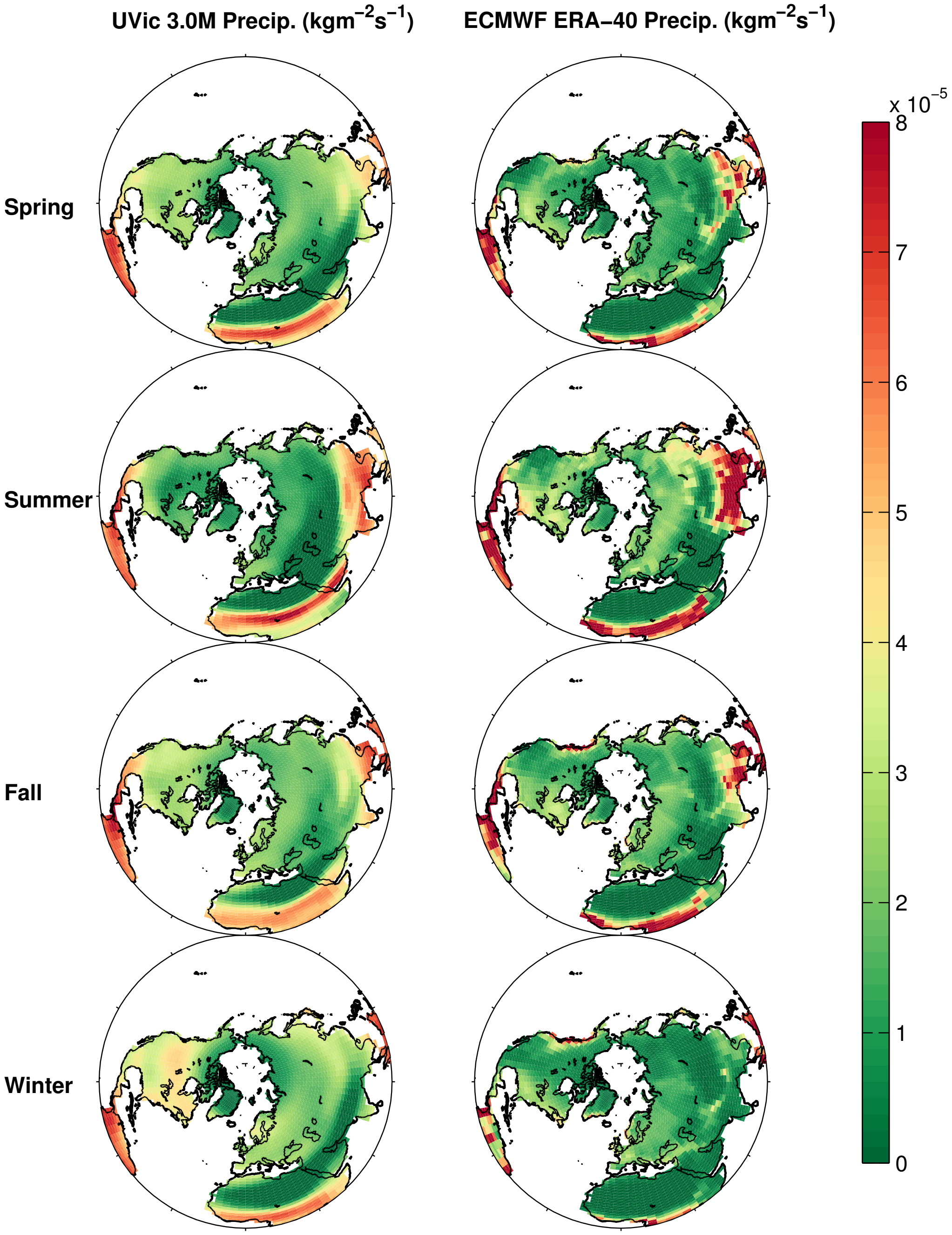


Figure 3.4: Seasonal mean precipitation rates (kg m$^{-2}$s$^{-1}$) in UVic 3.0M and the ESCM ERA-40 reanalysis.

months, the ESCM typically underestimates high latitude precipitation. A possible reason for the low simulated summer precipitation is the model's inability to represent convective precipitation; a detailed analysis of the ERA-40 Arctic precipitation regime (Serreze and Hurst (2000)) indicates that summer convective precipitation accounts for a substantial portion of the total precipitation.

### 3.1.3 Snow Cover

The snow scheme in the ESCM's land surface scheme is very simple, consisting of a single snow layer with constant snow density and thermal conductivity and no snow heat capacity. Processes such as compaction, retention and refreezing of liquid water in the snowpack and blowing snow are also not represented. The direct impact of snow on ground temperature is therefore determined solely by the snow depth and the constant snow density and thermal conductivity; in a more realistic snow scheme the snow thermal conductivity, heat capacity and density would also vary in time. With constant snow density, the modelled snow depth is directly proportional to the snow water equivalent (SWE - measured in metres of water). Here, the SWE distributions of the coupled model (U3.0M) and the land surface scheme forced by the ERA-40 reanalysis (U3.0E) are examined.

The offline configuration of the LSS, U3.0E, was spun up by forcing it with atmospheric fields from the ERA-40 data. The model was repeatedly forced with an annual cycle of ERA-40 data, with data supplied as monthly fields of atmospheric variables. Each of these monthly fields was a mean taken over the entire reanalysis period. In terms of model output, annual and seasonal mean fields were generated from one year's worth of output once the model had reached equilibrium. The spin up procedures for U3.0M and U3.0E are not identical: the coupled model evolves through the 1957-2002 period under time-varying forcings and the simulated climate is therefore not in equilibrium with the forcings. In the second case, the land surface scheme equilibrates to a mean atmospheric climate for the 1957-2002 period. Nonetheless, despite the difference in spin-up procedures, I felt that a comparison of the state of the land surface scheme in the two configurations would allow for a useful assessment of the impact of atmospheric biases on the land surface scheme.

The SWE fields from the two configurations of the ESCM are compared with the simulated SWE from the ERA-40 reanalysis and satellite-derived estimates of northern hemisphere SWE. Snow is modelled in the ERA-40 land surface scheme as a single layer above four soil layers, with snow heat capacity and thermal conductivity dependent on the variable snow density (van den Hurk et al. (2000); Viterbo and Beljaars (1995)). Snow depths from surface synoptic reports are also assimilated as part of the reanalysis (Clifford (2010)). The satellite snow water equivalent data covers the period from 1978 to 2003 and uses brightness temperatures derived from microwave-band measurements from the SMMR (1978-1987) and SSM/I (1987-2003) sensors (Armstrong et al. (2007)) to determine SWE values; visual band measurements are also recorded to determine snow covered area (SCA) (Clifford (2010)).

Clifford (2010) compared the satellite-derived SWE and SCA values with those derived from the ERA-40 reanalysis and Hadley Centre HadCR3 model; panels a) and b) in Figure 3.5 are taken from her paper. The annual cycle in total Northern Hemisphere SCA (panel b, Figure 3.5), is nearly identical in the ERA-40 reanalysis and HadCM3, both of which show a greater SCA throughout the year than the satellite retrieval results. Clifford (2010) suggests that the satellite data likely underestimates the maximum SCA at any given time as satellite snow cover detection is felt to be reliable except in regions of very thin snow cover where sensors have detection issues.

Panel d) of Figure 3.5 shows the annual cycle in SCA in three versions of the UVic model: U3.0M, U3.0E and UVic ESCM version 3.0[4]. The seasonal cycle in SCA in U3.0E closely follows and is consistently slightly higher the ERA-40/HadCR3 curves. SCA values are consistently overestimated in all versions of the UVic model as a consequence of the model's land grid which identifies too much exposed Northern Hemisphere land mass as shall be discussed in the following section. The U3.0M SCA more closely tracks the satellite data with peak SCA in February (vs. January in U3.0E), but again SCA values are higher than observed both in HadCR3/ERA-40 and satellite data.

Snow extent datasets are more commonly used in the literature for model validation than SWE datasets. However, as the snow water equivalent at a site determines the amount of fresh water released during spring melt, the degree of insulation of the ground and the amount of energy required for snowmelt amongst other factors, accurate simulation of SWE is important in assessing the performance of land surface schemes in cold regions. Assessing the distribution of SWE in models and satellite output is challenging owing to the difficulty in ground truthing data. Panel a) in Figure 3.5 shows monthly mean SWE in the Northern Hemisphere from HadCR3,

---

[4] The version of the ESCM without any changes implemented by the author - results from this configuration of the model shall be addressed shortly.

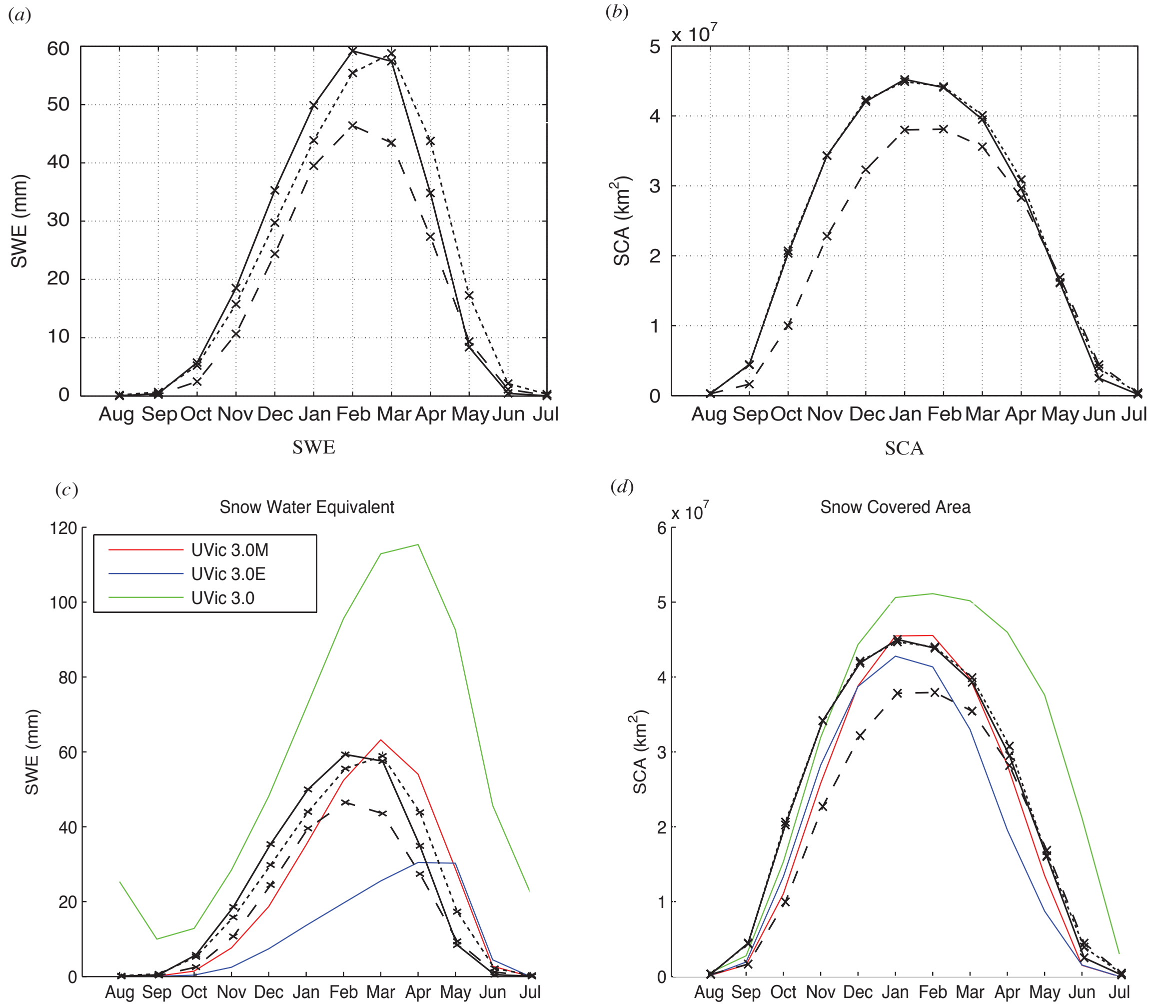


Figure 3.5: Annual cycle in Northern Hemisphere average climatologies of a) and c) snow water equivalent (SWE) (mm) and b) and d) total snow covered area (SCA) (km$^2$). Panels a) and b) are taken from Clifford (2010). Solid line: ECMWF ERA-40, dashed line: satellite SSM/I sensor, dotted line: HadCM3. Red line: UVic 3.0M, blue line: UVic 3.0E, green line: UVic ESCM 3.0. Note that plot y-axis ranges are different between panels a) and c) as well as b) and d)

ERA-40 and satellite SSM/I data. Again, ERA-40 and HadCR3 values track each other closely and are consistently substantially higher than the SSM/I data. Panel c) of Figure 3.5 shows the hemisphere-mean SWE in the different configurations of the UVic model. The general shape and peak value of mean SWE in U3.0M agrees well with the results shown in panel a), but the simulated SWE in U3.0M tends to be lower than the other timeseries in panel a)from September through to January and higher than the other timeseries from January through to June. In contrast, the U3.0E values SWE equivalent values are considerably lower than the other data products for all but three months of the year (May-July).

The spatial distribution of SWE varies substantially between datasets as shown in the comparison between seasonal mean SWE values between ERA-40 and satellite data in Figure 3.6. The discrepancy between ERA-40 and satellite data is highest in central Siberia, Eastern Canada and Alaska where ERA-40 shows substantially higher SWE values than the satellite output. At present, it is unclear whether this discrepancy stems from issues with algorithms used to determine SWE from satellite measurements or a poor representation of high-latitude precipitation and/or snow processes in the reanalysis (Clifford (2010)). Equivalent plots for U3.0M and U3.0E are shown in Figure 3.7. U3.0E values are consistently much lower (especially at the highest latitudes) than the ERA-40 values despite comparable snow-covered areas although regions of peak SWE occur in similar regions in both U3.0E and ERA-40 plots. U3.0M does a considerably better job at representing SWE. The U3.0M SWE fields are more diffuse than U3.0E, reflecting a similar character in U3.0M's precipitation. The U3.0M SWE plots more closely resembles the satellite fields than the ERA-40 fields although U3.0M does not simulate peak winter values that are as high as the satellite fields and appears to noticeably underestimate peak SWE values at sites close to the Arctic Ocean, particularly in winter months. Given the large disparity between reanalysis and satellite products in terms of SWE, I feel that U3.0M does a reasonable job of representing SWE.

The large difference between U3.0E and ERA-40 snow fields is particularly curious. Either the ERA-40 SWE values are very strongly influenced by the assimilation of surface synoptic reports[5] or, more likely, there is some issue with the representation of snow in version 3.0M of the UVic model code, which is common to both versions 3.0M and 3.0E. To investigate this issue, snow cover and SWE from the earlier version

---

[5] i.e. the surface scheme of ERA-40 too would tend to show the low SWE values as in U3.0E were it not for the influence of the assimilated data

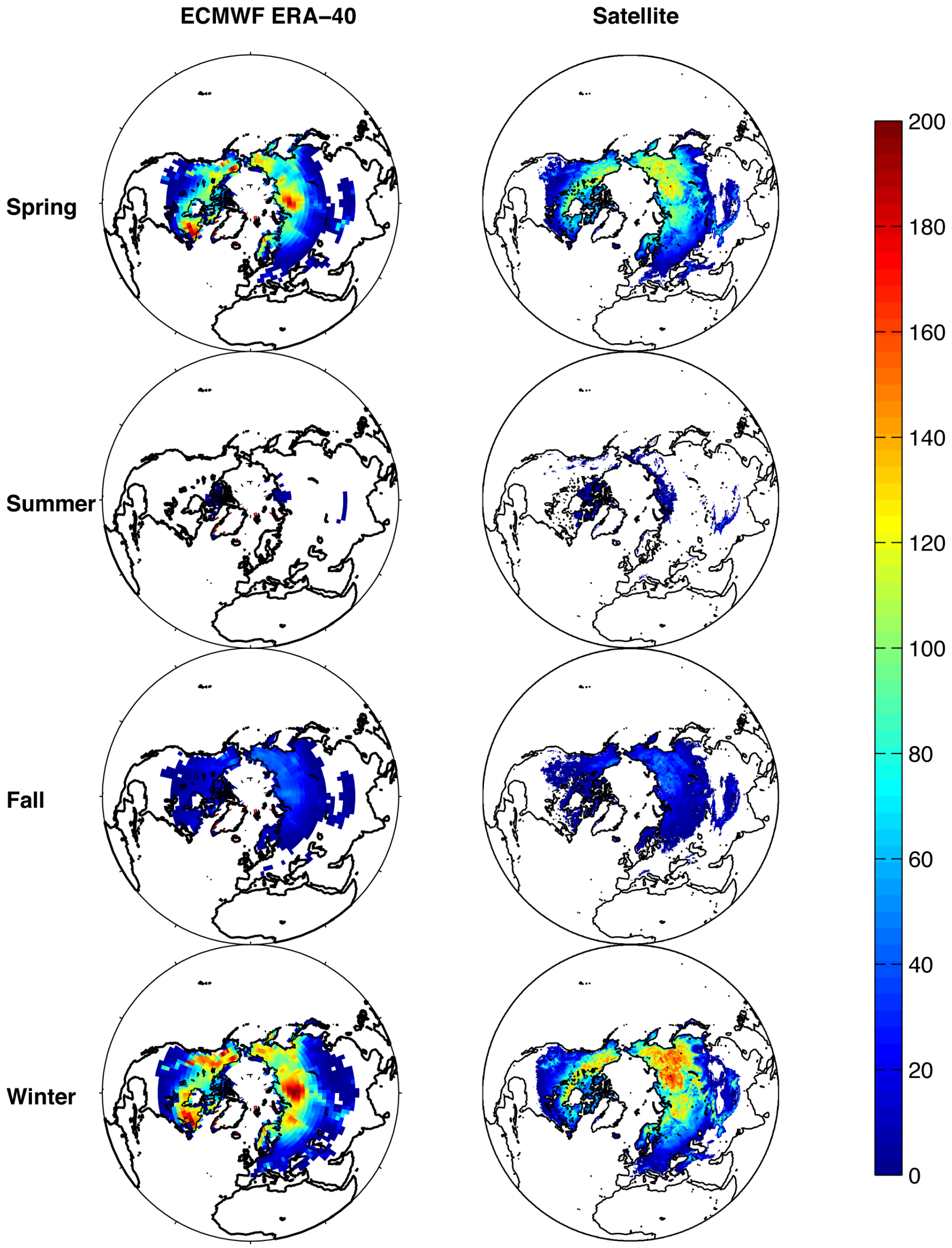


Figure 3.6: Seasonal mean snow water equivalent (mm) from the ECMWF ERA-40 reanalysis and satellite observations.

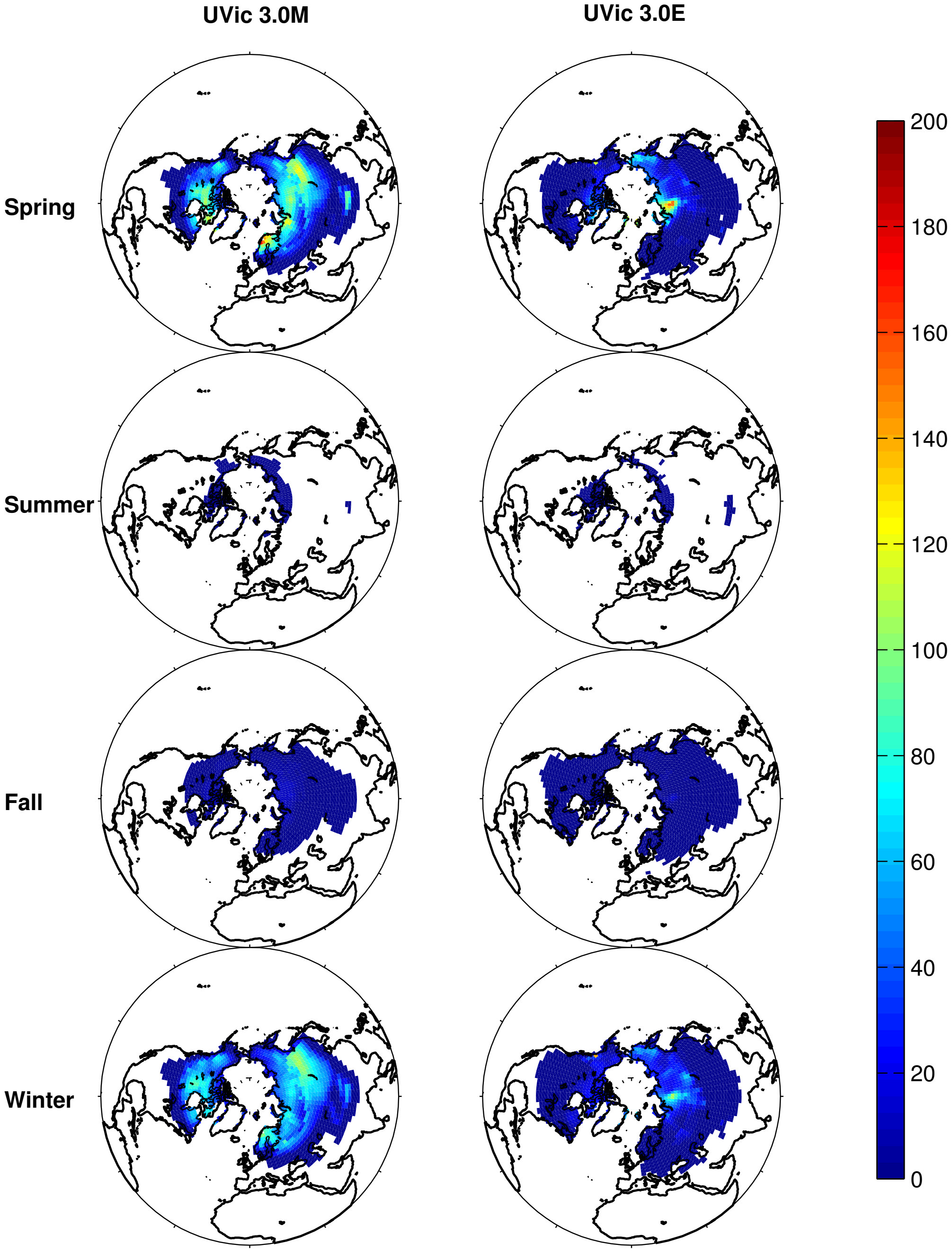


Figure 3.7: Seasonal mean snow water equivalent (mm) in the coupled UVic model 3.0M (U3.0M) and UVic 3.0M driven by ERA-40 forcing data (U3.0E).

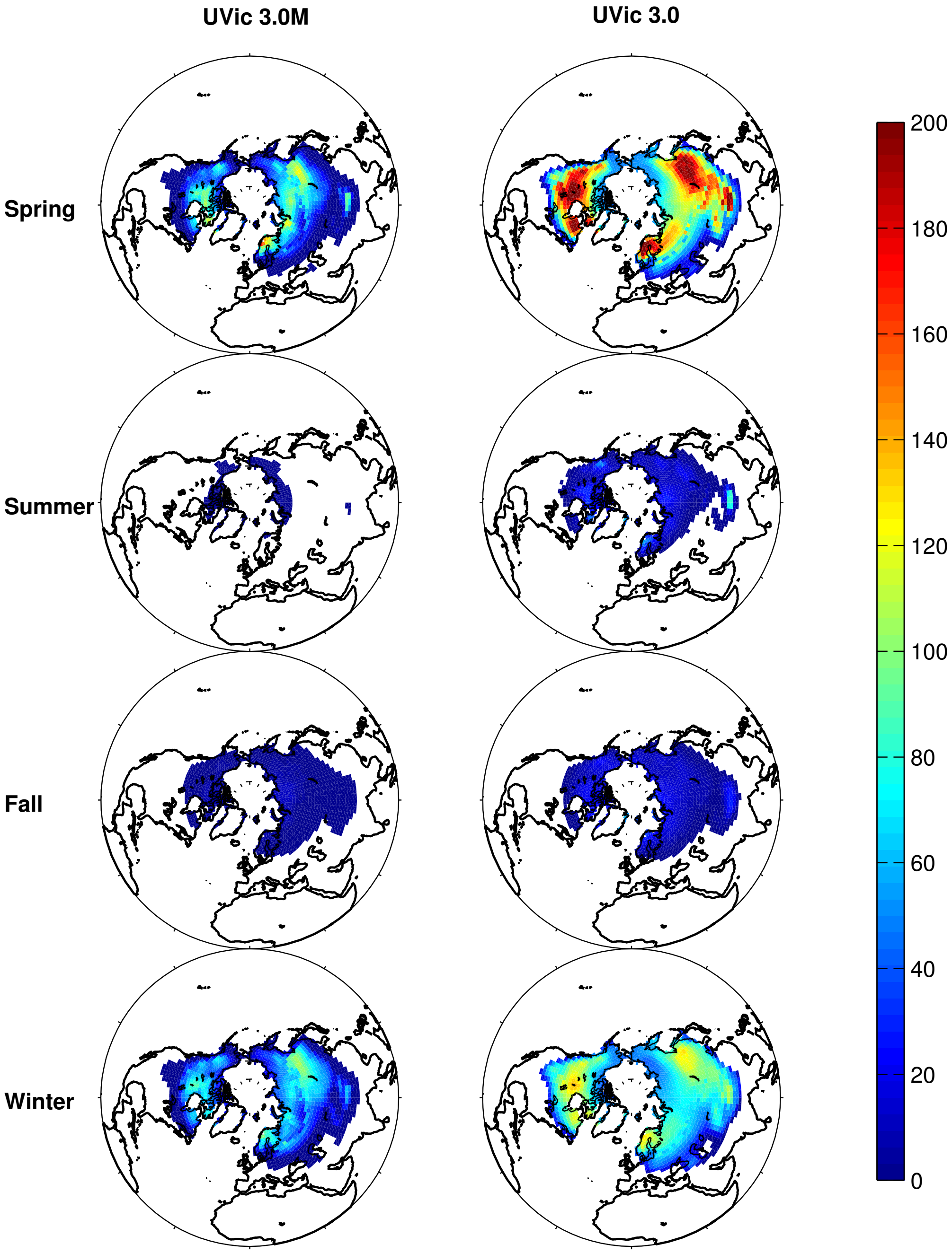


Figure 3.8: Mean snow water equivalent (mm) in the UVic model 3.0M and UVic 3.0.

3.0 of the model were examined - modifications in the model snow code from U3.0 to U3.0M were detailed in Chapter 2. It should be recalled as well that annual cycle in high-latitude SAT and precipitation was found to be very similar between U3.0 and U3.0M. Hemisphere mean SWE values in U3.0 (panel c, Figure 3.5) are consistently very high at all times of the year (Figure 3.8) relative to all other displayed SWE plots. The excessive SWE values in U3.0 might be expected given that the model tends to simulate too-high precipitation values in the circum-Arctic fall and winter (Figure 3.4). A consequence of these high SWE values, there appears to be a delayed retreat of snow in U3.0 likely due to the energy required to melt the extra snow. Panel d) of Figure 3.5 indicates that U3.0 Northern Hemisphere SCA agrees well with observations during the snow accumulation season (Aug-Jan), but SCA values lag observations from March-Jul as the snow retreats.

These observations may be explained if the snow scheme in UVic 3.0M simulates too much snow ablation. That is, it seems that UVic 3.0M simulates too much sublimation or snow melt when forced with unbiased atmospheric data, explaining the abnormally low SWE values simulated by U3.0E. However, the heightened snow ablation rates appear to compensate for the too-high precipitation rates in during the fall and winter in U3.0M resulting in a reasonable distribution of snow cover and SWE despite the atmospheric biases. Given this reasonable snow distribution in U3.0M and potential large amount of work involved in further revising and testing the snow cover scheme in the model, it was decided that further work on the snow component of the LSS was outside of the scope of this thesis. However, if a future version of the UVic ESCM is developed that strongly reduces the precipitation biases identified in this work, then the snow scheme will likely need to be adjusted, lest too little snow mass be simulated in the model.

## 3.2 Permafrost and Seasonally Frozen Ground

The permafrost and seasonally frozen ground distribution and characteristics in U3.0M and U3.0E are now compared against available observational datasets. At the outset, it is useful to outline the possible reasons for discrepancies between model and observational data.

1. **Atmospheric variables**: Biases in atmospheric state variables in the UVic ESCM have already been discussed and apply to U3.0M, under the assumption

that the atmospheric fields driving U3.0E realistically represent the state of the atmosphere in high latitudes. Temperature biases impact the heat flow into the ground and precipitation biases affect the simulated snow cover distribution, soil moisture distribution as well as influencing the surface energy balance by altering the amount of energy that is absorbed in evaporation and sublimation. The soil moisture distribution in turn affects the soil heat capacity and the rate of soil freeze/thaw through latent heat effects.

2. **Land surface scheme processes and constants**: Soil and bedrock characteristics in circumpolar regions remain poorly constrained. While the best available soil data suitable for inclusion in a global climate model was used in this study, there are certain to be discrepancies at local and regional scales between the characteristics specified in the model and the true characteristics. As well, land surface processes in the model may not be properly simulated. Snow processes almost certainly fall within this category as evidenced by the fact that the snow distribution is simulated reasonably well in U3.0M but not in U3.0E. The simulation of vegetation in the model is also not ideal as a consequence of the use of only five plant functional types to simulate the full diversity of terrestrial vegetation as well as on the the impact of atmospheric and subsurface biases on the distribution of vegetation. Vegetation influences the surface energy budget through its impact on terrain roughness, surface albedo and control on transpiration. Subsurface processes such as heat and moisture transport through the soil and soil freeze-thaw too are likely handled in an overly simplified manner but these processes are even harder to assess given the lack of reliable global scale measurements of soil moisture content and temperatures.

3. **Small scale variability**: The Arctic terrain is characterized by small-scale heterogeneity in vegetation, topography and snow cover amongst other factors and these can have significant impacts on the state of the soil. As an example, it has been demonstrated (Hinkel and Nelson (2003); Nelson and Shiklomanov (1999)) that active layer thickness can vary substantially on scales as small as a meter owing to such heterogenity. Consequently, it is unclear whether available point measurements of soil variables such as temperature or active layer thickness are representative of typical regional characteristics. Consequently, when comparing model output against such point measurements, there is a scale mismatch. The model simulates mean characteristics over a large geographic region

which may encompass as few as one or two point measurements.

4. **Temporal coverage**: Available frozen ground data does not span a consistent temporal range. For instance, the only pan-Arctic permafrost map includes data from the mid-1960s to the early 1990s. In recent years, a relatively large network of ground thermal monitoring stations has been established under the Circumpolar Active Layer Monitoring program. Data from these stations began to become available beginning in 1990 with some stations offering data coverage since then, and others with data from only one year in this period. I decided to use model runs from the previous section which spanned the years from 1957-2002 in the case of U3.0M and the output of the equilibrated U3.0E, rather than attempting to separately compare the models against different time slices of the latter half of the 20th century to match the all of observational data. This mismatch in temporal coverage between data and observations will undoubtedly account for a slight amount of the discrepancy between model and observations as observed active layer thicknesses and soil temperatures have generally been observed to slightly increase at ground monitoring sites over the latter half of the Century.

5. **Relict permafrost**: As described in Chapter 1, relict permafrost is permafrost that is not in equilibrium with the climate and may yet be warming from the Little Ice Age or the Last Glacial Maximum. If relict permafrost were allowed to fully equilibrate with the present climate, then it would warm, possibly become thinner or even vanish in its entirety in a region depending on the degree of disequilibrium. It is not known how much permafrost is relict; certainly large amounts of the permafrost in the sporadic and isolated zones is relict and Williams and Smith (1989) and French (2007) suggest that large portions of the discontinuous permafrost zone too are relict. Simulation of such relict permafrost would be impossible barring very long transient permafrost simulations beginning at LGM or longer ago and therefore the presence of relict permafrost remains an unquantified uncertainty in assessing the permafrost distribution in models. However, a model that simulates permafrost starting from preindustrial conditions should simulate too little permafrost, if anything, compared to the total permafrost area as it is unable to adequately represent relict permafrost.

Evidently, it is impossible to thoroughly assess the impact of all of these factors on the modelled frozen ground distribution. Where model discrepancies are known, these will be addressed in comparing model output to observations. To further illuminate some of the land surface scheme factors that may govern the modelled permafrost distribution and are poorly constrained by observations, a sensitivity analysis is conducted at the end of this chapter.

### 3.2.1 Areal Distribution of Permafrost

At present, the only hemisphere-scale map of permafrost characteristics is the circum-Arctic map of permafrost and ground ice conditions produced by the International Permafrost Association (Brown et al. (1998)). The map was developed by assimilating available maps and other relevant data on permafrost characteristics - maps that are specifically cited extend from the mid 1960s to the early 1990s (Heginbottom et al. (1993)) and so the resulting IPA map is treated as a best estimate of the state of Permafrost over the latter half of the $20^{th}$ Century. The digital form of the IPA map includes information on permafrost zonation, ground ice content and thickness of overburden. Permafrost zonation (plotted in Figure 3.9) describes the extent of the ground underlain by permafrost and reflects the fact that the high degree of small-scale heterogeneity in ground characteristics in the Arctic can have large impacts on the depth-to-permafrost and the presence or absence of permafrost at these small scales.

Zonation is specified in terms of *continuous* (90-100 % of ground underlain by permafrost), *discontinuous* (50-90 %), *sporadic* (10-50 %) and *isolated* (0-10 %) permafrost regions and the zonation from the IPY map is shown in Figure 3.9. Zhang et al. (1999) calculated the area of the permafrost zones from the IPA map and Zhang et al. (2000) calculated the ground area within permafrost zones actually underlain by permafrost (determined by multiplying the permafrost zone area by the percentage of ground underlain by permafrost within each zone); these results are all reproduced in Table 3.1. Zhang et al. (2000) argue that the ground area actually underlain by permafrost (as opposed to the area occupied by permafrost zones) should be the metric used in assessing models of permafrost.

Figure 3.9 also shows the distribution of seasonally frozen ground from the work of Zhang et al. (2003) who established a relationship between fields of mean monthly air temperature and the number of days that soil is frozen ground based measurements. In their work, regions of seasonally frozen ground are identified as those experiencing more than 15 days of freeze per year, while intermittently frozen ground consists of

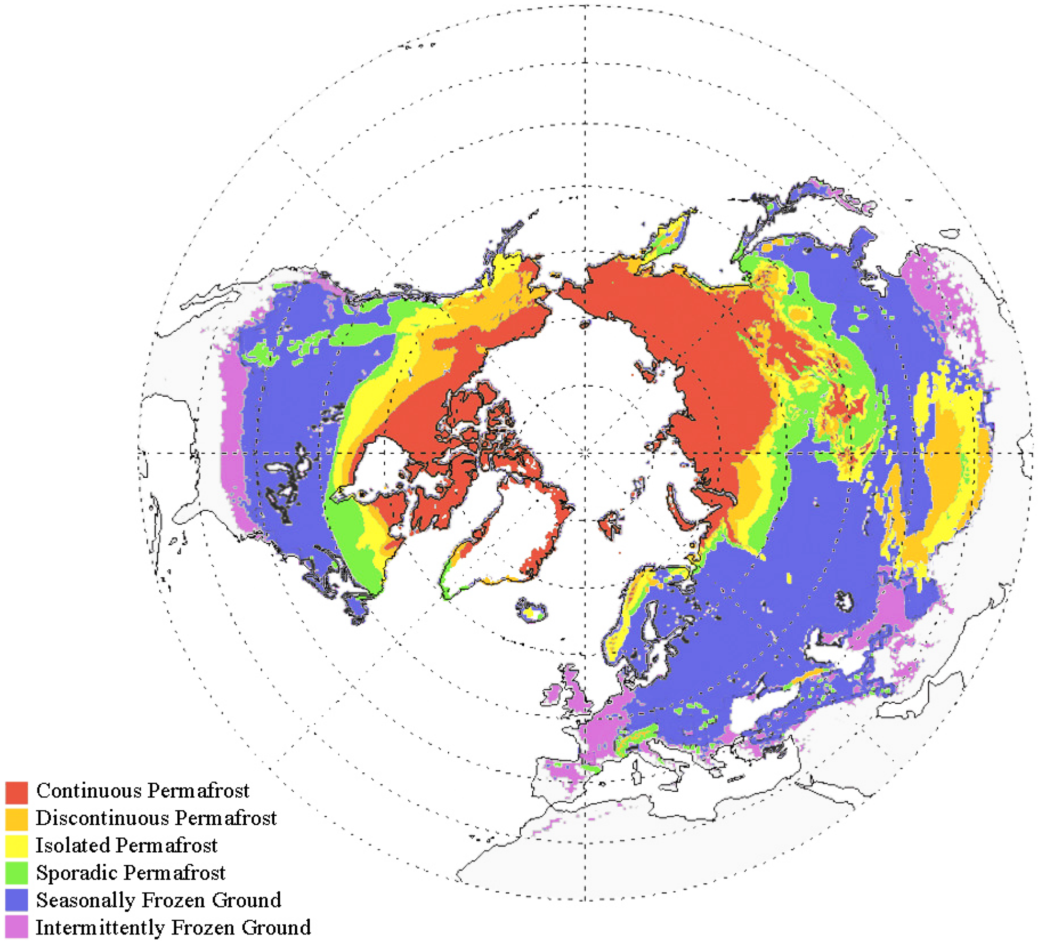


Figure 3.9: Circum-arctic map of permafrost and seasonally frozen ground. This map is a composite of two separate maps: a map of permafrost zonation from the digital IPA permafrost map (Brown et al. (1998)) produced by Zhang (2005) and a map of regions of permafrost and seasonally frozen ground produced by Zhang et al. (2003). Maps were composited together by the author using photo editing software as Zhang indicated that data from Zhang et al. (2003) had been lost.

| **Permafrost (PF) Zone** | **PF Zone Area** | **Min. PF Area** | **Max. PF Area** |
|---|---|---|---|
| Continuous | 10.69 x $10^6$ km$^2$ | 9.62 x $10^6$ km$^2$ | 10.69 x $10^6$ km$^2$ |
| Discontinuous | 4.38 x $10^6$ km$^2$ | 2.20 x $10^6$ km$^2$ | 3.95 x $10^6$ km$^2$ |
| Sporadic | 3.90 x $10^6$ km$^2$ | 0.39 x $10^6$ km$^2$ | 1.96 x $10^6$ km$^2$ |
| Isolated | 3.82 x $10^6$ km$^2$ | 0 km$^2$ | 0.38 x $10^6$ km$^2$ |
| Total | 22.79 x $10^6$ km$^2$ | 12.21 x $10^6$ km$^2$ | 16.98 x $10^6$ km$^2$ |

Table 3.1: Area of permafrost zones and estimates of the minimum and maximum ground areas underlain by permafrost in the Northern Hemisphere. Data was produced by Zhang et al. (1999) and Zhang et al. (2000) working from the digital IPA permafrost map (Brown et al. (1998)).

soil that experiences between 1 to 15 days of freeze per year. The authors include the active layer above permafrost as part of the seasonally frozen ground and identify the maximum Northern Hemisphere extent of seasonally frozen ground as 48.12 x $10^6$ km$^2$, of intermittently frozen ground as 6.27 x $10^6$ km$^2$ and of ice sheets as 2.12 x $10^6$ km$^2$.

| **Region** | **Area: U3.0M** | **Area: U3.0E** |
|---|---|---|
| Ice sheets | 2.46 x $10^6$ km$^2$ | 2.46 x $10^6$ km$^2$ |
| Permafrost | 15.59 x $10^6$ km$^2$ | 23.01 x $10^6$ km$^2$ |
| Seasonally Frozen Ground | 47.92 x $10^6$ km$^2$ | 51.46 x $10^6$ km$^2$ |
| Intermittently Frozen Ground | 2.09 x $10^6$ km$^2$ | 2.52 x $10^6$ km$^2$ |

Table 3.2: Area of cryotic regions in the Northern Hemisphere as simulated by coupled UVic 3.0M (U3.0M) and ERA-40 driven UVic 3.0M (U3.0E) configurations. Ice sheet areas are not simulated but are specified from data.

The distribution of frozen regions in U3.0M is plotted in the top panel of Figure 3.10 and the areal extent of these regions is presented in Table 3.2. There were slight changes in these regions simulated by U3.0M from 1957-2002, consequently a region is identified as being part of a given zone in the figure if it occupies this zone for more than 50 % of this time interval. Given that mean ground conditions are simulated in a grid cell, it is suggested that the model should ideally be able to reproduce continuous and discontinuous permafrost (C-D permafrost - where permafrost is present in more than 50 % of terrain in a region) zones, but not sporadic or isolated permafrost zones. Comparing the top panel of Figure 3.10 with Figure 3.9, it is found that there is generally fairly strong agreement between regions of C-D permafrost and the U3.0M

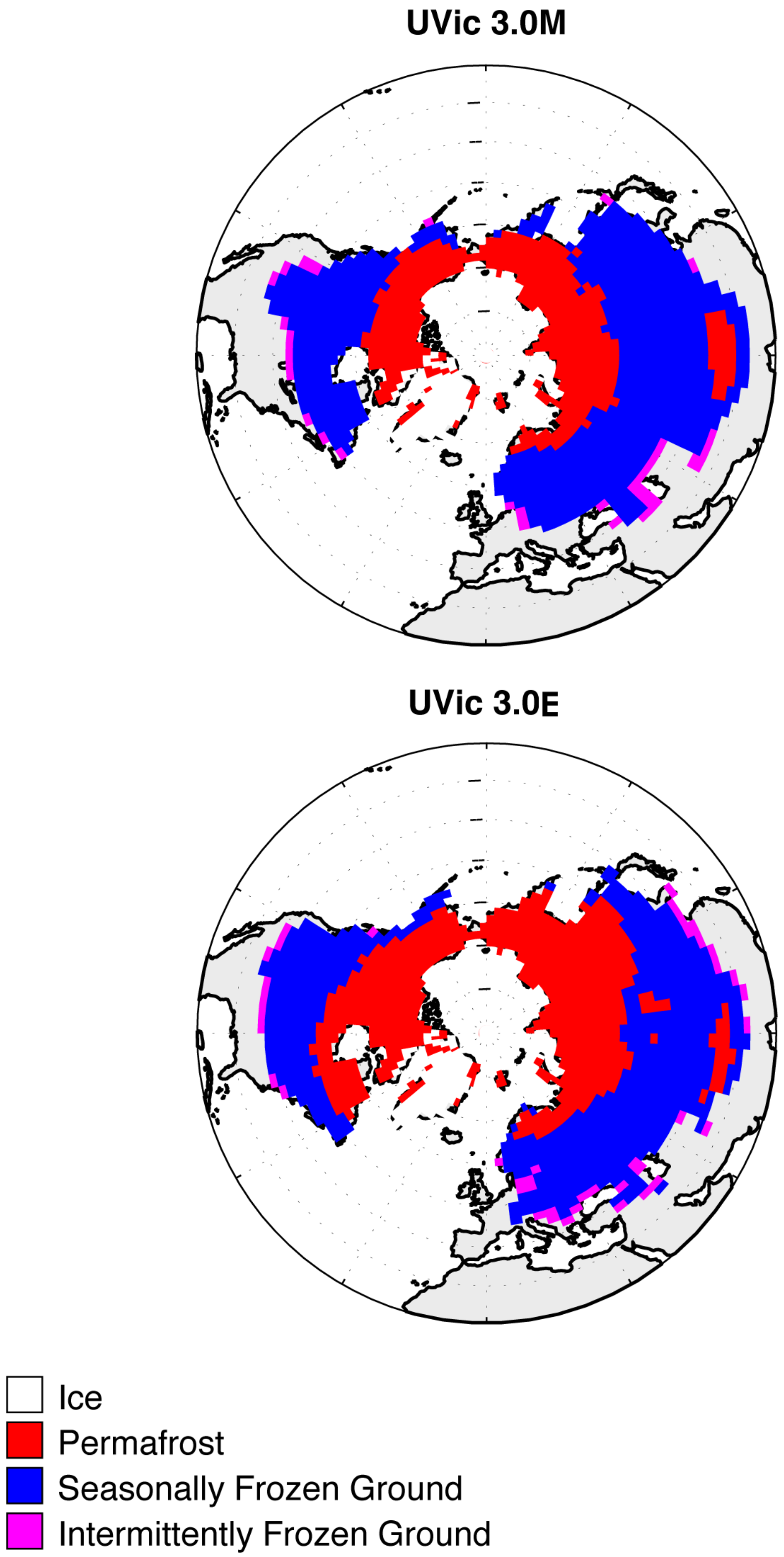


Figure 3.10: Regions of ice, permafrost, seasonally frozen ground and intermittently frozen ground in UVic 3.0M and UVic 3.0E . The ice sheet distribution is specified from data and not simulated. Regions are identified based on the criteria in Zhang et al. (2003).

permafrost. However, the simulated permafrost does not extend as far south in Central/Eastern Canada (eg. there is no simulated permafrost in the vicinity of Hudson's Bay or in Quebec) nor in Eastern Siberia (eg. no permafrost in the South Eastern portion of the Central Siberian plateau near the Russia-Mongolia and Russia-China borders). Simulated permafrost extends too far to the south in Western and Central Eurasia (eg. there is simulated permafrost in regions of isolated/sporadic permafrost and seasonally frozen ground in the Nordic countries and in Russia, extending from the longitude of the White Sea to the Central Siberian Plateau). I suggest that these discrepancies are, to a large extent, linked to the temperature biases in the model as regions where simulated permafrost is absent agree well with the southern limits of warm air temperature biases in the model, while regions where permafrost is too extensive are associated with cold SAT biases in the model (Figure 3.2).

The total areal extent of permafrost in U3.0M (15.59 x $10^6$ km$^2$) agrees fairly well both with the area of the Continuous and Discontinuous permafrost zones (14.98 x $10^6$ km$^2$) as well as the total area of ground containing permafrost (12.21 - 16.98 x $10^6$ km$^2$). However, it should be noted that the ESCM (both in U3.0M and U3.0E configurations) over-represents the exposed land mass in the Northern Hemisphere. This is a consequence of the fact that the ESCM does not simulate lakes and gridcells containing a large number of small-scale land features such as islands are represented as being completely filled with land. The net result is an over estimate of exposed land area within permafrost regions. To assess the extent of this over estimate, regions containing permafrost in the Northern Hemisphere were identified from the IPY digital permafrost map (Figure 3.11, top panel). This map was then interpolated to match the grid of the UVic model (Figure 3.11, bottom panel). Regions containing lakes or ocean points in the IPY map corresponding to land-filled cells in the ESCM were marked as bad data and missing values were determined by extrapolation from nearby neighbouring cells. The IPY map indicates that permafrost zones encompass 22.76 x $10^6$ km$^2$ of Northern Hemisphere landmass contains permafrost. Using the IPY map interpolated to the model grid to calculate extent of permafrost zones yields a value of 25.66 x $10^6$ km$^2$ of exposed permafrost bearing landmass in the Northern Hemisphere, a 12.7 % over-estimate. Thus, the simulated permafrost extent figures cited in this work should be understood to be too high by a factor on the order of 10 %.

The model's simulated extent of seasonally frozen ground (SFG) is calculated to be 47.92 x $10^6$ km$^2$, a slight underestimate (considering the above-mentioned model

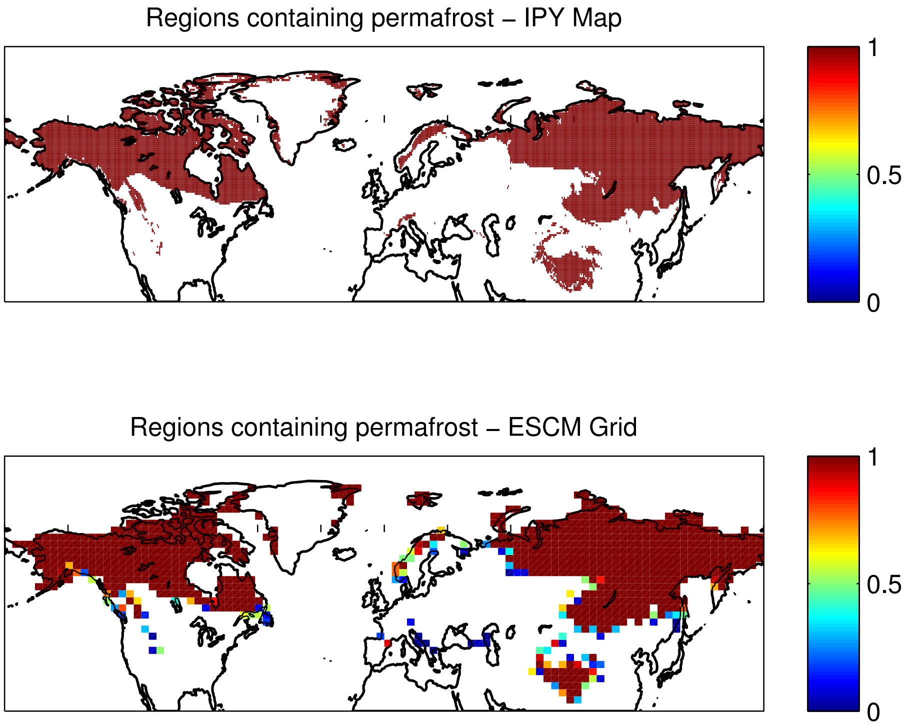


Figure 3.11: Depiction of the overestimation of exposed land surface area in the UVic ESCM model grid. The top panel shows the areal extent of permafrost from the International Polar Year permafrost map in terms of fractions of grid cells containing permafrost. In the case of the IPY map regions either contain permafrost of some form (fraction = 1) or contain no permafrost (fraction = 0). Though it may not be apparent given the scale of the plot, extensive inland water bodies are not represented as containing permafrost in this map. The bottom panel shows the result of interpolating this map to the UVic model grid. The ESCM LSS does not include a representation of lakes and some high latitude gridcells contain more land mass than is actually present. Both of these factors result in the model over representing exposed Northern Hemisphere landmass.

grid setup) of the value 48.12 x $10^6$ km$^2$ cited by Zhang et al. (2003). The simulated Intermittently Frozen Ground (IFG) value of 2.09 x $10^6$ km$^2$ is substantially lower than Zhang et al. (2003)'s value of 6.27 x $10^6$ km$^2$. Both SFG and IFG do not extend sufficiently far to the south nearly everywhere in the Northern Hemisphere. Here again, the model's SAT anomalies are likely a major factor as much of the southern limit of SFG and IFG North America and Eurasia lie in regions of strong winter warm biases (bottom panel, Figure 3.2) that extend to the southernmost limits of SFG and IFG regions. It should also be remembered that the SFG and IFG results presented by Zhang et al. (2003) are estimates produced based on fields of monthly mean air temperature, rather than direct observations and so Zhang et al. (2003)'s results may not reflect the true distribution of SFG/IFG entirely accurately.

The distribution of cryotic regions at the end of U3.0E's equilibration run is shown in the bottom panel of Figure 3.10. U3.0E simulates permafrost nearly everywhere C-D permafrost is present including some geographically isolated patches such as in central Mongolia. However, the model simulates extensive permafrost in regions of sporadic and isolated permafrost and seasonally frozen ground. This is reflected in the total areal extent of permafrost in U3.0E of 23.01 x $10^6$ km$^2$ which is comparable to the total areal extent of all permafrost zones (22.79 x $10^6$ km$^2$) but is substantially higher than the estimate of maximum ground area actually underlain by permafrost (16.98 x $10^6$ km$^2$). U3.0E simulates a SFG extent of 51.46 x $10^6$ km$^2$ (compared with 48.12 x $10^6$ km$^2$ from Zhang et al. (2003)). A visual inspection of Figures 3.9 and 3.10 indicates that U3.0E does a much better job at simulating SFG than U3.0M although regions of IFG still do not extend sufficiently far to the south in many areas.

It may at first seem puzzling that areas of SFG are well reproduced by U3.0E but the simulated permafrost distribution is far too extensive. The answer lies in the distribution of snow in U3.0E. I have previously shown that U3.0E generally does a good job of reproducing snow covered areas throughout the year, but consistently underestimates snow thickness. Snow is a powerful thermal insulator with the degree of this thermal insulation dependent on snow thickness. If snow thickness were too low during winter, then the ground would cool at an unrealistically high rate. This heightened rate of cooling would be most pronounced in areas where the snow insulation effect should be greatest (i.e. where the snow is thickest). Thus, simulated regions of permafrost are too extensive, but regions of SFG (where snow is thin and the insulation effect is weak) are reasonably reproduced. It remains unclear why IFG is poorly captured by U3.0E. One possibility is that the atmospheric data driving the model is specified as monthly mean fields, with the surface air temperature supplied to the model determined by interpolation between monthly fields. As the presence of IFG requires that the ground be frozen for at least one day, it is possible that the model's air temperature fields do not allow it to capture very short freezing intervals associated with the coldest winter temperatures at a particular grid cell.

### 3.2.2 Active Layer Thickness

I next compare simulated active layer thicknesses (ALT) against observations. While active layer measurements have been made throughout the $20^{th}$ Century, the number of measurement sites has greatly expanded in recent years through the establishment of the Circumpolar Active Layer Monitoring program - one of the first large scale efforts at monitoring the global state of permafrost (Brown et al. (2000)) that was established in the early 1990s. Data from CALM sites is available from 1990 - 2010; not all sites have measurements available for this full period - at some sites measurements are available for only a single year, while other sites span the full 21-year period. The Canadian government also maintains an active layer monitoring program that includes additional sites that are not part of the CALM program. Active layer monitoring from these Canadian non-CALM sites is available from 1991-2007 (Nixon (1998. Updated 2009.)). Both the CALM and Canadian ALT datasets were retrieved from the National Snow and Ice Data Center website. To compare these measurements against the simulated ALTs, I first calculated a mean ALT for each site from all available years' data at that site. While this approach is not ideal as mean ALT measurements span varying numbers of years, restricting the measurement comparison to years when data was available from all sites would have substantially reduced the number of sites available for comparison with the model. Figure 3.12 shows a plot of the mean ALTs at each site location. While the CALM site network is quite extensive by the standards of the permafrost community, it is clear from the plot that sampling is relatively limited compared to the expansive Northern Hemisphere permafrost regions.

The ALTs simulated by U3.0M and U3.0E are shown in Figure 3.13. The ALT distribution in regions in which both model configurations simulate permafrost is fairly similar, although ALTs at the southern margin of permafrost in U3.0M are typically greater than corresponding regions in U3.0E as are ALTs in the region bordering the Arctic Ocean from the Northwest Territories to the west coast of Greenland -

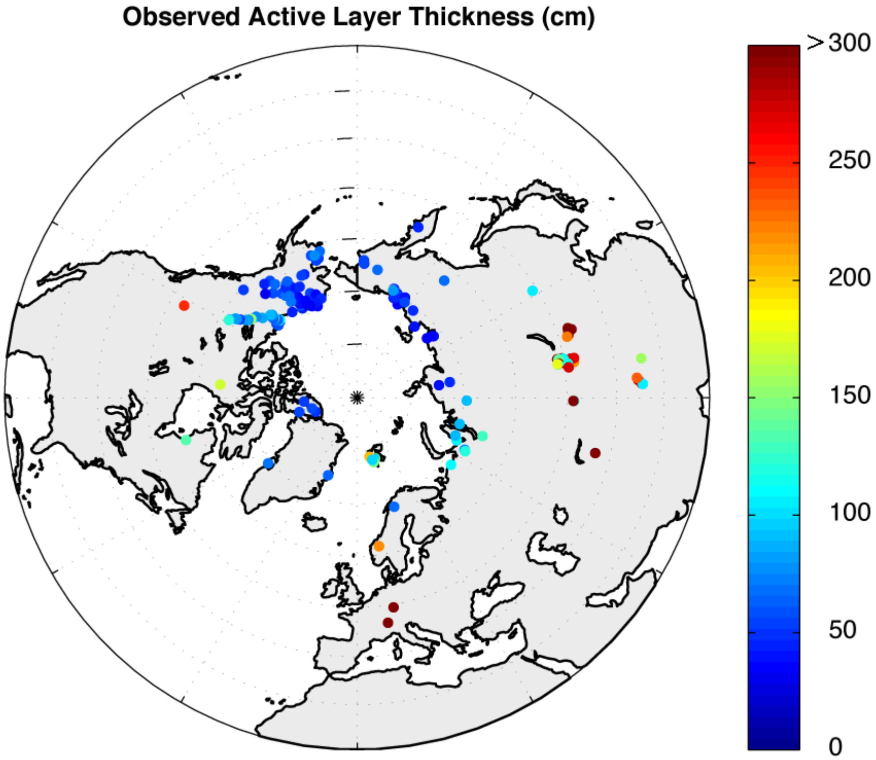


Figure 3.12: Active layer thicknesses from the Circumpolar Active Layer Monitoring Program with additional Canadian Arctic and Subarctic Sites. The ALTs were determined as the mean ALT from all available years of data during the interval from 1990-2010.

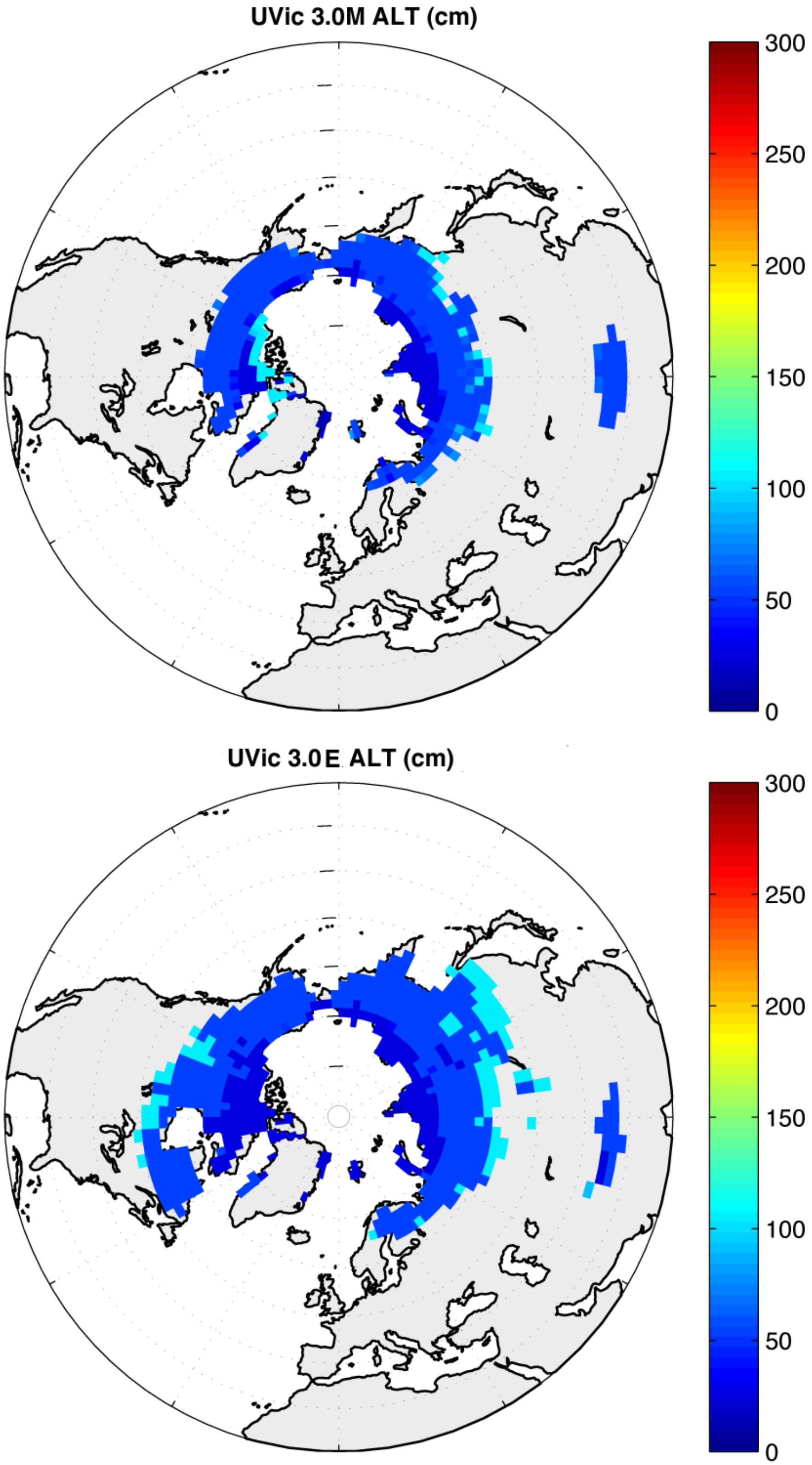


Figure 3.13: Simulated active layer thicknesses. Top panel: Mean ALT from 1957-2002 in UVic 3.0M. Bottom panel: ALT at the end of equilibration run of UVic 3.0E.

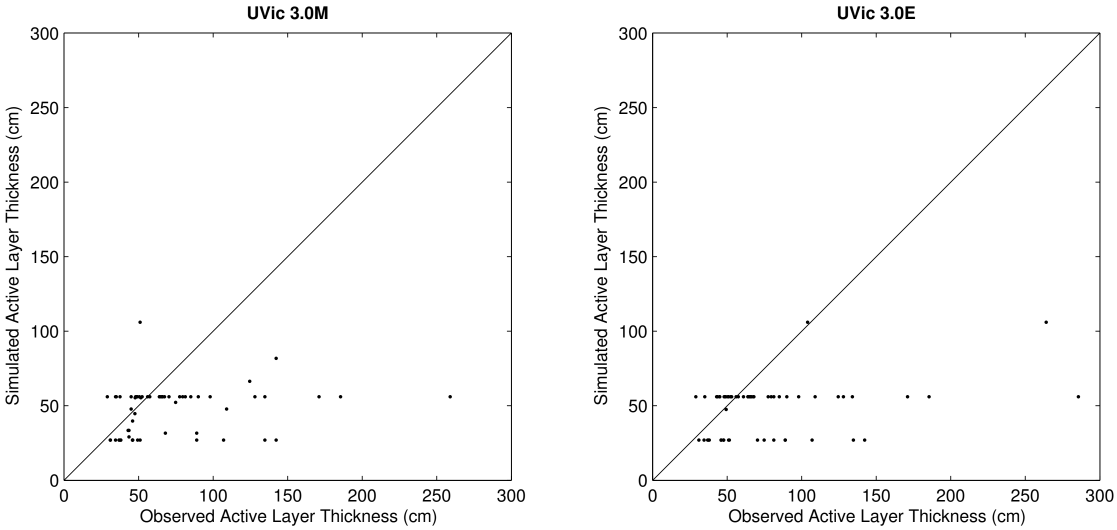


Figure 3.14: Scatter plot comparing active layer depths in UVic 3.0M and 3.0E. As UVic 3.0E configuration simulates a more extensive permafrost area, there are more plot points in this plot. The ALTs in the coupled model (U3.0M) are mean ALTs from 1952-2002 while the ALTs in the ERA-40 forced model (U3.0E) are the ALTs at the end of the equilibration run.

these are again likely due to the U3.0M warm biases and U3.0E snow thickness issues discussed earlier. Both models appear to crudely reproduce the predominantly latitudinal variation in ALT shown in Figure 3.12.

To better compare the ALTs from the point measurements against the model, a scatter plot of simulated vs. observed ALTs was produced. A mean observation-based ALT measurement was produced for each model grid cell by identifying CALM and Canadian sites that fell within a particular grid cell and then averaging the ALT values for these sites; there ranged between 1 and 36 observation sites per model grid cell. In cases where there was more than one site per grid cell there was often a considerable degree of variability between sites: typically the standard deviation of the ALT values that were averaged was about 30 % of the mean value. The mean observed value was then plotted against the corresponding simulated ALTs as shown in Figure 3.14.

In part, some of the discrepancy between simulated and observed ALTs must be due to scale incompatibilities: I am implicitly treating the average of a small number of ALT observations as though it is representative of mean ALT conditions over a region the size of a model grid cell. Given the small sample size of ALT observations within this regions, it is inevitable that the calculated ALT for the region is biased relative to the true mean ALT in this region. Nonetheless, despite this scale mismatch, it is clear that ALTs in both configurations of the model are generally shallow compared with observations. This indicates that there is either a selection bias in the CALM sites towards sites with deep ALTs within a region (unlikely given the number of points plotted) or that some aspect of the model results in ALTs that are too shallow.

It should be noted that the ability of the model to realistically simulate active layer thicknesses is restricted by its relatively coarse vertical resolution. The simulated ALTs are identified as the bottom of the deepest soil layer that fully thaws annually. Consequently, the model will tend to underestimate the true active layer thickness as the simulated ALT can only increase in incremental steps. This is reflected in Figure 3.14 by the fact that simulated active layer thicknesses are clustered in horizontal bands corresponding to the depth of the bottom of soil layers in the model. Thus, the true ALT should lie between the depth of the bottom of the deepest thawed layer and the bottom of the layer below that. A version of the model with higher vertical resolution is presented in a subsequent section and shows better agreement with observations for this reason.

Another factor in the discrepancy between simulated and observed ALTs is the temporal coverage of the model runs (1957-2002) relative to the observations (1990-2010). To examine this, the coupled version of the model forced from a year-1800 equilibrium state through to 2011 and the model's simulated 1990-2011 ALTs were compared with observations (not shown). While a few grid cells showed deeper ALTs and better agreement with observations, the vast majority of cells still showed too-shallow ALTs, indicating that the mismatch in temporal coverage is not the dominant cause of the discrepancy.

The fact that ALTs are too shallow in both U3.0M and U3.0E indicates that a factor common to both configurations of the model (such as the vertical resolution) is responsible for the discrepancy. Another possible explanation may be that the model exhibits excessive thermal dampening, as a consequence of the ground thermal diffusivity being too low, preventing the thawing front from penetrating as deeply into the ground as it should. If the ground thermal diffusivity were incorrect, this might have a large number of possible causes. For instance, soil dry thermal conductivities and heat capacities might not properly reflect the characteristics of Arctic soils or the variability in these characteristics with soil moisture content might be poorly parameterized. As well, if there was a bias in soil moisture content then this might have a major impact on the thermal conductivity and heat capacity (as well as on the seasonal cycle in these quantities). Lastly, the depth to bedrock may be incorrect. Bedrock layers have a much different thermal parameters and, in particular, do not contain any moisture that must freeze / thaw over the course of the year. Thus, if bedrock were consistently too deep in the model, then this might result in ALTs that are too shallow:

If the ground thermal dampening in permafrost regions is indeed too large in the model then this must be kept in mind when considering the transient response of permafrost to long term climatic warming as the permafrost might be expected to thaw too slowly in this case. Finally, it is noted that, while the discrepancies between the simulated and observed ALTs may appear to be fairly large, they are on the order of the discrepancies shown in other large land surface schemes that are capable of simulating permafrost (see Figure 2 in Dankers et al. (2011), and Figure 4 in Lawrence et al. (2011)), and this reflects the difficulty in realistically simulating permafrost in a global scale climate model.

### 3.2.3 Ground Temperatures

As part of International Polar Year (IPY - 2007-2009), a major effort was undertaken to enhance the network of permafrost ground observation stations. This included the installation or reactivation of a number of borehole sites to better monitor subsurface temperatures. As a result of this effort, a large number of ground temperature readings are available spanning 2007-2009 at variable depths including both shallow ($<$ 10 m) and deep ($>$ 10 m) sites. Sites for which there was at least one mean annual ground temperature measurement available from the years 2007-2009 are shown in Figure 3.15.

To compare the borehole temperatures with the model, I again identified all observation sites within a model grid cell, splitting the sites into shallow and deep boreholes. At each site, the mean annual ground temperature (MAGT) and the depth at which that temperature was recorded was specified. For each borehole depth category, I then determined a gridcell average depth at which temperatures were measured and average temperature reading. I retrieved mean annual ground temperatures from the models and determined the model temperature at a depth corresponding to the mean observational depth for the gridcell, interpolating between the model temperatures at

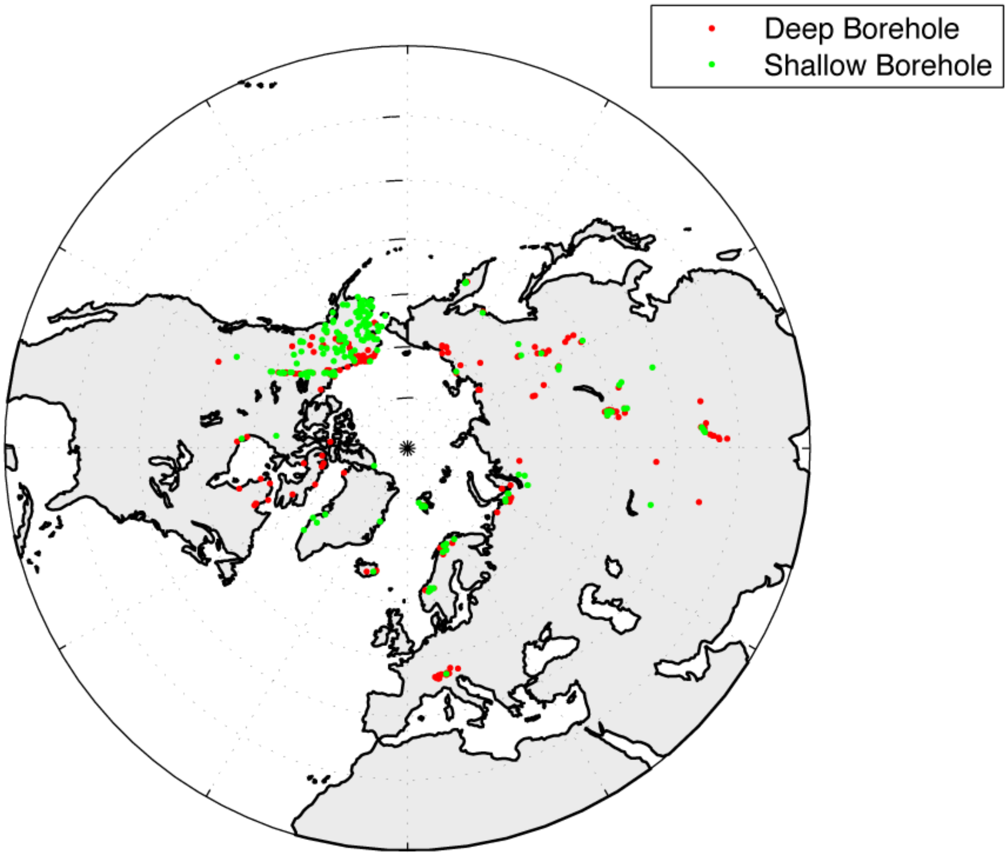


Figure 3.15: Location of borehole sites with ground temperature readings used in comparison with UVic ESCM. The borehole temperature readings were taken during the years 2007-2009. Sites marked shallow indicate temperature readings at depths of less than 10 metres, while deep boreholes indicate temperature readings at greater depths.

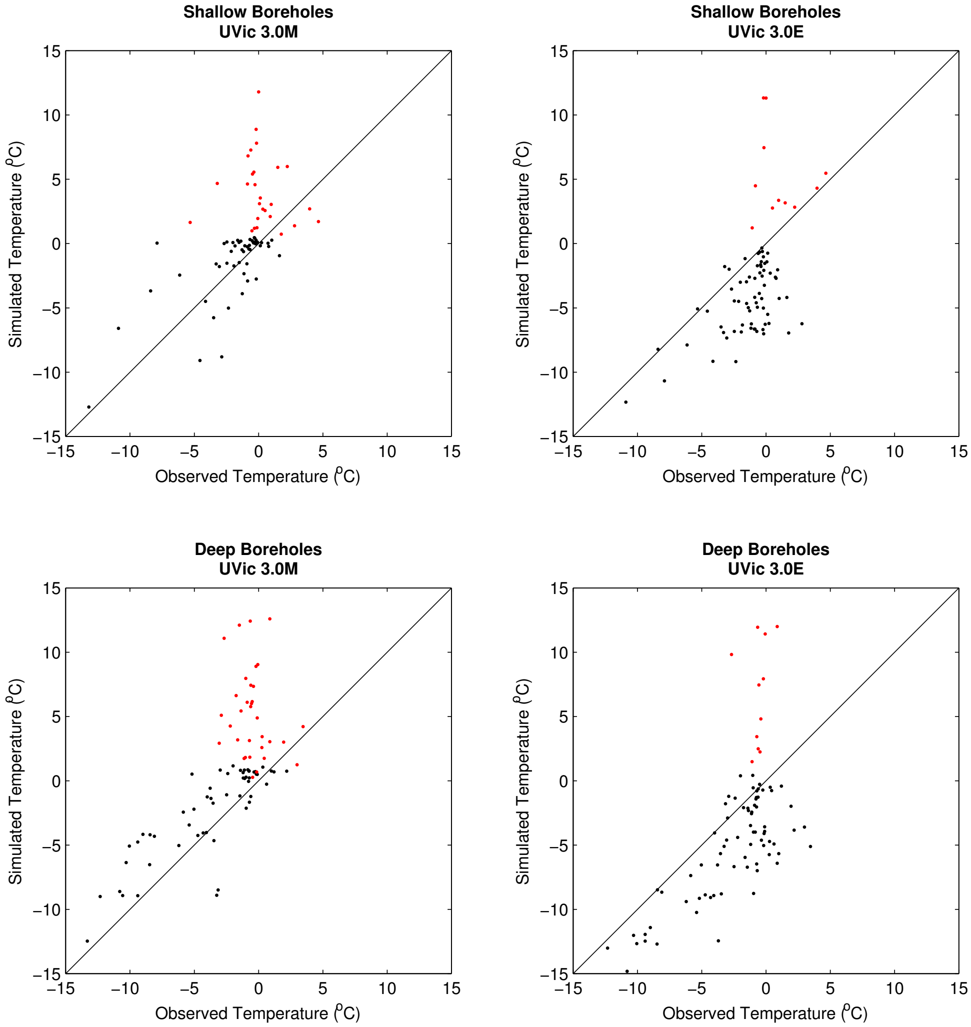


Figure 3.16: Scatter plot comparing mean annual ground temperatures in coupled UVic ESCM 3.0M and ESCM 3.0M driven by ERA-40 reanalysis data with readings from boreholes taken during International Polar Year (2007-2009). Boreholes are all located in permafrost regions. Black dots correspond to gridcells where the models simulate permafrost and red dots correspond to gridcells where permafrost is absent in the models.

soil levels bracketing the observed depth. The resulting interpolated model temperatures are plotted against the average MAGT from the sites for the gridcell in Figure 3.16.

MAGTs in U3.0M generally agree quite well with shallow borehole temperature readings, though are slightly warmer than average. When the analysis is restricted to sites where the model simulates permafrost (black dots), the model shows little overall bias when compared with shallow boreholes, and a slight warm bias when compared against deep boreholes. MAGTs are too high in U3.0M at nearly all sites where permafrost is not simulated in the model (red dots). Most of these sites are located within zones of isolated and sporadic permafrost. Hence, I suggest that the ground temperatures where the boreholes are located must be lower than the average ground temperature in the region and consequently, the simulated ground temperatures would appear high relative to the observations. In the case of deep borehole temperature readings, U3.0M again shows a slight warm bias, though there is again fairly good agreement with observations. The warm bias in both the near surface and deep soil temperatures are not surprising as most of the sites lie in regions where the model has a strong winter (and overall annual) warm bias (compare Figure 3.15 against Figure 3.1).

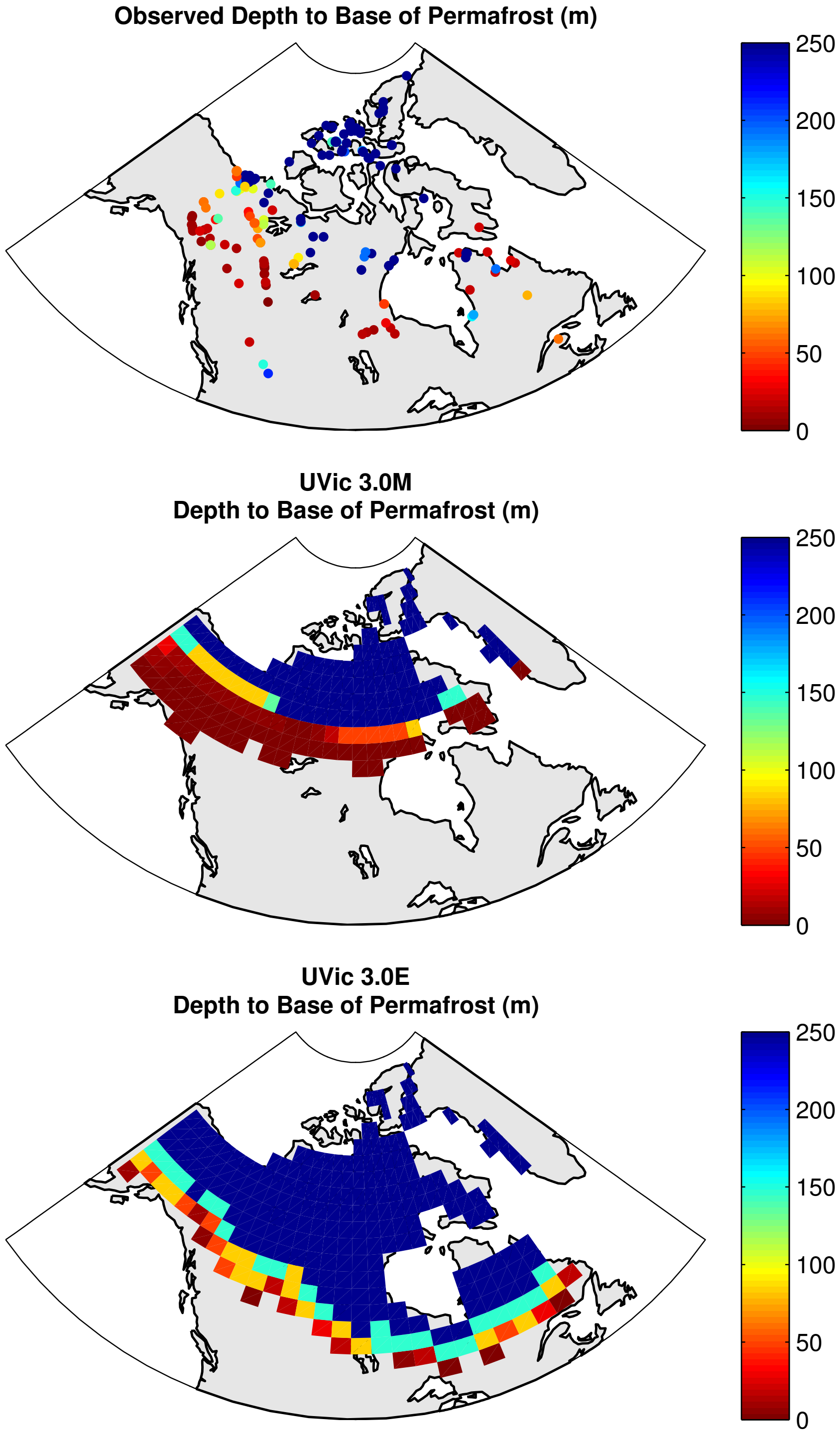
U3.0E on the other hand shows a stronger bias compared with observations and tends to be too cold (sites in sporadic and isolated permafrost - most of the red dots - not withstanding). This is presumably due to the same issue that produces excess permafrost in this model configuration: snow depths that are generally substantially too thin, resulting in excessive winter cooling. The fact that U3.0M and U3.0E show the opposite temperature bias compared to observations at both shallow and deep borehole sites, yet both have active layer thicknesses that are too shallow supports the notion that an issue with the dampening of the annual temperature cycle or vertical resolution is the cause of the active layer thickness bias.

### 3.2.4 Permafrost Depth

Permafrost thickness measurements have been made worldwide, but, as the compilation of all available thickness measurements would be an onerous task, I have opted not to examine this over the whole pan-Arctic domain. However, A comprehensive digital database of permafrost thickness has been produced for Canada by Smith and Burgess (2002) and includes two different measurements of the depth of the bottom boundary of permafrost. Most of the measurements in this database were made between 1960 and the mid-1980s. The base of permafrost is formally defined as the depth of the 0° C isotherm and this measurement is available at a number of the sites in the measurement database. The bottom boundary of the permafrost layer is also estimated at some sites based on measurements of the greatest depth at which ice is present in the ground. As a consequence of the suppression of the freezing point due to the effects of pressure and soil forces , this depth is shallower than the 0° C isotherm. Smith and Burgess (2002) indicate that given typical geothermal gradients, the base of ice-bearing permafrost may be between 40 to 275 m above the 0° C isotherm. As model layers contain permafrost if soil temperatures remain below 0° C, I opted to restrict the comparison to sites in the database where the permafrost bottom boundary is identified by temperature and not the absence of ice. Sites where the bottom boundary measurements were identified as 'unreliable' were also not included in the comparison.

The observed depths to the base of permafrost (DBP) from the database of Smith and Burgess (2002) are plotted in the top panel of Figure 3.17. The most striking characteristic of the plot is the relatively abrupt transition from sites where the DBP is relatively thick (>100 m) to sites where it is relatively thin (< 50 m). This transition zone roughly follows the spatial distribution of discontinuous permafrost in Canada (Figure 3.9). The simulated DBP in U3.0M is shown in the middle panel of the Figure. The model does a good job of simulating the correct order of magnitude of permafrost depth in both the thin and thick permafrost regions and also simulates the abrupt transition between the zones, although the transition region appears to be slightly too far to the north is this configuration. An equivalent plot for U3.0E is shown in the bottom panel. Here, the model simulates reasonable DBPs in its northern most and southernmost gridcells, but the transition zone between thick and thin permafrost is clearly displaced too far to the south, reflecting the cold sub-surface bias in this model configuration.

To examine the agreement between simulated and observed DBP in more detail, I produced scatter plots for the two configurations of the model (Figure 3.18). Where observations fell within a model gridcell containing permafrost, a mean gridcell DBP value was calculated following the same methodology as the active layer depth scatter plots. When examining the resulting plots, it is important to recall that the bottom boundary of the model occurs at a depth of 250 m; consequently, the model cannot reproduce DBP values deeper than this seen that occur in the database. As well



Observed Depth to Base of Permafrost (m)

UVic 3.0M
Depth to Base of Permafrost (m)

UVic 3.0E
Depth to Base of Permafrost (m)

Figure 3.17: Depth to the bottom of permafrost (m) in Canada. Top panel: depth measurements from the database of Smith and Burgess (2002). In many cases depths exceeded 250 m, but the colormap was restricted in this plot to facilitate comparison with the ESCM. Middle panel: mean depth to base of permafrost from 1957-2002 from UVic ESCM 3.0M. Bottom panel: depth to base of permafrost at the end of the equilibration run of UVic ESCM 3.0E.

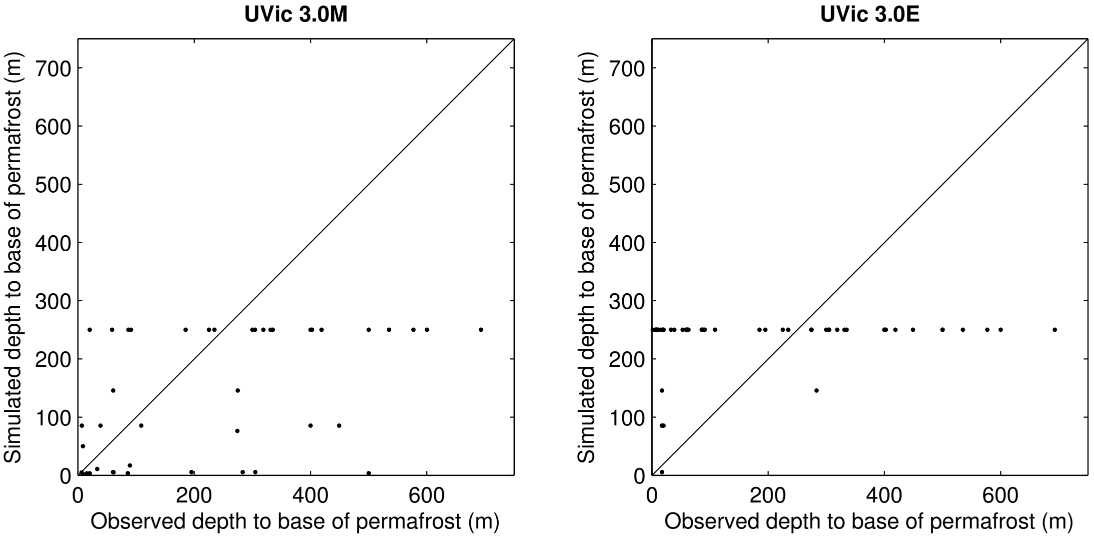


Figure 3.18: Scatter plot comparing the depth to the base of permafrost in coupled UVic ESCM 3.0M and ESCM 3.0M driven by ERA-40 reanalysis data with observations from the Canadian database of Smith and Burgess (2002). As the ERA-40 configuration simulates a more extensive permafrost area, there are more plot points in this plot.

the ability to reproduce DBP is limited by the fairly coarse vertical resolution of the model. Restricting observations to points where DBP is observed to be less than 250 m, it appears that U3.0M does not show any significant overall bias in DBP relative to observations. Although deep soil temperatures were found to be slightly warm overall relative to observations in Figure 3.16, there is no clear tendency towards thin simulated permafrost thicknesses. This however may be a function of the relatively limited number of sampling points. Despite the absence of any clear overall bias, there is a lot of scatter present in the comparison. As with the active layer thickness plot, a significant degree of this scatter presumably reflects spatial heterogeneity in surface conditions, but here there is also likely some influence from the model's uniform depth to bedrock and uniform bedrock characteristics which may poorly reflect actual subsurface conditions at specific sites. Overall, however, Figure 3.17 suggests that U3.0M does a reasonably job at capturing variations in permafrost thickness across the Canadian Arctic, especially given the paucity of detailed subsurface data available to inform the model.

In contrast, U3.0E clearly shows a tendency to overestimate DBP and permafrost thickness relative observations as FIgure 3.18 reveals. This is unsurprising, given the fact that the transition zone between thin and thick permafrost occurs too far to the south as well as the fact that soil temperatures generally appear to be too cold relative to observations (Figure 3.16) in this configuration.

## 3.3 Sensitivity Analysis

In this section, I present five sensitivity analyses to better understand how perturbing some of the characteristics of the land surface scheme impacts the model's simulation of regions of frozen ground. The U3.0E configuration of the model was used in this section to eliminate land-atmosphere feedbacks and simplify analysis.

### 3.3.1 Vegetation Distribution

To get a sense of how well the UVic ESCM reproduces the global distribution of vegetation, I plotted the six MOSES land cover types (5 PFTs + bare soil) against corresponding fields derived from the International Geosphere Biosphere Programme (IGBP) vegetation map (Loveland and Belward (1997)). The IGBP map is a 17-class land surface map that includes a number of vegetation categories as well as inland water, urban and ice classes. Vegetation, water, urban and ice fields from the IGBP product were mapped to MOSES land cover types by Dunderdale et al. (1999). As the UVic ESCM does not represent inland water, urban and ice fields within MOSES, gridcell fractions of these fields in Dunderdale et al. (1999)'s map were set to zero and the remaining six MOSES land cover types gridcell fractions were adjusted upwards, maintaining the relative proportion between types. The resulting IGBP-derived landcover fields are plotted in Figure 3.19 and are compared against the annual mean of the UVic ESCM vegetation fields that were used in conjunction with the ERA-40 data to drive the U3.0E version of the model. These UVic vegetation fields were derived from the vegetation distribution of UVic 3.0M in the final year of its 4,000 spinup under perpetual year-1800 forcing.

I produced a map showing an estimate of regional discrepancies between UVic and IGBP vegetation fields, following the methodology of Meissner et al. (2003). Within a model gridcell, the correlation between the fractional cover of the n$^{th}$ modelled PFT and corresponding IGBP PFT was calculated as follows:

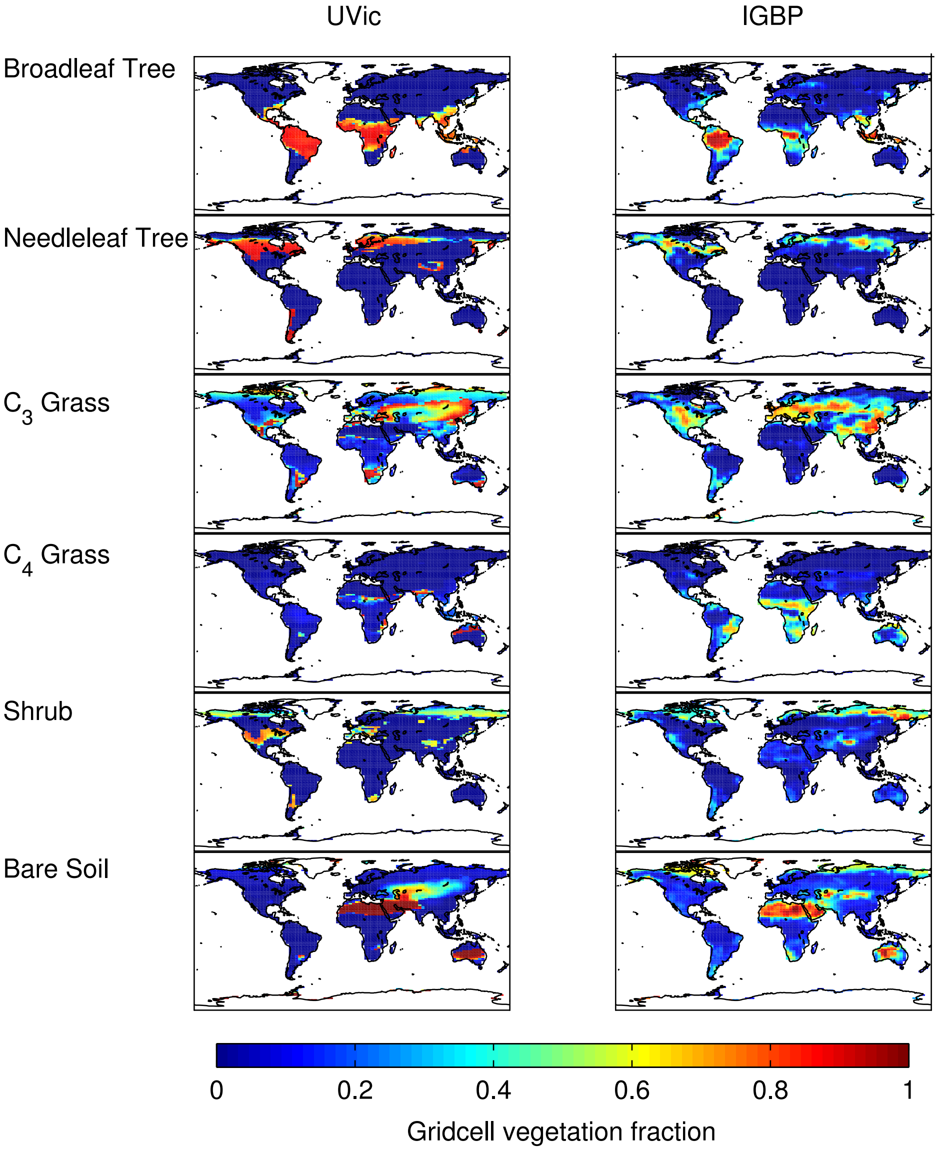


Figure 3.19: Distribution of land cover types in the UVic model and IGBP vegetation datasets. The UVic fields represent the annual mean vegetation distribution (derived from the year-1800 equilibrium state of the model of U3.0M) that was subsequently used to force U3.0E.

$$corr(PFT_n) = 1 - |fraction(PFT_{n,model}) - fraction(PFT_{n,IGBP})| \tag{3.6}$$

The total correlation within a gridcell is then defined as the minimum of the correlations for the 5 PFTs and an equivalently-calculated bare soil correlation. This total correlation is plotted in Figure 3.20. Much of the disagreement between the UVic and IGBP vegetation fields in temperate and low-latitudes may be attributed to the differing temporal coverage of the model and observations. The model allows for specified fractions of crop and pasture land within gridcells - as discussed in the following chapter, these fractions are derived from maps of estimates of historical crop and pasture lands supplied with the Representative Concentration Pathway datasets. Within crop and pasture lands, the model allows only $C_3$ and $C_4$ grasses to compete. Crop and pasture lands expanded substantially in the $19^{th}$ and $20^{th}$ Century and consequently, the ESCM land surface fields displayed in Figure 3.19 simulate more extensive regions of trees and shrubs relative to the IGBP dataset. A separate comparison (not shown) of the state of simulated vegetation of UVic 3.0M at year 2000 shows substantially better agreement in mid-to-low latitudes with an overall degree of correlation at a global scale comparable to previous versions of the model (see Figure 2 in Meissner et al. (2003)).The regions showing the most extensive change in crop and pasture land over the past 200 years generally do not correspond with simulated permafrost regions in this model configuration.

Another way of comparing surface cover fields is to compare the dominant surface cover[6] as in Figure 3.21. Within permafrost regions, U3.0E and IGBP surface covers generally agree reasonably well. The most noticeable discrepancy is that $C_3$ grasses tend to dominate at the highest latitudes where the IGBP dataset indicates should be largely bare soil. As well, needleleaf trees do not extend sufficiently far north into Alaska and are nearly absent in central Siberia. These anomalies reflect the difficulty of representing vegetation at a global scale using only five plant functional types in a climate model that is known to have significant regional temperature and moisture anomalies. I experimented with altering the TRIFFID in U3.0M vegetation parameters to obtain a better vegetation distribution in the Northern high latitudes, but found that an improved high latitude vegetation distribution came at the expense of a significantly worse vegetation distribution elsewhere in the world. Consequently,

---

[6] e.g. the surface cover having the greatest areal fraction within a model cell

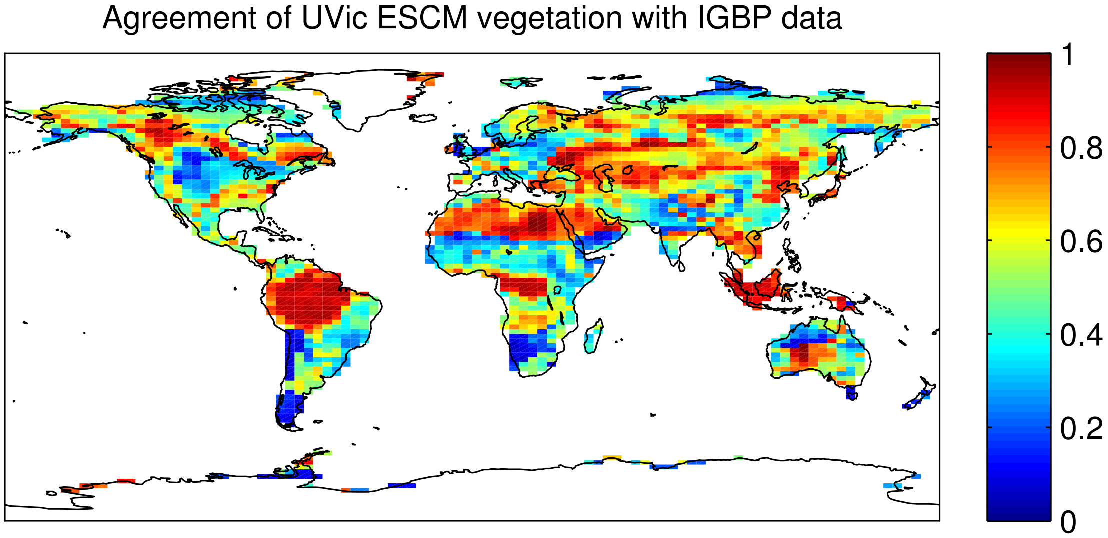


Figure 3.20: An estimate of the agreement between between UVic ESCM vegetation distribution and IGBP vegetation fields.The figure plots the correlation between the two land surface types in a grid cell with the correlation metric defined in the text.

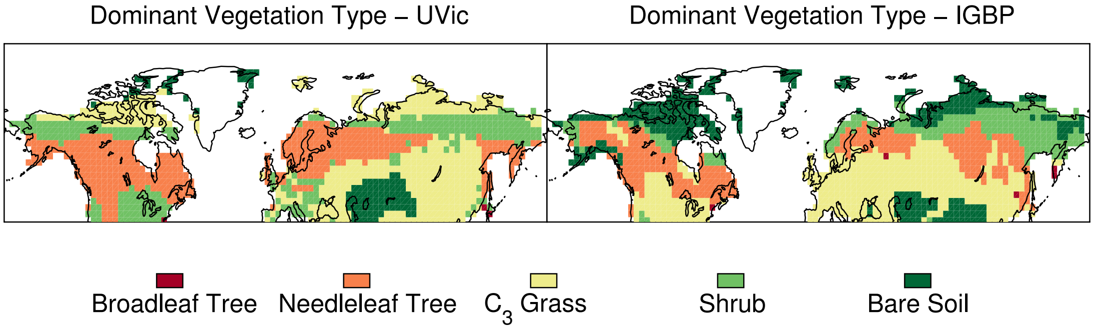


Figure 3.21: Dominant land surface types in the U3.0E vegetation fields and the IGBP dataset. $C_4$ grasses do not dominate anywhere in the mid-to-high latitudes of the Northern Hemisphere in either dataset and hence are not represented as a category.

while the vegetation distribution in the high latitudes is not perfect, it was achieved as the result of an attempt to have an overall reasonable global vegetation distribution.

As a means of understanding how significantly these vegetation anomalies influence the simulated permafrost distribution, I conducted a set of three sensitivity experiments in which the land surface distribution in the model was replaced with a single land surface type. The experiments shall be denoted *grass* (complete coverage with $C_3$ Grass), *tree* (complete coverage with Needleleaf Tree) and *bare soil* experiments (no vegetation whatsoever). In the case of the grass and tree experiments, both vegetation types had fixed leaf area indices and heights - these were determined as areal-weighted averages of the LAI and heights from the U3.0M vegetation distribution. A fourth experiment (*IGBP*) was conducted in which the vegetation and bare soil areal fractions were prescribed from the IGBP vegetation dataset with each PFT having a constant global LAI and height.

Table 3.3 compares the areal extents of permafrost and seasonally frozen ground in the different surface cover configurations. The extent of permafrost and seasonally frozen ground is comparable to the standard (ie. U3.0E) version of the model in the grass, shrub and IGBP configurations but the permafrost area is about 20 % lower in the bare soil configuration. Annual mean ground temperatures (not shown) are also comparable to the standard version in the grass, shrub and IGBP configurations with the exception of regions where bare soil is the dominant vegetation type in U3.0E. In these regions, replacing bare soil with vegetation results in the annual mean ground temperature decreasing by up 5 K.

| **Area (x $10^6$ km$^2$)** | **Standard** | **Bare Soil** | **Grass** | **Tree** | **IGBP** |
|---|---|---|---|---|---|
| Permafrost | 23.01 | 18.71 | 23.30 | 24.59 | 24.04 |
| Seasonally Frozen | 51.46 | 49.97 | 51.73 | 51.24 | 51.42 |
| Intermittently Frozen | 2.52 | 2.03 | 2.35 | 3.80 | 2.35 |

Table 3.3: Areas of permafrost and frozen ground in the different land surface configurations of the model. Standard denotes the model configuration U3.0E, described above.

As bare soil and vegetated surfaces show the greatest contrast, I focus on explaining these changes. Soil, as parameterized in the model, has a higher albedo than vegetated surfaces. Consequently, when vegetated surfaces are replaced with bare soil, there is a reduction in absorbed shortwave radiation, which would lead to cooling of the ground if other fluxes remained the same. The albedo difference between the bare soil and standard configuration is greater in the winter months (Figure 3.22, top panel). This is in part due to the fact that, where snow is present, bare soil is completely covered by snow (and the gridcell thus has the albedo of snow), whereas when vegetation is present the albedo lies between the snow-free and snow-covered-vegetation albedo both of which are lower than the albedo of snow. The albedo contrast between model configurations is highest where the standard configuration simulates boreal forest, owing to the sharp contrast in albedo between needleaf tree and bare soil.

The second major contrast between bare soil and vegetation pertains to the aerodynamic roughness, which governs sensible and latent heat fluxes. In this set of sensitivity experiments, each vegetation type has a constant height. Consequently, the aerodynamic resistance for a particular land surface class depends only on wind speed. For the same wind speed, the ratio of the aerodynamic resistance of bare soil to that of grass is 4.16, and the ratio of bare soil to needleleaf tree is 13.30. The impact of the greatly reduced aerodynamic roughness must be assessed in terms of the surface energy budget and the skin temperature. The skin temperature (T*) is the temperature that balances the surface energy budget. It is determined through a single iteration of the Newton-Raphson method (as discussed further in Appendix B) based on the surface energy budget evaluated at the atmospheric temperature ($\mathrm{b}(\mathrm{T}_a)$) and the derivative of the energy budget evaluated at this temperature:

$$T^{*} = T_a - \frac{b(T_a)}{\frac{db}{dT}} \tag{3.7}$$

The derivative of the surface energy budget is a negative-definite quantity which depends on the aerodynamic roughness ($\mathrm{r}_a$):

$$\frac{db}{dT} = -4\sigma T^3 - \frac{\rho_a c_p}{r_a} - \frac{\rho_a (L_c)}{r_a + r_c}\frac{dq_{sat}}{dT} - \frac{\lambda_{soil}}{\frac{\Delta z}{2}} \tag{3.8}$$

If the aerodynamic resistance increases (e.g. going from vegetation to bare soil), then $\frac{db}{dT}$ becomes less negative and it follows from Equation 3.7 that the skin temperature will be further from the air temperature. Thus, in going from the standard version to the bare soil configuration of the model, the skin temperature will become cooler during times of the year when $\mathrm{T}^* < \mathrm{T}_a$ and warmer during times of the year when $\mathrm{T}^* > \mathrm{T}_a$. Figure 3.23 shows the difference between the skin temperature and the air temperature in the standard version of the model in January and July. In

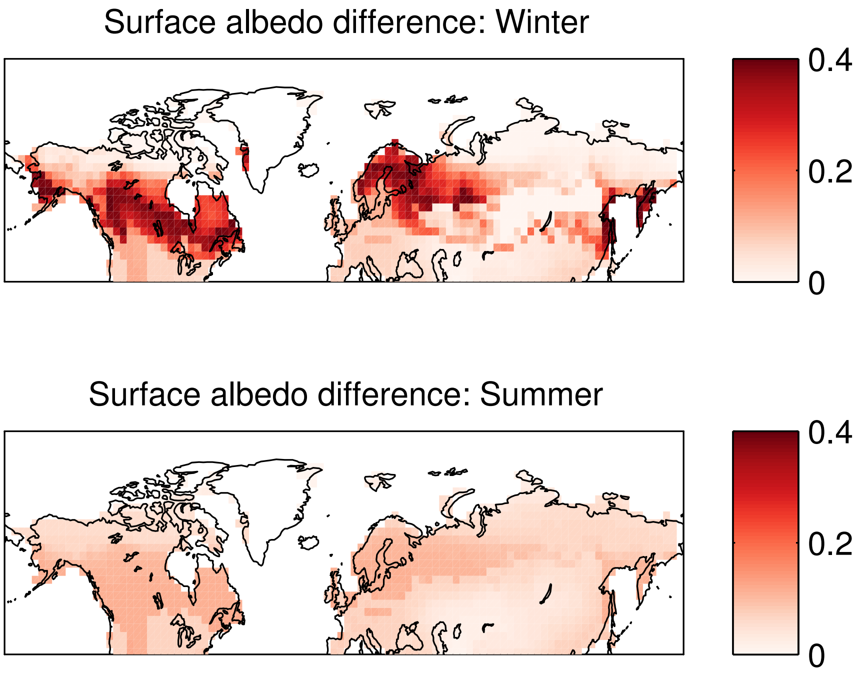


Figure 3.22: Difference in land surface albedo between bare soil configuration and standard configuration. Top panel: January albedo; bottom panel: July albedo

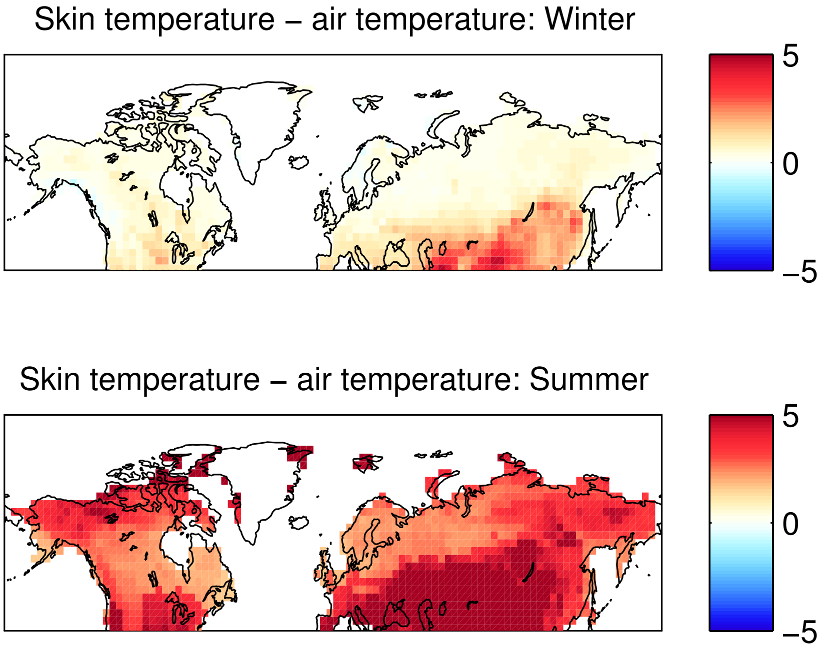


Figure 3.23: Difference between skin temperature and air temperature in standard version (U3.0E). Top panel: January difference, bottom panel: July difference.

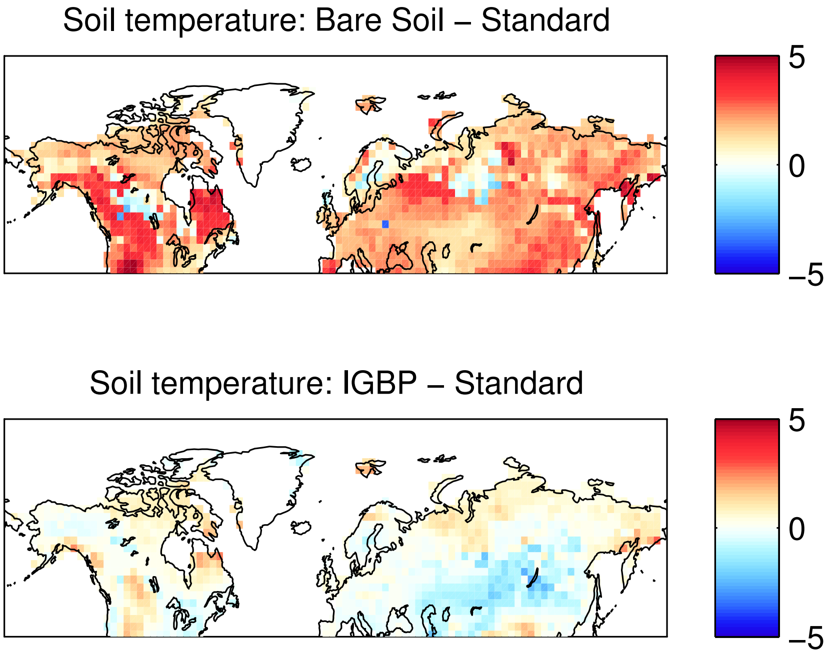


Figure 3.24: Difference in annual mean near surface (top 1m) soil temperature. Top panel: Bare soil - standard; bottom panel: IGBP - standard.

the winter, the skin temperature is fairly close to the air temperature in the Northern mid-to-high latitudes, generally equal to the air temperature or slightly warmer. In July (and much of the spring, summer and fall), the skin temperature is much warmer than the air temperature everywhere. Consequently, the net annual effect of increasing the aerodynamic roughness is an increase in the skin temperature leading to a corresponding increase in the soil temperature.

There are therefore two competing effects on soil temperature when the bare soil configuration is compared with the standard configuration. As the top panel of Figure 3.24 reveals, the albedo-cooling effect tends to counteract the roughness-cooling impact in the boreal forest region. However, warming is sufficiently pronounced in the southern limits of permafrost to substantially reduce the overall permafrost area simulated relative to the standard configuration of the model.

In the case of the *tree* simulation, the albedo is typically lower than the standard configuration, particularly in regions of extensive snow cover. Nonetheless, annual mean ground temperatures are lower throughout the permafrost region by around 0.5 °C associated with a strong increase in sensible heat flux leading to enhanced cooling of the ground. This indicates that here as well, the change in aerodynamic roughness tends to have the dominant effect. The slight cooling of the ground in the all-tree run results in a slightly more extensive permafrost area than in the standard model. The overall permafrost area in the all-grass configuration is nearly identical to the standard version. Soil temperatures are ∼ 2-3 °C higher (and the southern limit of permafrost shifts slightly northward) where boreal forest is replaced with grass; soil temperatures are slightly cooler (and the southern limit of permafrost shifts slightly southward) where bare soil dominated terrain in central Siberia is replaced with grass.

It should be emphasized that the model configurations in which the model land surface contains only a single land surface type are meant to be extremes, designed to assess the model sensitivity to land surface cover. The model is clearly most sensitive (in terms of the distribution of permafrost) to the presence or absence of bare soil, as a consequence of the relatively strong albedo differences between soil and vegetation as well as the strongly reduced aerodynamic roughness over bare soil. Given this sensitivity, it may be worth investigating in future work whether the single values for soil albedo and roughness length are reasonable values for soil characteristics at a global scale.

As a means of assessing the impact of vegetation distribution discrepancies, I compared the state of the standard version of the model to a version of the model where the distribution of vegetation was specified from the IGBP dataset. Soil temperatures (bottom panel, Figure 3.24) are slightly warmer in the IGBP configuration in regions where the UVic ESCM simulates an overabundance of trees (Central North America, the West Coast of British Columbia and Alaska, Northern Quebec, the Eastern Coast of Siberia), and where the model simulates too little bare soil (most sites bordering the Arctic Ocean). Soil temperatures are cooler in the IGBP configuration where the model simulates and overabundance of grasses and an under abundance of trees (e.g. Siberia in the vicinity of Lake Baikal, Canada in the vicinity of Great Slave Lake). These temperature anomalies are generally no larger than about 2 °C and result in a slightly more extensive permafrost area in this model configuration.

It should be noted that the representation of vegetation in the UVic ESCM is simplistic as the surface is simply treated as a bulk surface/vegetation layer with the skin temperature representing the temperature of the combined layer. More sophisticated land surface schemes include a separate ground layer and (one or more) vegetation layer(s) with radiative coupling between the layers. The ESCM consequently cannot represent radiative interactions between the vegetation and the ground surface and effects such as vegetation's ability to shield the surface from sunlight. There is also no distinction between turbulent fluxes of sensible and latent heat at the canopy and ground surface nor of the effect of vegetation on reducing wind speed from the canopy towards the ground nor of the tendency for dense vegetation to act as ‘snow fences’ to retain snow. Consequently, while the above sensitivity analysis accurately accounts for the impact of vegetation on the characteristics of the ground within the ESCM, it likely fails to realistically capture the full sophistication of vegetation-ground interactions in the actual climate system.

### 3.3.2 Snow Thermal Conductivity

The second sensitivity experiment explores the impact of snow thermal insulation properties on the state of the subsurface through varying the fixed snow thermal conductivity. I chose this experiment as the strong difference in permafrost area and soil temperature between U3.0M and U3.0E suggested that the land surface scheme was strongly sensitive to snow processes. As well, a relatively recent paper (Cook et al. (2007)) using the NCAR CAM3-CLM3 model demonstrated that varying snow thermal conductivity through a range of realistic values could result in strong cooling or warming of the soil, with concomitant impacts on soil moisture storage and rates of photosynthesis through changes in frozen soil moisture concentrations.

Snow in the UVic ESCM has a constant density; consequently, the model’s snow thermal conductivity is a fixed value, unlike many other land surface schemes that allow for snow compaction and density-dependent snow thermal conductivity (see, for example, Slater et al. (2001)). In both U3.0M and U3.0E (the “standard” model in this section), this thermal conductivity is set at 0.25 W $m^{-1}$ $K^{-1}$. In reality, snow thermal conductivity generally increases from fall to spring as snow compacts and air-filled pore spaces are reduced. I tested two additional configurations of the model with low and high snow thermal conductivity values. The low conductivity / high insulation configuration has a thermal conductivity of 0.10 W $m^{-1}$ $K^{-1}$, which is typical of that of fresh snow, while the high conductivity / low insulation configuration has a conductivity of 0.50 W $m^{-1}$ $K^{-1}$, representative of that compacted or partially melted snow. These values were chosen to be compatible with those in the study of Cook et al. (2007).

I first compared soil temperatures near the top of the soil column (using an average soil temperature for the top 1 m of soil) and the bottom of the soil column (using the temperature of the bottommost layer, which extends from a depth of 5.8

m to 10.0 m) between the three model configurations, as shown in Figure 3.25. As anticipated, in the high insulation run, near surface soils show increased soil temperatures of up to $\sim$ 10 K in winter throughout the snow covered regions of the Northern Hemisphere (top left panel, Figure 3.25), with soil warming generally most pronounced where snow is thickest (refer to Figure 3.7). Soils generally remain warm through the summer (top right panel, Figure 3.25), though there are a few regions where the soil is cooler in summer in the high insulation configuration. These cells are associated heightened summer evapotranspiration rates in the higher insulation run; warmer winter temperatures result in increased liquid moisture availability in the spring / summer which in turn results in heightened rates of evaporative cooling. The opposite tendency is shown in the low insulation run: near surface soils cool substantially in the winter months (second row, left panel, Figure 3.25) with cooler soils persisting into summer (second row, right panel, Figure 3.25). Deeper soils layers (bottom two rows, Figure 3.25) show the same warming and cooling tendencies in the high and low insulation configurations, respectively, although there is no appreciable annual cycle in the deeper layers.

The differences in the annual cycle in near surface and deep soil temperatures in the different thermal conductivity configurations is clearly seen in the central panels of Figure 3.26. Here, soil temperatures were averaged over all land grid cells above 45 ° N to roughly encompass regions that experience significant snowfall rates. The plots confirm that near surface soil temperature differences are greatest near the end of winter and are fairly minor in the summer, whereas deep soil temperature differences remain at about the same level, year round. Interestingly, although the difference between the standard and high insulation conductivities (0.15 W $m^{-1}$ $K^{-1}$) is greater than the standard and low insulation conductivities (0.25 W $m^{-1}$ $K^{-1}$), the temperature differences between standard and high insulation simulations is greater. Presumably, this is because there is a limit to the extent that snow cover can prevent soils from equilibrating with skin temperature. One can extrapolate from these results and imagine that if the snow thermal conductivity were incrementally increased that the annual soil temperature curve would then asymptotically approach the skin temperature annual cycle.

The soil temperature results described above are qualitatively similar to those presented by Cook et al. (2007) (see their Figure 7, top right panel, for instance) and their explanation for the changes in soil temperature would appear to apply to my results as well - the explanation relates to changes in the ground heat flux - and an

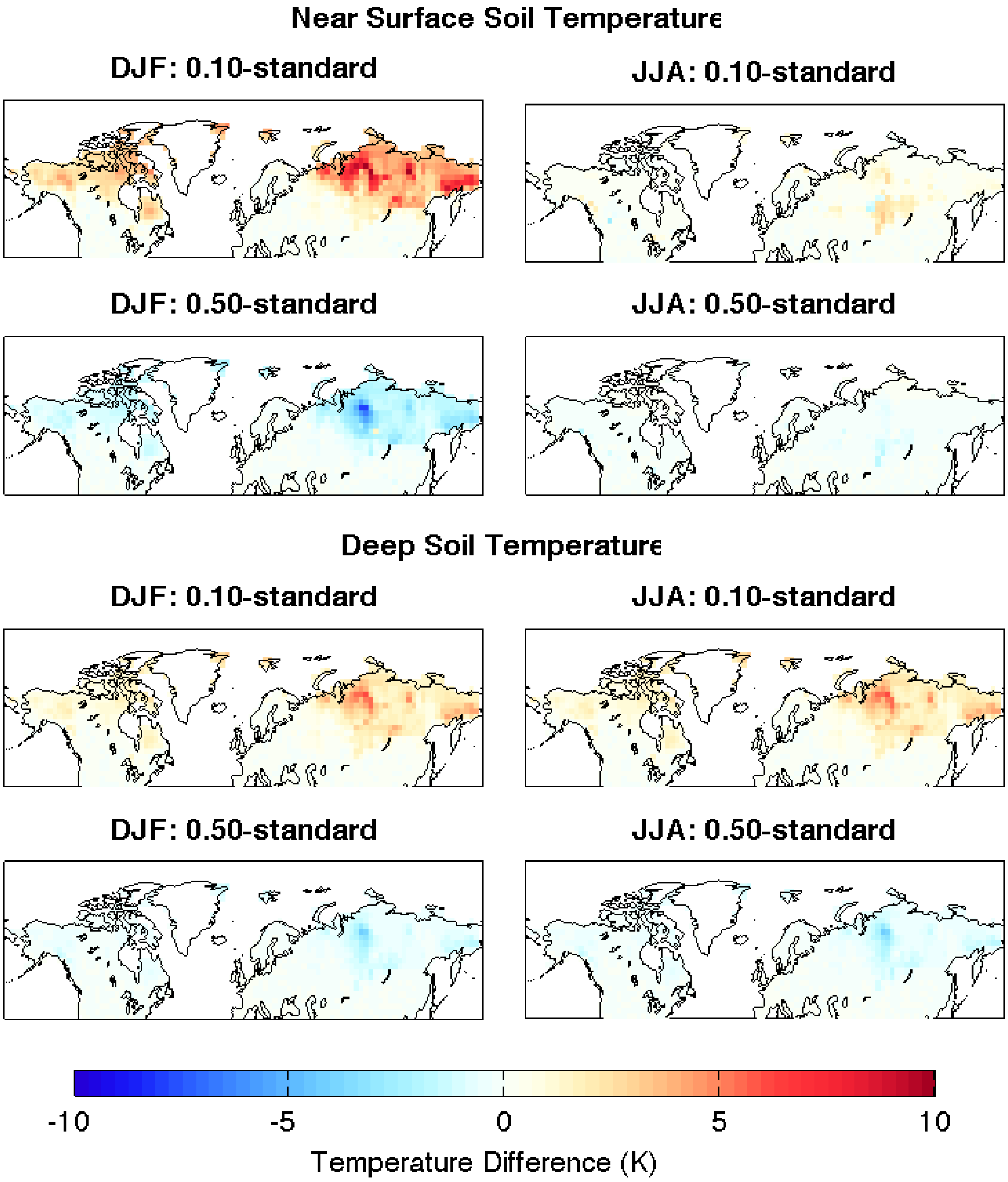


Figure 3.25: Near surface and deep soil temperature differences in winter (DJF) and summer (JJA) associated with changes in snow thermal conductivity. Low conductivity / high insulation (0.10 W $m^{-1}$ $K^{-1}$) and high conductivity / low insulation (0.50 W $m^{-1}$ $K^{-1}$) configurations are compared against the standard configuration (0.25 W $m^{-1}$ $K^{-1}$) )

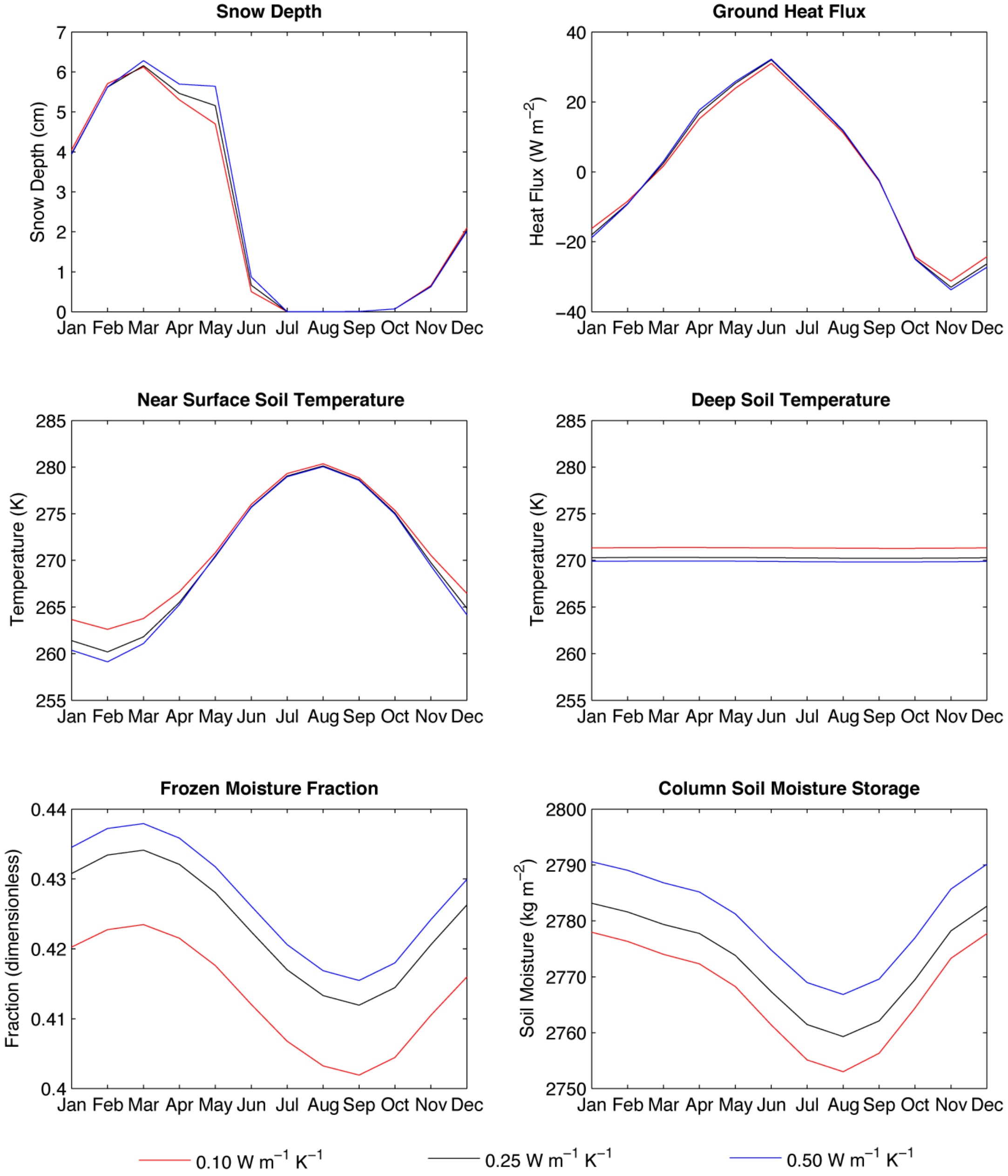


Figure 3.26: Impacts of soil thermal conductivity changes on land surface state. All variables presented here are averages tabulated over all land cells above 45° N.

average (> 45° N) of this quantity is shown in the top right panel of Figure 3.26).

An increased snow thermal conductivity results in increased transfer of heat from the ground to the atmosphere during winter months (negative ground heat flux), resulting in colder soil temperatures. During the summer months, the colder soil temperatures, associated with the heightened snow thermal conductivity, result in a stronger temperature gradient between the soil and the atmosphere and consequently a greater transfer of heat from the atmosphere to the ground (positive ground heat flux). The same logic applies in reverse to a decrease in snow thermal conductivity. While the heat flux differences between the various model configurations appear to be relatively small at any given time, over the course of a season, they can lead to fairly strong cooling and warming of the soil with an overall annual cooling of the soil in the configuration with high conductivity / low insulation and warming of the soil in the low conductivity / high insulation configuration. The changes in the soil temperature also influence the snow depth as shown in the average (> 45 ° N) snow depth plot in the top left panel of Figure 3.26). While the snow depths are virtually identical across the three model configurations during the fall and winter as snow accumulates, the depths begin to diverge as the snow melts. The snow depth decreases most rapidly in the low conductivity configuration as a result of the warmer soil temperatures, and decreases least rapidly in the high conductivity configuration. This must then play a role in retarding (enhancing) the spring/summer warming of the soil in the high (low) conductivity configuration, reinforcing the soil temperature trends induced by the original conductivity change.

Cook et al. (2007) found that the soil temperature differences in the different conductivity configurations could result in substantial differences in the soil ice fractions in their soil column (see the bottom right panel in their Figure 7), with differences in the fraction of soil moisture that is ice (soil ice fraction) of up to 20 % in winter. These results are substantially greater than shown in the UVic ESCM (bottom left panel, Figure 3.26) and I attribute these differences to the fact that the soil column in the CAM3-CLM3 model used in their study has a depth of 3.43 m vs. the 10 m soil depth (and 250 m total column depth) in the UVic ESCM. This results in a greater annual soil temperature cycle in CAM3-CLM3 relative to the UVic ESCM which leads to greater variations in frozen / unfrozen soil moisture content at all depths. As well, the soil ice fractions I present are averaged over the entire soil column (to a depth of 10 m) and deeper soil layers show substantially reduced annual variations in frozen soil moisture content throughout the year. Despite the fact that the two models show substantially different magnitudes of changes in ice fraction with snow thermal conductivity, the sign of the changes is compatible. In the high conductivity configuration, cooler soils result in a greater soil ice fraction (particularly in the winter months), with the opposite true of the low conductivity configuration. An increase in soil ice fraction is also associated with increases in total column soil moisture storage (including both liquid and frozen moisture - bottom right panel, Figure 3.26). This is a result of decreased hydraulic conductivity (and less runoff) and decreased evapotranspiration associated with increased ground ice content. The changes in average soil moisture storage in the various ESCM configurations are relatively minor compared to the total column soil moisture content, however.

Finally, I investigated how the different thermal conductivity configurations impacted the simulated distribution of permafrost and seasonally frozen ground (Table 3.4). As one might expect, permafrost area was greatest in the high conductivity (cooler soils) configuration and lowest in the low conductivity (warmer soils) configuration, although the difference in permafrost area between these two extremes is only 1.7 x $10^6$ km$^2$ (less than 10 %). Although the near surface soils showed substantial seasonal temperature differences between the various thermal conductivity configurations, the deep soil layers' temperature changes were substantially lower and it is these deep soil layers that remain perpetually frozen: deep soil layers warmed by, on average, ~1 K in the low conductivity configuration (relative to the standard configuration) and cooled by, on average, ~0.5 K in the high conductivity configuration. The maximum extent of seasonally frozen ground was also identical between the model configurations although regions of intermittently frozen ground were slightly lower in the low conductivity configuration relative to the other configurations.

| **Area** | **Low cond.** **(0.10 W m$^{-1}$ K$^{-1}$)** | **Standard cond.** **(0.250 W m$^{-1}$ K$^{-1}$)** | **High cond.** **(0.50 W m$^{-1}$ K$^{-1}$)** |
|---|---|---|---|
| Permafrost | 21.91 x $10^6$ km$^2$ | 23.01 x $10^6$ km$^2$ | 23.61 x $10^6$ km$^2$ |
| Seasonally Frozen | 51.46 x $10^6$ km$^2$ | 51.46 x $10^6$ km$^2$ | 51.46 x $10^6$ km$^2$ |
| Intermit. Frozen | 2.47 x $10^6$ km$^2$ | 2.52 x $10^6$ km$^2$ | 2.52 x $10^6$ km$^2$ |

Table 3.4: Areas of permafrost and frozen ground with different snow thermal conductivity values. Standard denotes the model configuration U3.0E, described above.

Active layer thicknesses were virtually identical in the three thermal conductivity configurations (not shown). While the presence of permafrost reflects the condition of deep soil layers, the active layer thickness reflects the state of layers closer to the surface. These surface layers showed substantial differences in winter soil temperatures, however the summer soil temperature differences in near surface soil layers was relatively minor. As the summer temperatures determine the extent to which the soil column thaws, it is not surprising that the active layer thickness remains nearly identical in the different configurations.

### 3.3.3 Resolution

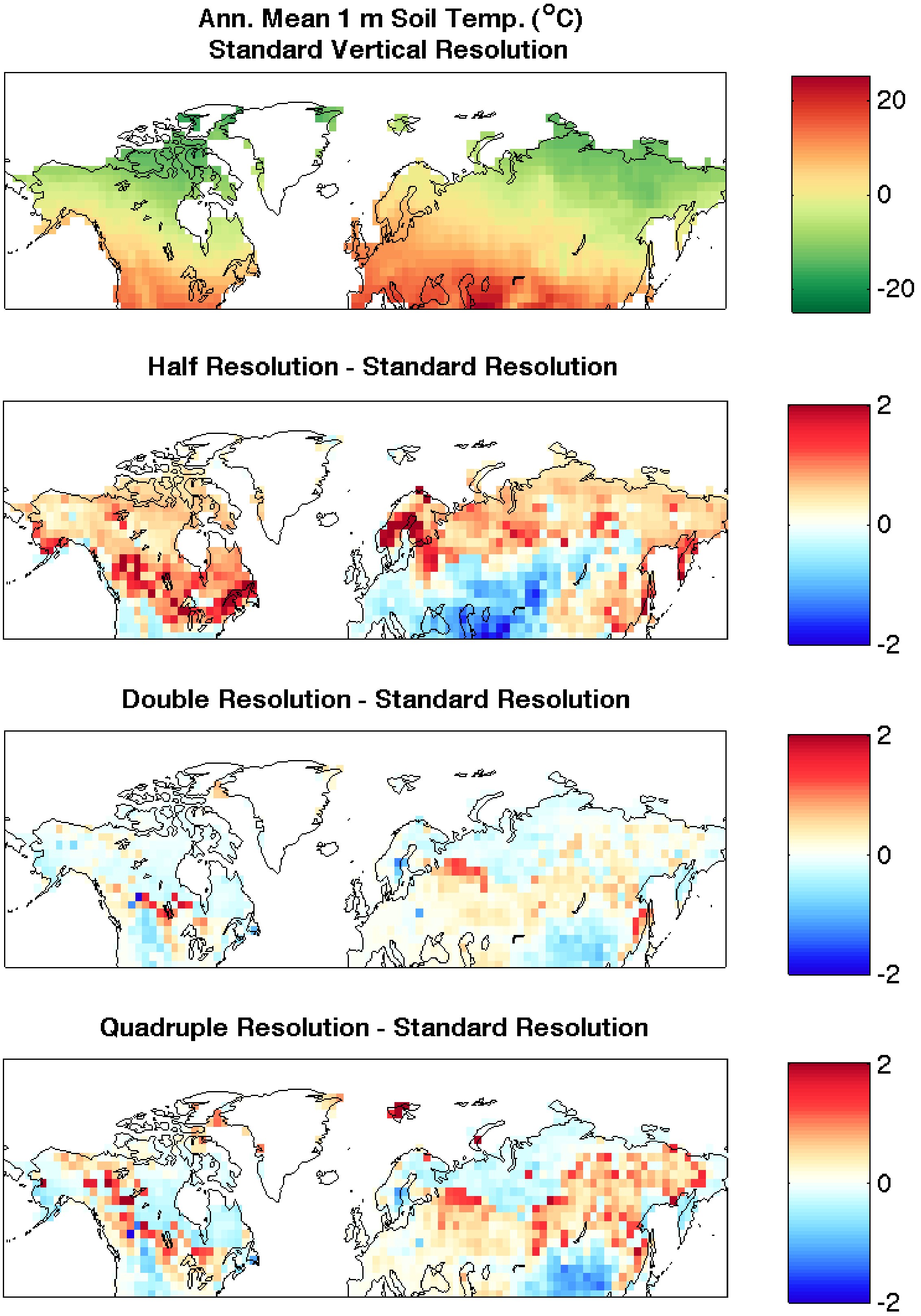
The subsurface vertical resolution of the ESCM's land surface scheme was chosen somewhat arbitrarily early in the project. When I first added multiple soil layers to the model, I experimented with a variable number of soil layers, though in each case, I allowed the soil layer thickness in increase exponentially with depth to a maximum depth of 250 m. I found that the simulated permafrost extent area and equilibrium soil temperatures generally did not change substantially once the number of subsurface layers increased beyond 12-14. As the model's computational requirements increased with the number of subsurface layers, I did not see any clear merit early on in substantially increasing the number of subsurface layers beyond 14.

In this section, I investigate the impact of altering the LSS's vertical resolution. For this sensitivity experiment, I developed three additional configurations of the model with altered vertical resolutions. A *double resolution* version was developed by splitting each soil layer in the standard version into two separate soil layers with half the original thickness; similarly, a *quadruple resolution* version was developed by dividing each original soil layer into four new layers with equal thickness. A *half resolution* version was produced by combining layers in pairs from the standard version (e.g. layers 1 and 2 in the standard version became layer 1 in the half resolution version).

To investigate differences in the subsurface state of the four resolution configurations, I plotted differences in the annual mean near surface (top 1 m) soil temperature, as shown in Figure 3.27. The differences in soil temperature tend to be greatest between the half resolution configuration and the standard configuration and, in most grid cells, temperature differences decrease as the resolution increases[7]. There is no

---

[7] ie. the temperature difference in a grid cell in between the quadruple resolution and double resolution configurations is less than the difference between the double resolution and standard resolution which is less than the difference between the standard resolution and the half resolution configurations.


Ann. Mean 1 m Soil Temp. (°C)
Standard Vertical Resolution

Half Resolution - Standard Resolution

Double Resolution - Standard Resolution

Quadruple Resolution - Standard Resolution

Figure 3.27: Differences in annual mean soil temperature in four models with varying vertical resolution. The top panel shows the annual mean soil temperature in the standard configuration. The bottom panels show soil temperature differences in different configurations of the model relative to the standard configuration.

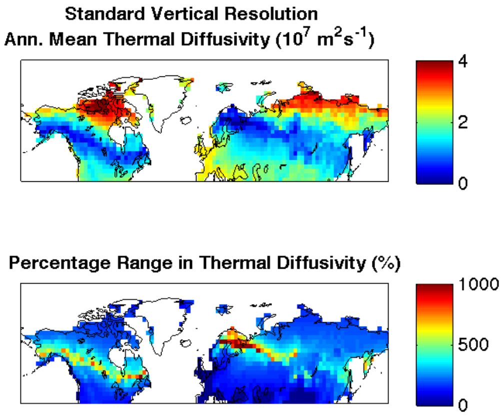
immediately clear overall pattern in terms of regions which tend to warm vs. cool as the resolution is altered, although the warming and cooling changes are largely consistent as the resolution is increased (ie. regions which are warmer in the half resolution configuration relative to the standard resolution configuration are typically cooler in the double resolution configuration relative to the standard resolution configuration).

Further investigation showed that the warming and cooling trends could be linked to changes in ground thermal diffusivity, which governs the rate of heat transfer into / out of the soil. The ground thermal diffusivity is a characteristic that is more fully explored in Appendix B, but it essentially accounts for the effects of both changes in thermal conductivity and heat capacity on heat transfer through the soil. Here, I calculate an 'effective' thermal diffusivity ($k_{eff}$, $m^2/s$) for the top meter of the soil. This is defined as the effective thermal conductivity for heat transport across the first four layers of soil (that span the top 1 m) divided by the average heat capacity of these layers. The effective thermal conductivity term (numerator in equation 3.9) is calculated by treating the layers as though they were resistors in series, and summing the resistances (inverses of the thermal conductivities) to find an overall thermal resistance across the top four layers. The effective heat capacity for the top meter of soil is simply the depth-weighted average heat capacity of the layers.

$$k_{eff} = \frac{\frac{\sum_{n=1}^{4} \Delta z_n}{\sum_{n=1}^{4} \frac{\Delta z_n}{k_n}}}{\frac{\sum_{n=1}^{4} C_n \Delta z_n}{\sum_{n=1}^{4} \Delta z_n}} \tag{3.9}$$

Figure 3.28 compares the effective thermal diffusivity of the top 1 m of soil in standard and quadruple resolution configurations. The annual mean and range in effective thermal diffusivity is generally quite similar between the two versions, though there are regional differences which are heightened in particular seasons. To illustrate these differences more clearly, I plotted seasonal differences in the thermal diffusivity between the quadruple resolution and standard resolution configurations (Figure 3.29).

Regions where there is warming of the soil in the quadruple resolution configuration are associated with either a reduction in the effective thermal diffusivity in the winter (ground cools less readily in winter) or a increase in the thermal diffusivity in summer (ground warms more readily in summer). The reverse is true for grid cells that tend to cool in the quadruple resolution configuration of the model. The effective thermal diffusivity of the top 1 m of soil is a complicated function that depends on



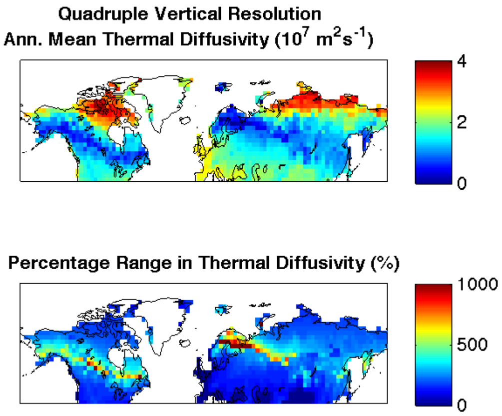


Figure 3.28: Mean and annual range in ground thermal diffusivity. Left column: model with standard vertical resolution; right column: model with quadruple vertical resolutiuon.

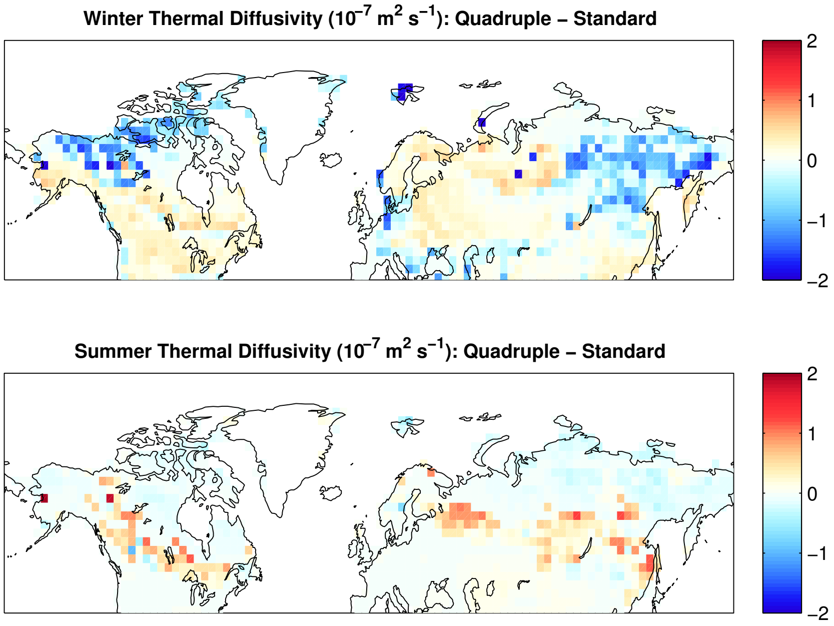


Figure 3.29: Difference in winter and summer effective thermal diffusivity between quadruple and standard resolution configurations of the model.

the distribution of soil moisture throughout the top metre of the soil as well as on the form of soil moisture (which in turn depends upon the temperature distribution). It is a highly non-linear function of soil moisture content owing to the fact that the effective thermal conductivity of the top 1 m of soil is a non-linear function of the thermal conductivity of individual soil layers and the thermal conductivity within a particular soil layer is a non-linear function of layer soil moisture content. Where there is a vertical gradient in soil moisture content or form in the high resolution model that is masked in the standard resolution model, the effective thermal diffusivity in the two different configurations will generally not be the same.

The point that I wish to make is that explanations for particular seasonal changes in the thermal diffusivity of model configurations having different vertical resolution are potentially extremely complicated and I could not find an obvious, straightforward explanation as to why the thermal diffusivity of a certain grid cell might seasonally change as a consequence of a change in model resolution. However, the resolution-dependent changes in thermal diffusivity can explain the observed soil temperature differences between the different model configuration.

I next compared the simulated state of permafrost in the four versions of the model against observations to assess whether increasing the vertical resolution altered the representation of permafrost in the model. The Northern Hemisphere areal extent of frozen ground regions in the four model configurations is tabulated in Table 3.5. The table indicates that the simulated extent of permafrost and seasonally frozen ground is generally comparable across the four versions with the exception of the half resolution version, where the simulated area of seasonally frozen ground is clearly substantially lower than in the other model versions. The thicker top soil layer in the half resolution version should take longer to freeze than the thinner top layers in the higher resolution configurations. Consequently, some grid cells in the southern most region of seasonally frozen ground which ought to freeze during the winter may not be simulated as freezing in the half resolution version, explaining the reduced seasonally frozen ground area.

The simulated active layer thickness was also examined (Figure 3.30). As discussed earlier, the simulated ALTs represent the bottom of the deepest soil layer that *fully thaws* in the model such that if the true ALT in a finite resolution model occurs at a depth between soil layers, the ALT that is identified is the depth of the layer above the true ALT. Consequently, in any configuration of the model, the ALT can have only a fixed number of possible values set by the thickness of the soil layers. If

| **Area (x $10^6$ km$^2$)** | **Half Res.** | **Stand. Res.** | **Double Res.** | **Quad. Res.** |
|---|---|---|---|---|
| Permafrost | 21.88 | 23.01 | 22.55 | 22.26 |
| Seasonally Frozen | 46.73 | 51.46 | 53.404 | 54.62 |
| Intermittently Frozen | 2.82 | 2.52 | 2.25 | 1.78 |

Table 3.5: Northern Hemisphere areas of permafrost and frozen ground with different model vertical resolutions.

there were no soil temperature changes associated with a change in resolution, then the simulated active layer thickness should either remain the same or increase as the model resolution is increased. As there are soil temperature changes accompanying changes in vertical resolution, these manifest themselves as a deepening or shoaling of the active layer superimposed upon the general trend towards deeper active layer thicknesses. The results shown in Figure 3.30 are consistent with these observations.

When I compared the simulated active layer thicknesses of the standard resolution configuration against observations, I observed that ALTs were generally shallower than observations and attributed this to the fact that the ERA-40 offline configuration of the model showed a cold soil bias as well as to the fact that the model would tend to underestimate active layer thickness as a consequence of its discrete vertical resolution. Figure 3.31 compares the active layer thicknesses in the four resolution configurations against observations. The agreement between the observations and the models clearly improves as the resolution increases, illustrating that a substantial amount of the apparent cold bias that might be inferred from the shallow simulated active layer thicknesses is simply an artefact of the model resolution. At the same time, even in the high resolution configuration of the model, the model still appears to show a shallow bias in the simulated ALTs. This is probably to be expected as a consequence of the fact that soil temperatures appear to be too cool in the offline configuration of the model, but it would be interesting to see if this apparent shallow bias were to decrease if the model's resolution were increased further.

I also compared near surface and deep soil temperatures against borehole temperature measurements in the four model configurations using the same methodology described earlier in this work (Figure 3.32). While the different model configurations showed patterns of warming and cooling of up to $\sim 2$ K between resolution configurations, the borehole scatter plots are generally quite similar between the four different configurations. All four resolution configurations also maintain the general U3.0E cold soil temperature bias both near the surface and at depth relative to the

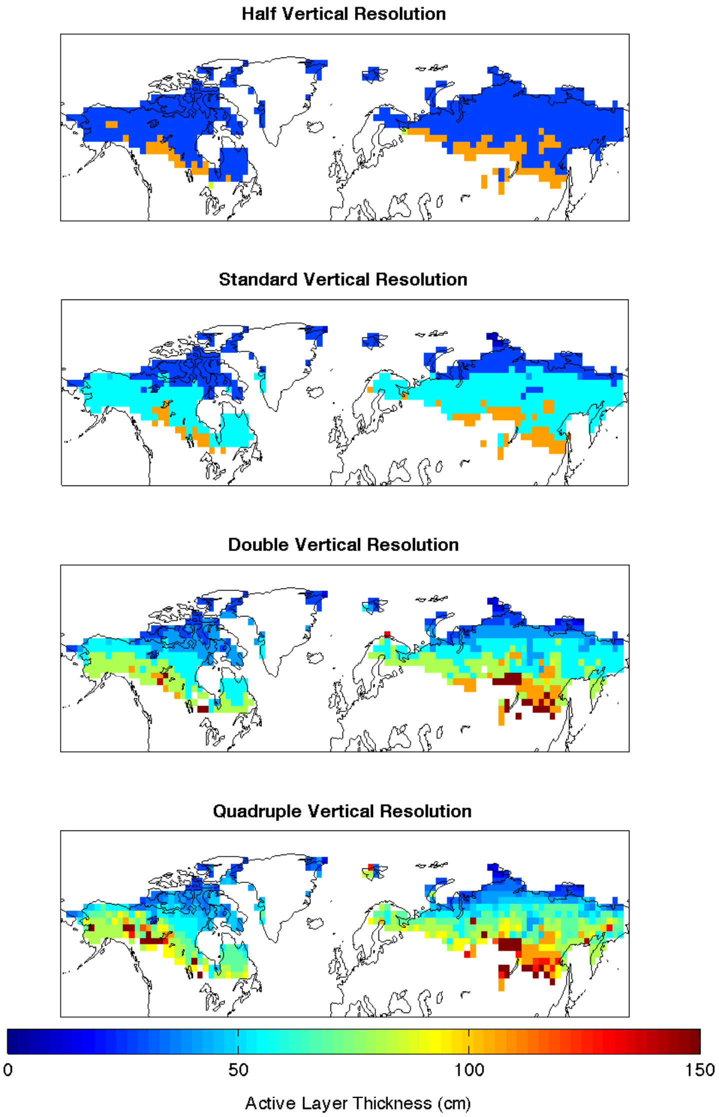


Figure 3.30: Active layer thickness in four model configurations having different vertical resolution.

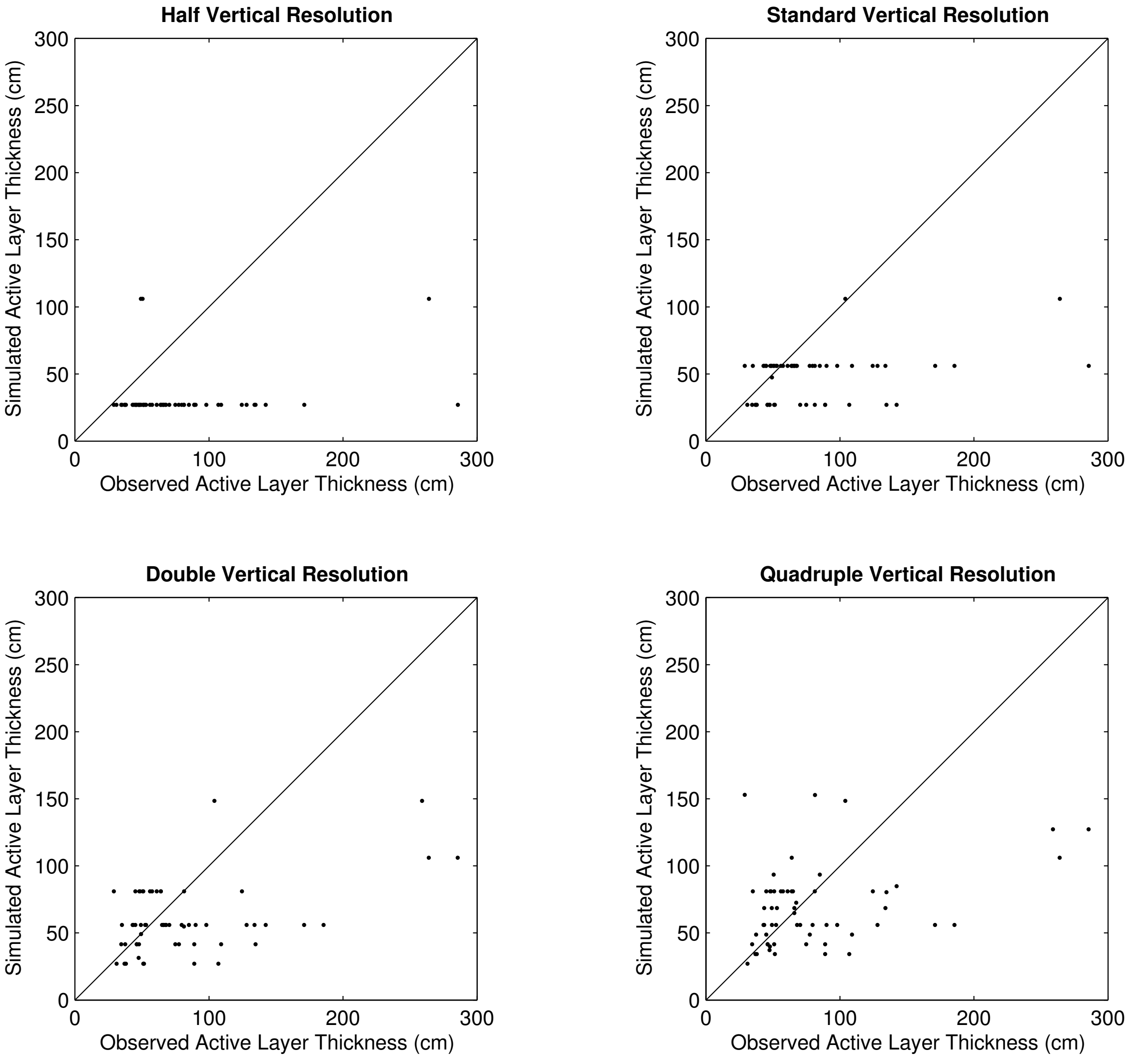


Figure 3.31: Scatter plots comparing simulated active layer thicknesses in four different model configurations with differing vertical resolution with observations.

observations.

While there are regional temperature differences between the four different resolution configurations of the model linked to changes in ground thermal diffusivity, the overall extent of permafrost and seasonally frozen ground tends to be fairly similar between the four models. The most noticeable impact of increasing the vertical resolution relates to the improved representation of the active layer thickness in the higher resolution configurations. If the model is to be used in an application where accurate representation of the active layer thickness is very important, then a higher resolution configuration may be preferred. This could potentially be important, for example, where the model's soil carbon stores are concerned. As bacterial decomposition depends on the amount of unfrozen moisture present in the soil in addition to temperature, it may be important to better represent the seasonal thawing of the ground for this reason. Such improvements must be traded against the considerably higher computational cost of the high resolution model configuration. As a consequence of these heightened computational demands and the fact that the overall permafrost extent tended to be fairly similar between the standard and quadruple resolution version of the model, I chose to pursue the remainder of my simulations with the standard configuration.

### 3.3.4 Depth to Bedrock

As discussed in Chapter 2, the model's uniform depth-to-bedrock (or, thickness of the soil layers) of 10.00 m was chosen as a consequence of the lack of information on this quantity at a global scale. To test the sensitivity of the LSS to this parameter, I produced four additional soil thickness configurations of the model. In all cases, the total depth (250 m) and number (14) of ground layers was kept constant but the number of subsurface layers that were soil layers was varied. The four additional configurations of the model had soil thicknesses of 0.27 m (two soil layers), 1.06 m (four soil layers), 5.81 m (seven soil layers) and 29.32 m (ten soil layers). While the depth to bedrock may exceed 30 m in some locations, results with these initial four soil depth configurations suggested that deeper depth-to-bedrock configurations would not be substantially different than the configuration with ten soil layers, as shall be shown.

I first considered the difference in the annual mean, near surface (top 1 m) ground temperature in the different soil thickness configurations relative to the standard

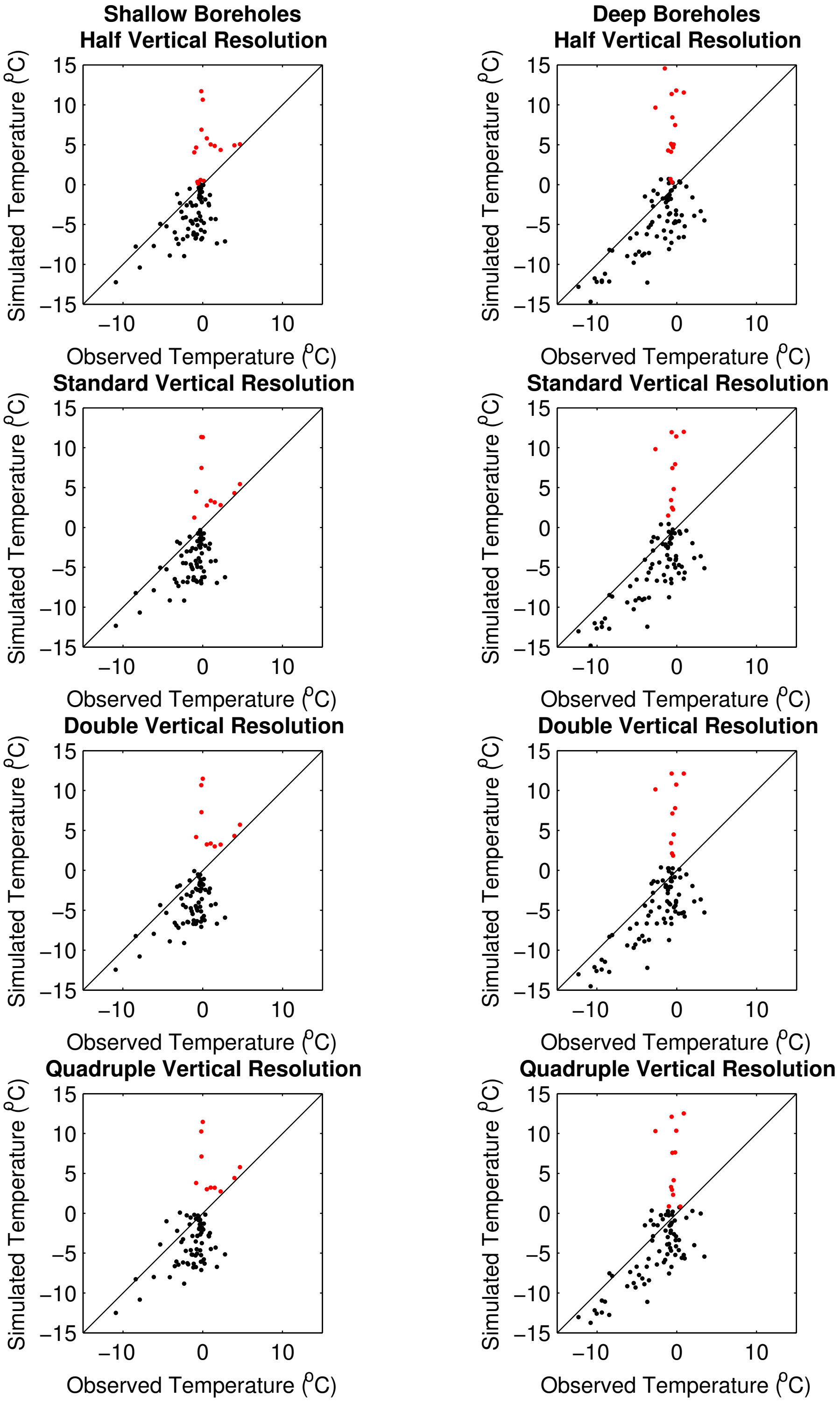


Figure 3.32: Scatter plots comparing near surface and deep borehole temperature measurements with simulated temperatures in four model configurations with varying vertical resolution.

versions of the model (8 soil layers, 10.00 m soil thickness) as shown in Figure 3.33 (left column). Also plotted in this figure is the difference in the annual range in near surface ground temperature (right column). The figure reveals that the thin 0.27 m and 1.06 m soil thickness configurations shown strong differences in both quantities relative to the standard model configuration while differences are relatively minor in the thicker 5.81 m and 29.32 m soil thickness versions. Consequently, I will focus my discussion on the results with the thin soil model configurations[8].

The 0.27 m configuration shows greater mean near surface ground temperatures nearly everywhere where there is permafrost and seasonally frozen ground (top left panel), while the 1.06 m configuration is typically cooler throughout the same region (second panel, left column) with the exception of the southern limit of permafrost and the high arctic where ground temperatures are slightly warmer than the standard configuration. Both thin soil configurations also show interesting changes in the the annual range of near surface soil temperatures: the range strongly increases in the regions of permafrost where the active layer is relatively thick (and there are extensive amounts of freezing and thawing of the top layers of the ground, annualy), while the range in ground temperatures is typically strongly reduced, elsewhere.

In understanding these changes, it is helpful to compare the constant thermal conductivity and heat capacity of the bedrock against the range of values typically shown in soil. The bedrock was taken to be granitic in nature and its thermal conductivity set to 3.4 W $m^{-1}$ $K^{-1}$ and volumetric heat capacity to 2.4 x $10^6$ J $m^{-3}$ $K^{-1}$. Dry soil thermal conductivities for soil layers in the model globally range between 0.187 - 0.270 W $m^{-1}$ $K^{-1}$. Thus, the thermal conductivity of bedrock is uniformly higher than that of soil (including soil whose pore spaces might be completely filled with ice as the thermal conductivity of ice is set at 2.24 W $m^{-1}$ $K^{-1}$). Dry soil heat capacities in soil layers in the model globally range between 0.91 - 1.47 x $10^6$ J $m^{-3}$ $K^{-1}$ for mineral soils and the dry soil heat capacity is 0.25 x $10^6$ J $m^{-3}$ $K^{-1}$ for pure organic soil and the specified volumetric heat capacities of water and ice are 4.18 x $10^6$ J $m^{-3}$ $K^{-1}$ and 2.10 x $10^6$ J $m^{-3}$ $K^{-1}$. Consequently, the volumetric heat capacity of bedrock may be higher than soil in soils with relatively low moisture contents, but may be lower than that of soil in regions with relatively high unfrozen moisture contents. However, when phase changes occur, the effective heat capacity soars by several orders of magnitude.

---

[8] A more detailed examination of the differences between the two addition thick soil configurations and the standard configuration was conducted and it was found that additional metrics too showed relatively minor differences as compared with the thin soil configurations of the model.

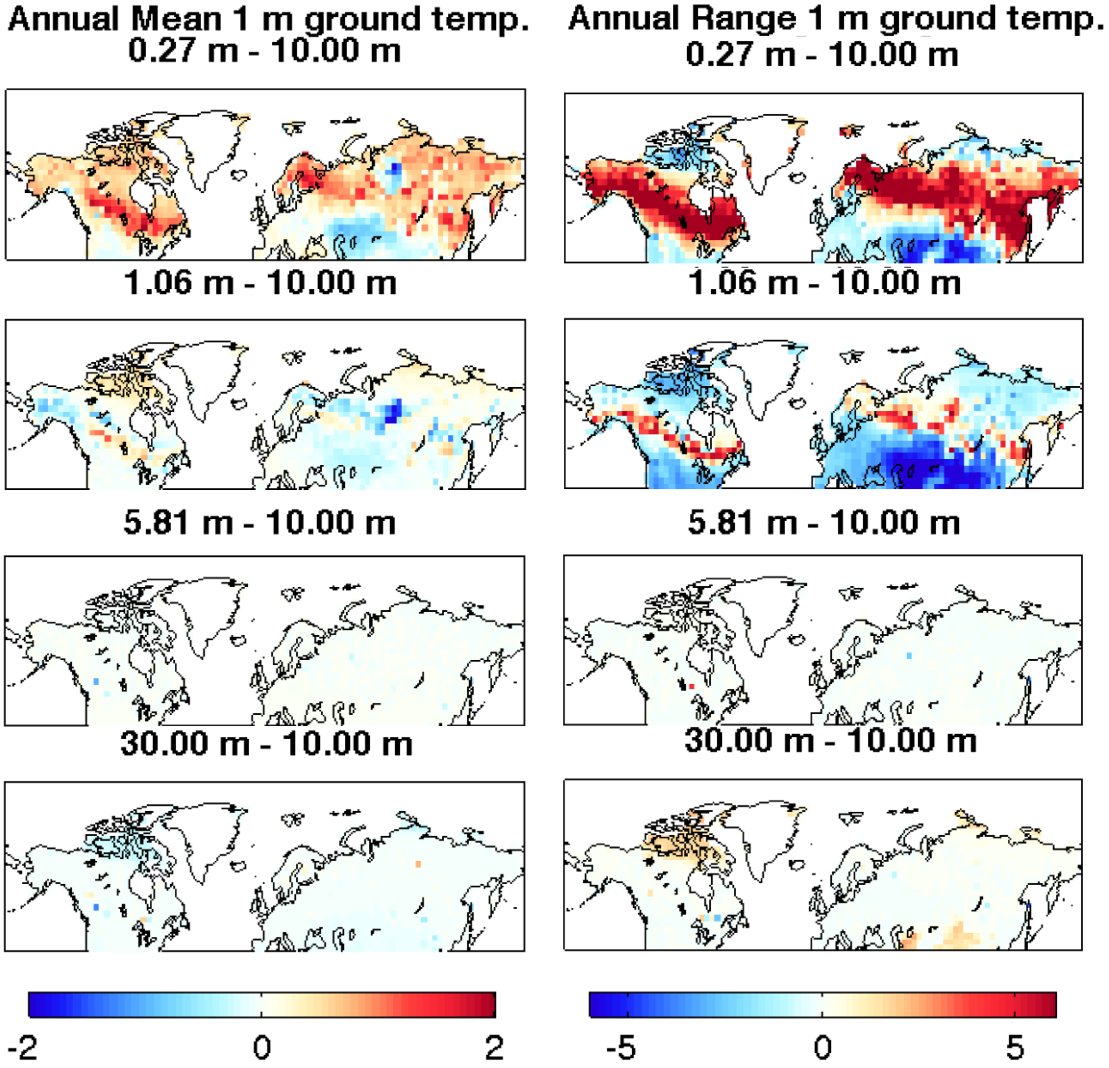


Figure 3.33: Changes (°C) in annual mean and range in near surface (1 m) ground temperature in model configurations with varying depth to bedrock, relative to the standard (10.00 m) depth to bedrock.

When phase changes are included as well, then the bedrock heat capacity will be substantially lower than the soil heat capacity. As well, the bedrock layers show no variation in thermal parameters in response to changes in moisture content or form throughout the year.

To explain the changes in annual mean soil temperature and range, I examined winter and summer mean near surface ground temperatures relative to the standard version of the model (Figure 3.34). It is also helpful to refer back to the top panels of Figure 3.28 which illustrates the mean and range in the soil effective thermal diffusivity - regions where there are extensive phase changes in the top layers of the soil in the standard configuration of the model are highlighted by a very low annual mean diffusivity and high percentage range in diffusivity. In both thin soil configurations of the model, the top soil temperatures are substantially cooler in winter and substantially warmer in summer in regions where there are substantial phase changes in the top of the soil column throughout the annual cycle. Here, the total heat capacity in the near surface layers of the ground owing to a reduction of latent heat effects on account of there being less moisture to freeze/thaw in the ground column. As well, the mean near surface thermal conductivity in the thin soil configurations is higher than in the standard configuration of the model as the bedrock is close to the surface and bedrock layers have higher thermal conductivity than soil layers. The overall effect of both of these changes is to increase the ground thermal diffusivity year round. Thus, in regions of extensive annual phase changes, the thin soil configurations of the model allow for substantially greater heat flow out of the soil in the winter months (cooler winter temperatures) and greater heat flow into the soil in summer months (warmer summer temperatures) and the results (Figure 3.34) support this contention.

In very high latitudes, winter soil temperatures are warmer than the standard configuration of the model in both thin soil configurations. Here, the active layer is relatively thin and so latent heat effects have a substantially smaller impact on ground thermal behaviour. Consequently, in the absence of significant phase changes, the rock layers' higher heat capacity of rock layers increases the column heat capacity during winter relative to the standard version of the model, resulting a reduced rate of soil cooling and warmer soil temperatures. The same effect is seen in the midlatitudes in regions that are not subject to seasonal freezing / thawing where again the rock layers have a higher heat capacity. In the summer months, the reverse pattern in seen in the model configuration with a 1.06 m soil depth, but not in the 0.27 m configuration.

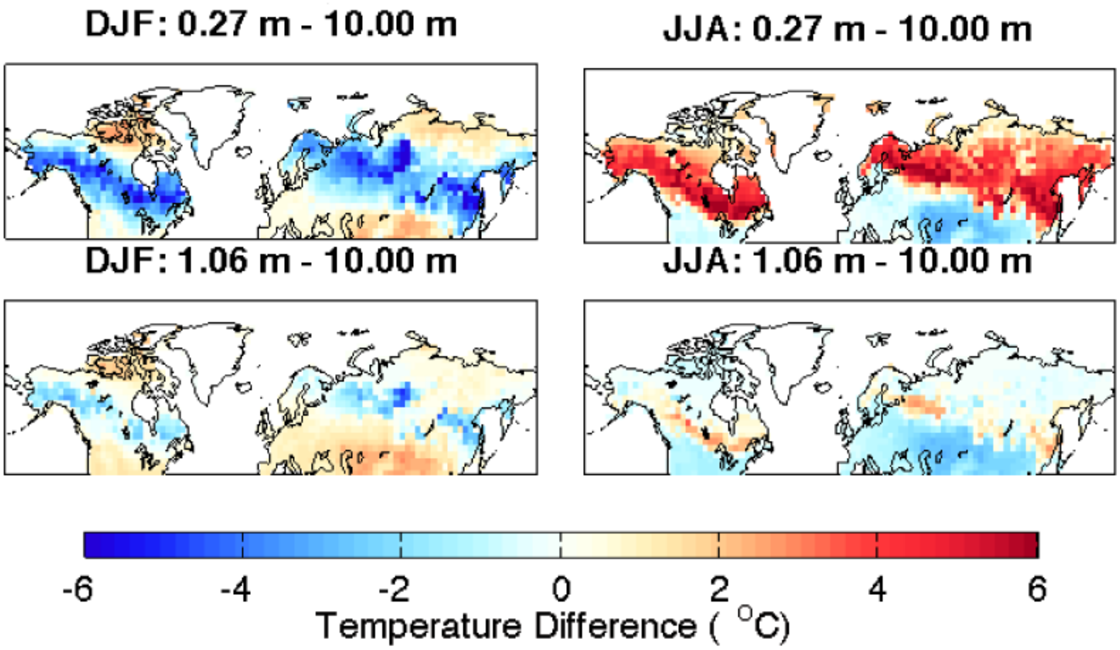


Figure 3.34: Winter (DFJ) and summer (JJA) changes in near surface soil temperature in thin soil configurations of the model relative to the standard configuration of the model (soil depth 10.00 m).

Soil temperatures are, however substantially cooler than the standard depth model in the 0.27 m configuration (and the 1.06 m configuration) at the start of the fall (not shown) which suggests that the timing of the seasonal periods of peak heat loss / gain may also be shifted as a result of the changes to the soil thermal parameters.

In the case of the 0.27 m configuration of the model, the net annual effect of the altered depth to bedrock is an overall warming of the soil from the southern boundary of seasonally frozen ground northwards (Figure 3.33, top left panel). Conversely, in the 1.06 m soil depth configuration (Figure 3.33, second panel, left column), there is a net cooling throughout the seasonally frozen ground and permafrost regions with the exception of the the southernmost limit of permafrost and the highest latitude cells. In general, the amplitude of temperature changes in the top ground layers resulting from changes in the thermal parameters of the deeper ground layers becomes less and less significant as these changes are made deeper and deeper in the soil column - a sensible result since the magnitude of the annual cycle in temperature and in heat flow tends to decrease exponentially with soil depth.

Both thin soil configuration show an increased range in soil temperatures (right column, Figure 3.33) in the regions subject to substantial soil freeze/thaw in the thin soil configurations of the model as one might anticipate given the observed seasonal soil temperature changes. Outside of these regions, the annual range in soil temperatures is decreased. Turning to the comparison of permafrost properties, I then plotted the active layer thickness of the five soil thickness configurations of the model, as shown in Figure 3.35. The five different active layer plots are very similar with the exception of the 0.27 m plot where the ALT is as much as an order of magnitude greater than in the other versions. The extensive area in this plot where the active layer thickness is much higher than in the other configurations matches the regions in Figures 3.34 and 3.33 where the summer soil temperature is substantially warmer than the standard configuration and where the annual range in near surface soil temperature is substantially greater. The substantially deeper ALTs in the 0.27 m configuration of the model also highlight the key role that moisture in the soil column plays in restricting the annual range of soil temperatures in regions that experience annual freezing and thawing. In the 1.06 m configuration, ALTs are deeper along the southern limit of permafrost where this configuration showed warmer summer temperatures and a greater range in soil temperature relative to the standard version. The ALT plots of the 5.81 m and standard configuration are essentially identical while the 29.32 m configuration reveals deeper ALTs at the southern most extent of permafrost where the model simulates more extensive permafrost relative to the standard version of the model.

Previously, the standard configuration of the model had been shown to show ALTs that were consistently too low relative to observations. This has been attributed in part to the coarse vertical resolution of the soil layers in the model but as the above plot suggests, if the standard configuration soil thickness is systematically too high, then this may also explain the ALT bias. To investigate this assertion in more detail, I produced scatter plots comparing the simulated ALT in the 0.27 m, 1.06 m, 10.0 m and 29.32 m configuration of the model against observations - these are displayed in Figure 3.36.

The ALT scatter plots generally reflect the trends shown in spatial ALT plots of Figure 3.35. The 0.27 m ALT plot is substantially different than the other plots while the 1.06 m, 10.00 m and 29.32 m plots are fairly similar with the exception of a few sites located along the bottom boundary of permafrost where the 1.06 m configuration simulates substantially deeper ALTs than the other model configurations. These points excepted, the latter three plots still generally tend to show the

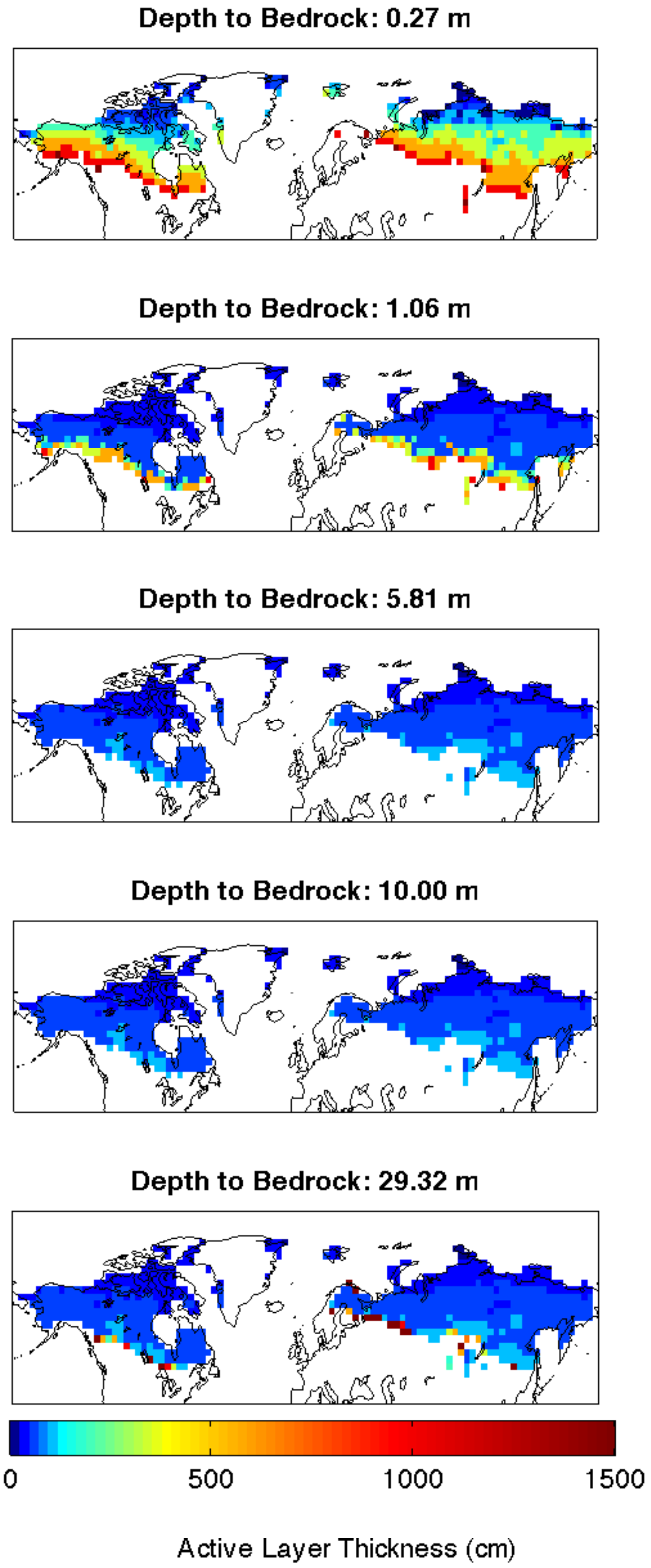


Figure 3.35: Active layer thickness in five model configurations having different depths to bedrock.

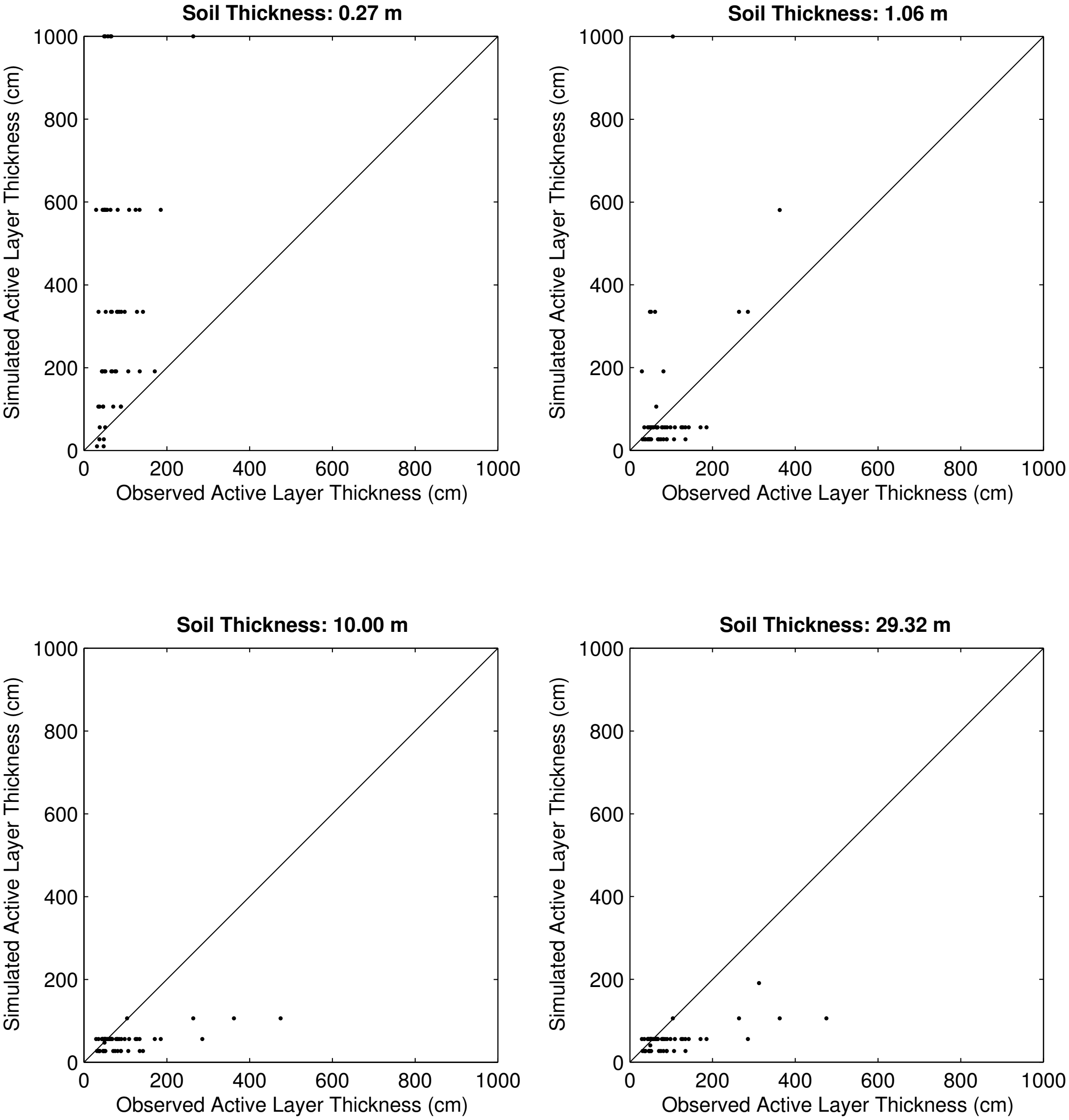


Figure 3.36: Scatter plot comparing simulated active layer thicknesses against observations for four model configurations having varying soil thicknesses.

systematic shallow ALT bias that has already been attributed, at least in part, to the coarse vertical resolution of the model and the tendency of soil temperatures to be too low in U3.0E. In contrast, the ALTs simulated in the 0.27 m configuration are substantially higher than observed at nearly all soil sites suggesting that here the soil thickness is systematically too low. Comparing the simulated ALTs between the 0.27 m and 1.06 m configuration also dramatically indicates that to accurate simulate the active layer thickness in permafrost, proper knowledge of the soil thickness may be essential, especially if bedrock is relatively close to the surface (say within 1 m of the surface). These results also suggest that, not only is accurate representation of soil thermal parameters important to accurately represent permafrost, but bedrock thermal characteristics may also be important - especially if the bedrock is relatively close to the surface.

Finally, I computed the extent of permafrost, seasonally frozen and intermittently frozen ground in the five soil thickness configurations (Table 3.6). The extent of seasonally and intermittently frozen ground is comparable in the five versions though the permafrost extent differs significantly, particularly in the shallow soil depth configurations. Here, the permafrost extent is considerably reduced relative to the deeper soil configurations, with the reduction occurring in the southern limits of permafrost. In these regions, permafrost, where present, occurs in relatively thin layers fairly close to the surface (resulting in a low depth to base of permafrost - see Figure 3.17). Consequently, as a result of the enhanced summer ground warming in the thin soil model configuration, thin layers of permafrost present in these regions in the standard configuration are able to thaw during the summer. One should also recall here, that such layers - when present in the 0.27 m configuration - may not necessarily be soil layers, as the term permafrost refers to any subsurface layer whose temperature does not exceed 0 °C throughout the year. The reverse effect presumably explains the slightly enhanced permafrost extent in the 29.32 m configuration of the model; a slight reduction in summer ground warming here allows for additional regions of permafrost to be present in the southern latitudes.

## 3.4 Summary

In this chapter, I examined the permafrost distribution in versions UVic 3.0M and 3.0E of the model. I firstly examined model climate biases in the high latitudes of the Northern Hemisphere, with the following key findings:

| **Area (x $10^6$ km$^2$)** | **0.27 m** | **1.06 m** | **5.81 m** | **10.00 m** | **29.32 m** |
|---|---|---|---|---|---|
| Permafrost | 19.06 | 21.57 | 23.05 | 23.01 | 24.52 |
| Seasonally Frozen | 50.26 | 51.06 | 51.46 | 51.46 | 51.42 |
| Intermittently Frozen | 2.79 | 2.53 | 2.47 | 2.52 | 2.47 |

Table 3.6: Areas of permafrost and frozen ground with different depth to bedrock values.

- Relative to the ERA-40 reanalysis, UVic 3.0M shows some key regional temperature biases in permafrost regions. North-eastern margins of North America and Asia are persistently warm, with the warm bias particularly pronounced in the winter, while the Tibetan Plateau and land points adjacent to the North Atlantic Ocean tend to be too cool. UVic 3.0M does a good job at simulating the annual mean Arctic and sub-Arctic precipitation, but the seasonality is poorly captured, with the model simulating highest precipitation rates in the fall and winter and lowest rates in the summer - the opposite of the seasonal cycle as represented in ERA-40.
- Simulated snow covered area and snow water equivalent values from UVic 3.0M and UVic 3.0E were compared with satellite observations and the snow distribution from the ERA-40 reanalysis. Both configurations satisfactorily represent the annual cycle in total snow covered area, but show very different simulated snow water equivalent values. UVic 3.0M's distribution of SWE agrees reasonably well with observations, given the level of disparity between ERA-40 and satellite data, but UVic 3.0E substantially underestimates SWE. I suggest that these observations may be reconciled if the multilayer LSS simulates too much ablation with snow. Since UVic 3.0M simulates too-high fall and winter precipitation in high latitudes, the too-high ablation rate still yields a fairly reasonable SWE distribution. Conversely, without the anomalously high accumulation season precipitation rates, UVic 3.0E simulates SWE values that are too low. These results are key to understanding differences in the permafrost and seasonally frozen ground distribution between the two model configurations.

Next, I compared simulation of characteristics of frozen ground against available observations, concluding that:

- UVic 3.0M does a good job of representing the areal distribution of permafrost, simulating an area of 15.59 x $10^6$ km$^2$ which agrees well with the estimates of the total ground area containing permafrost as well as the area of continuous and discontinuous permafrost zones, though the modelled permafrost area is an overestimate due to issues with the land grid. In contrast, the permafrost area (23.01 x $10^6$ $km^2$) in UVic 3.0E is far too high and I attribute this to the low simulated snow thicknesses that result is too little thermal insulation in winter. The extent of seasonally frozen ground is too low in UVic 3.0M due to the influence of warm temperature biases but is well simulated in UVic 3.0E. Both configurations underestimate the extent of intermittently frozen ground.

- Both UVic 3.0M and 3.0E are able to simulate the broad latitudinal pattern in terms of changes in Active Layer Thickness, but simulated ALTs are consistently too shallow compared with observations. A sensitivity analysis comparing model configurations with different subsurface layer thicknesses suggests that the relatively coarse vertical resolution of the model in its standard configuration is likely the major reason for the shallow ALTs. UVic 3.0M simulates soil temperatures in good agreement with both shallow and deep borehole readings, though with a slight overall warm bias. I attribute this bias to the fact that most of the borehole sites lie in regions where the ESCM shows warm atmospheric biases. Conversely, UVic 3.0E shows a cold bias when simulated soil temperatures are compared with observations, likely due to that configuration's underestimate of lying snow. Finally, the observed depth to the base of permafrost in Canada is well simulated by UVic 3.0M, including realistic simulation of the location of the transition line from thin to thick permafrost, In UVic 3.0E, permafrost thicknesses are too great at southern sites and the transition line is located too far South.

Finally, I conducted a suite of sensitivity experiments with UVic 3.0E. The key results from each of the sensitivity experiments were as follows:

- **Land surface cover experiment**: The impact on regions of frozen ground of changing land cover class may be understood in terms of changes in surface albedo and surface roughness. I found that frozen ground distribution was generally fairly insensitive to changes in land surface cover save where vegetation was replaced by bare soil (or vice versa). Here, replacement of vegetation by bare soil generally resulted in a relatively strong warming of the ground and a reduction in the area of permafrost and seasonally frozen ground. Significantly, although the UVic vegetation distribution does not agree perfectly with the observation-based IGBP land cover data set, there was little change in the area of frozen ground in the sensitivity experiment where UVic land cover was replaced with IGBP land cover.

- **Snow thermal conductivity experiment**: In this experiment, I varied the snow thermal conductivity between two extremes representative of the range in thermal conductivity of snow throughout an annual cycle. My results here agree with those of an earlier study by Cook et al. (2007). An increase in snow thermal conductivity leads to increased loss of heat from the soil in winter months and cooling of the soil. The cooler soil temperatures persist into summer, driving a greater flux of heat into the soil owing to the greater temperature gradient between soil and atmosphere. The changes in soil temperature associated with variations in thermal conductivity did not greatly affect frozen ground areas, though the permafrost area was higher in the high conductivity configuration, and lower in the lower conductivity configuration. I speculate that the anomalously low snow depths in version U3.0E mean that, were this experiment to be repeated with U3.0M, the magnitude of impacts would be greater.

- **Resolution**: I developed three new resolution versions (half, double and quadruple) of the model to assess whether the limited vertical resolution of the model strongly affected simulations of frozen ground. All four configurations simulated comparable permafrost areas, but the seasonally frozen ground area was reduced in the half resolution version of the model, possibly because the thicker uppermost soil layer in this configuration of the model inhibited seasonal soil freezing. Changes in soil temperature between the resolution configurations were attributed to seasonal variations in thermal diffusivity in the different configurations, though the root cause of diffusivity variations was not identified owing to the complex dependence of that characteristic on the distribution and form of soil moisture content. The most significant difference between the different resolution configurations is an improved representation of active layer thickness as resolution increases. The quadruple resolution configuration shows a great reduction in shallow active layer bias over the standard version.

- **Depth to Bedrock**: One of the most uncertain characteristics of the land surface scheme is its uniform 10 m depth to bedrock. I developed two configurations of the model where the depth to bedrock was lower than the standard configuration and two configurations where the depth to bedrock was greater than the standard configuration. There were substantial differences between the thin soil configurations and the standard configuration, while differences between the standard configuration and the thicker soil configurations were negligible. Differences in the thin soil configurations may be explained by comparing the heat capacity of bedrock with that of soil. In regions where there are substantial phase changes on an annual basis, the annual range in soil temperatures is increased in the thin soil configurations as the heat capacity of bedrock is lower than that of soil subject to extensive phase changes; conversely, in regions where there were little or no phase changes on an annual basis, the annual range in soil temperatures decreases in the thin soil configuration as the heat capacity of bedrock is greater than that of soil. Both thin soil configurations show a reduction in the area of permafrost (especially in the case of the very thin soil configuration) and greater active layer thicknesses, especially along the southern limit of permafrost, but the area of seasonally frozen ground was comparable to the standard version. The permafrost and seasonally frozen ground area in the the thick soil configurations was also comparable to the standard configuration of the model. These results suggest that to properly simulate permafrost, detailed knowledge of soil thickness and bedrock thermal properties may be essential, especially in areas where bedrock is fairly shallow.

# Chapter 4

# Permafrost in a Warming World

In this chapter, the fully coupled ESCM (UVic 3.0M) is forced using data from the four recently released *Representative Concentration Pathways* in order to project the future distribution of permafrost over a range of potential radiative forcing scenarios. I first describe the generation of forcing fields for the model and then examine the response of temperature and precipitation fields in comparison with results from the IPCC Fourth Assessment Report. I next examine the rate of permafrost degradation in the four RCP scenarios and then conduct a suite of sensitivity experiments to determine the sensitivity of the permafrost distribution and rate of degradation to different configurations of the land surface scheme.

## 4.1 Representative Concentration Pathways

Representative Concentration Pathways (RCPs) are being used to drive climate simulations in support of the $5^{th}$ Coupled Model Intercomparison Project (CMIP5) and are successors to the scenarios from the IPCC *Special Report on Emissions Scenarios* (SRES - as described in Nakicenovic et al. (2000)) which have been used in previous intercomparisons. Model simulations forced by the SRES scenarios have also been extensively examined in the third and fourth IPCC Assessment Reports. The RCPs were generated as part of a new "parallel" development process wherein four radiative forcing trajectories of interest were selected to support modelling of a range of possible future climates.

Four emissions scenarios tracking greenhouse gas concentrations, land cover change and other factors influencing radiative forcing were identified from the literature as plausible pathways that could lead to each radiative forcing trajectory. As many possible scenarios could lead to the same radiative forcing trajectory, the resulting pathways are thus labelled as "representative" (Moss et al., 2010). The RCPs are named for their approximate year 2100 radiative forcing. This approach thus allows for climate modellers to conduct simulations following the forcing trajectories at the same time as integrated assessment modellers develop an ensemble of socioeconomic and emissions scenarios consistent with the RCP pathways. A timeseries of the radiative forcing associated with each RCP is shown in Figure 4.1, drawn from Meinshausen et al. (2011).

The focus of many modelling studies is on climate change over the coming century. This is reflected in the RCPs as the four concentration pathways include detailed forcing specifications to year 2100; they have also been extended to year 2500 (following *Extended Concentration Pathways* - ECPs) for the purpose of experiments examining long-term changes in the climate system, using a set of simple extension rules consisting of assumptions made about stabilization of emissions and / or concentrations of greenhouse gases as outlined in Meinshausen et al. (2011). While most modelling groups that investigate the warming response of permafrost have focussed exclusively on changes the $21^{st}$ century, I have opted to track changes over the entire period to 2500 as I believe that doing so will give a clearer indication of the response time of permafrost to transient climate forcings and give a better estimate of the potential long-term changes to the Arctic land surface associated with a spectrum of possible climate trajectories.

As earlier model experiments focusing on changes in permafrost over the coming century have been conducted using the SRES scenarios as forcing fields, it is helpful to compare the radiative forcing trajectories of the RCPs to the key SRES emission scenarios. Figure 4.2 shows projected radiative forcing and changes in global mean temperature for the SRES scenarios. RCP8.5 follows a radiative forcing trajectory comparable to the high emissions A2 and A1F scenarios which show year 2100 mean radiative forcings of around 8 and 9 $Wm^{-2}$, respectively. RCP6 and RCP4.5 are similar to the A1B and B1 scenarios, respectively which have respective year 2100 mean radiative forcing values of about 6.0 $Wm^{-2}$ and 4.5 $Wm^{-2}$. RCP3-PD has a radiative forcing trajectory lower than any of the illustrated SRES scenarios, in part because the SRES scenarios do not allow for any efforts to mitigate climate change, while the RCPs are designed to allow for this possibility (Moss et al., 2010).

In the experiments outlined in this chapter, different model configurations were

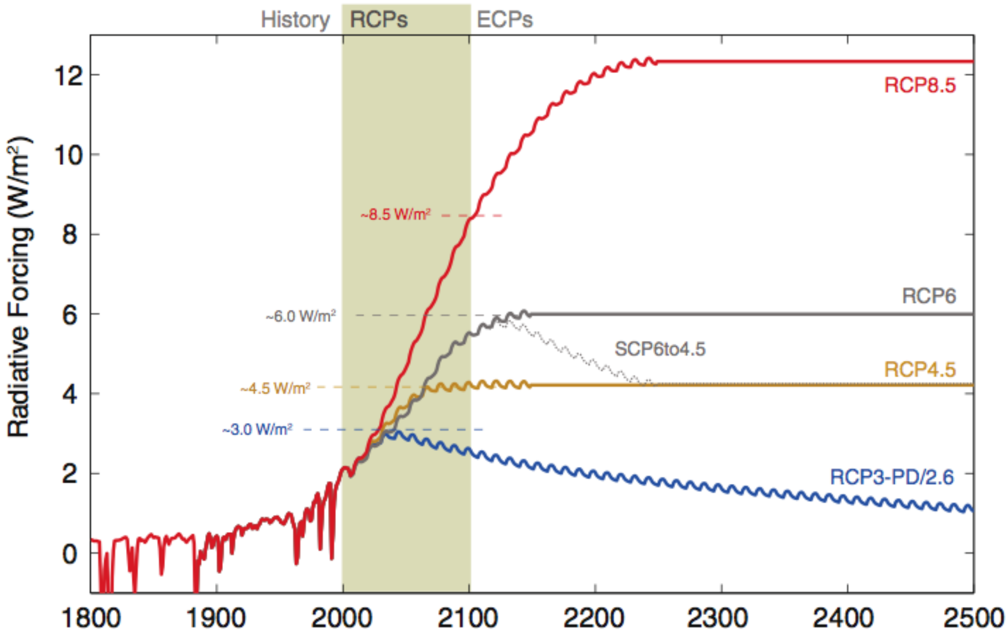


Figure 4.1: Radiative forcing (relative to 1765) over the historical period, from 2000-2100 following Representative Concentration Pathways and from 2100-2500 following Extended Concentration Pathways. Figure taken from Meinshausen et al. (2011).

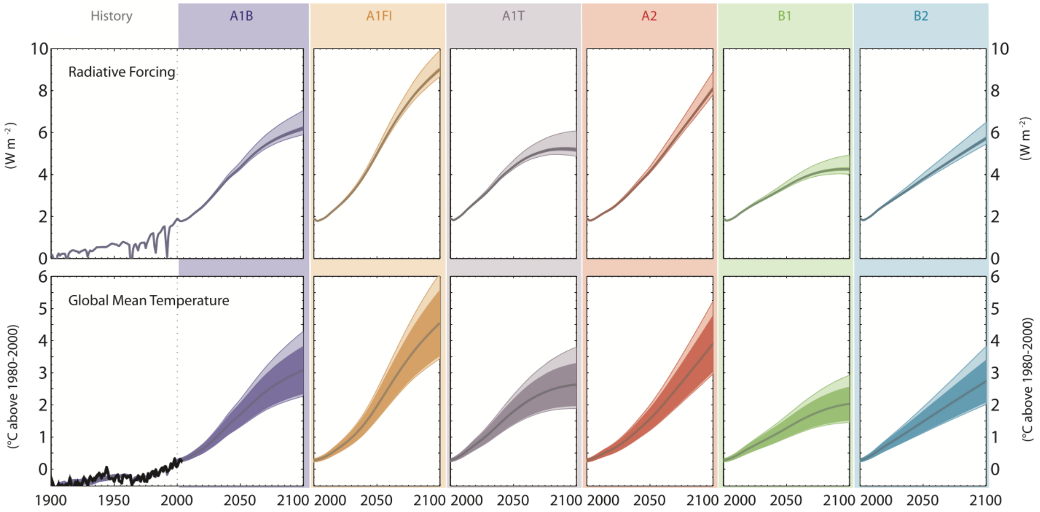


Figure 4.2: Projections of radiative forcing and global mean temperature change in key SRES emissions scenarios made using a simple climate model tuned to 19 General Circulation Models. This is a portion of Figure 10.26 from Meehl et al. (2007). Dark colored lines represent mean projections for the mid-range carbon cycle. Light shading represents the change in forcing and air temperature if carbon cycle feedbacks are lower or hight. The dark shading in the temperature plot represents ± 1 standard deviation in the 19 model tunings.

first allowed to reach an equilibrium state under perpetual year-1800 forcing for 4,000 years. I then forced the models over the historical period and to year 2500 following the extended Representative Concentration Pathway data by specifying atmospheric concentrations of $CO_2$, $CH_4$, $N_2O$, 12 fluorinated substances controlled by the Kyoto protocol and 16 ozone depleting substances controlled by the Montreal protocol. The model calculates only the radiative effect of $CO_2$ internally, but can account for the aggregated radiative forcing of additional greenhouse gases. I used the equations provided by Ramaswamy et al. (2001) (Table 6.2) and the radiative efficiencies of Forster et al. (2007) (Table 2.14) to determine this aggregated forcing field. Greenhouse gas concentrations over the historical period (1800 - 2005) were taken from data provided by Meinshausen et al. (2011); beyond 2005, I used the CMIP5 recommended data for the harmonized RCPs: RCP3-PD (van Vuuren et al., 2007), RCP 4.5 (Clarke et al., 2007; Smith and Wigley, 2006; Wise et al., 2009), RCP 6.0 (Fujino et al., 2006) and RCP 8.5 (Riahi et al., 2007).

Fractions of grid cells containing crop and pasture land were specified with data from the Land-Use Harmonization project which smoothly connects historical reconstructions of land use with future projections from Integrated Assessment Model implementations of RCPs (Hurtt et al. (2009)), which extend to year 2100. I handled crop and pasture land in the model by only allowing $C_3$ and $C_4$ grass plant functional types to exist and compete within the fraction of gridcells occupied by crops and pastures. Beyond year 2100, I held crop and pasture land fixed at their year 2100 distribution.

I also accounted for the direct effect of anthropogenic sulphate emissions using aerosol concentrations from the Community Atmospheric Model (Lamarque et al., 2010). I estimated the anthropogenic sulphate aerosol burden by subtracting the base year (1865) sulphate burden from the total sulphate burden in each year. I then converted this anthropogenic burden into a anthropogenic sulphate optical depth by multiplying by a specific extinction cross-section factor of 8 $m^2/gm$ (Reader and Boer, 1998) which is then used in the model to determine a perturbation to the local surface albedo as described by Matthews (2004). I also account for solar orbital variations (Weaver et al., 2001) and historical volcanic eruptions as globally averaged radiative forcings (Matthews, 2004).

## 4.2 Standard Experiment

In this chapter, I compare the response of different model configurations following the RCP scenarios. Each of these model configurations was based upon a *standard* model configuration in which the land surface scheme is fully coupled to all model components (labeled as version 3.0M in previous chapters). I first present results for this standard configuration followed by a sensitivity analysis that compares the response of different configurations of the model with the standard configuration.

### 4.2.1 Global Temperature and Precipitation Response

Before assessing the response of regions of frozen ground under the RCP scenarios, I first begin by comparing the ESCM's global temperature and precipitation response against results from comparable SRES emissions scenarios presented in the Fourth IPCC Assessment Report (AR4) to assess how well the model reproduces responses shown by ensembles of more sophisticated Atmosphere-Ocean General Circulation Models (AOGCMs). Here, changes in the model's climate are measured relative to the mean climate for the period from 1980-1999. The AR4 results presented here are derived from an ensemble of 23 global AOGCMs and focus on climate change over the $21^{st}$ Century following the B1 (low), A1B (medium) and A2 (high) SRES emission scenarios.

Figure 4.3 shows ensemble mean changes in global surface air temperatures for the decades 2020-2029 and 2090-2099 relative to 1980-1999. The sign of SAT change is positive everywhere and shows a high degree of spatial variability. While the patterns of surface warming are quite similar across scenarios early in the $21^{st}$ Century (centre column, Figure 4.3), end of century warming is very clearly scenario dependent (right column, Figure 4.3) and demonstrates that uncertainty in future anthropogenic radiative forcing is a key contributor to the overall uncertainty in future climate projections. Warming tends to be greater over land, increasing from coasts to continental interiors. The warming is also most pronounced in the high latitudes of the Northern Hemisphere, owing to the positive ice albedo feedback, with the exception of the North Atlantic Ocean, where deep convection allows for heightened absorption of heat by the ocean; this is also the case over the Southern Ocean. As shall be shown, this high latitude warming tends to be highly seasonal and is more pronounced in the winter months.

Equivalent plots were generated for the UVic ESCM for the RCP 3-PD, 4.5 (comparable to B1), 6.0 (comparable to A1B) and 8.5 (comparable to A2) scenarios and are shown in Figure 4.4. It should be borne in mind when comparing these plots, that the UVic ESCM is a less comprehensive climate model than many of the AOGCMs using in the IPCC AR4 intercomparison (with its atmospheric component, in particular, being less sophisticated). As well, individual models in the intercomparison may regionally differ substantially from the multi-model mean (e.g. see figures in Meehl et al. (2007), Supplementary Material). Like the AR4 plots, the pattern in spatial SAT change shown in the UVic ESCM is fairly similar across the four RCP scenarios in the early part of the $20^{th}$ Century, but late-century warming varies considerably according to scenario. Land warms at a faster rate than the ocean, though the land-ocean warming contrast is markedly lower than in the IPCC multi-model ensemble mean, possibly as a consequence of the model's overly diffusive atmospheric transport (Mike Eby, personal communication 2011).

The model also shows the more pronounced warming at the high latitudes of the Northern Hemisphere, though the contrast in temperature change at these latitudes and the global mean is lower than in the IPCC multi-model mean. This not an unexpected result as the UVic ESCM is known to have a low polar amplification that is close to the bottom of the range of polar amplifications of the models used in the fourth assessment report. This is shown in Figure 4.5 which shows a plot of model temperature response following the A2 scenario at year 2100. The zonally averaged temperature change relative to the global mean temperature change is plotted as a function of latitude with the response of the UVic ESCM plotted in blue and the mean response of IPCC AR4 models displayed in black. The temperature response at high latitudes in both the Northern and Southern Hemispheres in the ESCM is considerably lower than the mean AR4 response, though is within the range envelope of the AR4 models. This implication of the reduced polar amplification in the ESM is that the rate of degradation of permafrost in the UVic ESCM may be slower than in other climate models forced following comparable warming scenarios. As with the IPCC models, the greatest warming in the Arctic occurs over the ocean as a consequence of degradation of sea ice and associated reduction in surface albedo.

Despite the regional differences between the UVic ESCM and the IPCC models' mean temperature response, the ESCM's global mean SAT change following the RCPs over the $21^{st}$ Century agrees well with the IPCC models following the comparable SRES scenarios as shown in Figures 4.6 and 4.7. The multi-model ensemble mean warming for the decade 2090-2099 relative to 1980-1999 is reported as 1.8 °C, 2.8 °C,

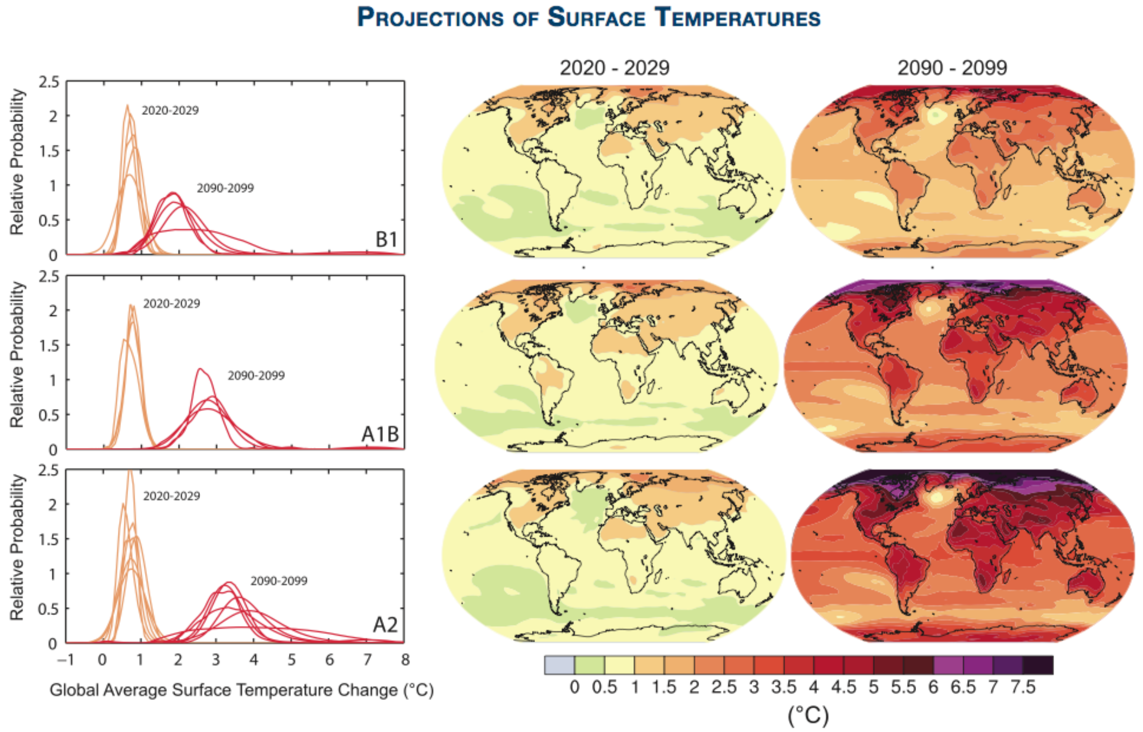


Figure 4.3: Projected surface air temperature (SAT) changes for the early and late $21^{st}$ century relative to the period 1980-1999 following three different emissions scenarios. Results presented are from a multi-model ensemble of Atmosphere-Ocean General Circulation Models (AOGCMs). Left columns: Relative probabilities of estimated global average SAT changes. Centre and right columns: multi-model ensemble average SAT change for the periods 2020-2029 and 2090-2099 relative to 1980-1999. Figure taken from IPCC (2007).

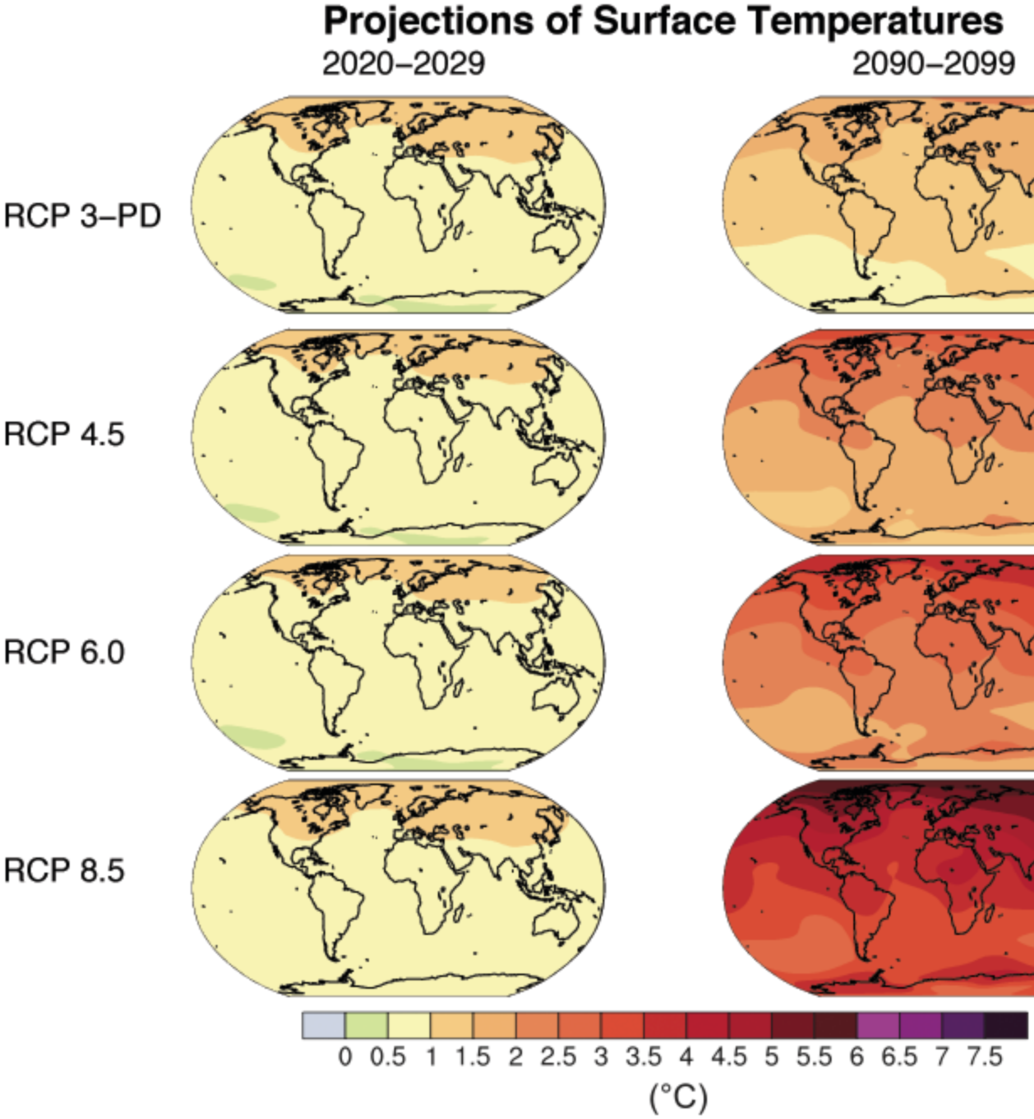


Figure 4.4: Patterns of global surface air temperature change in the UVic Earth System Climate Model. Temperature change following the RCP 3-PD, 4.5, 6.0 and 8.5 scenarios are shown for the decades 2020-2029 and 2090-2099 relative to 1980-1999.

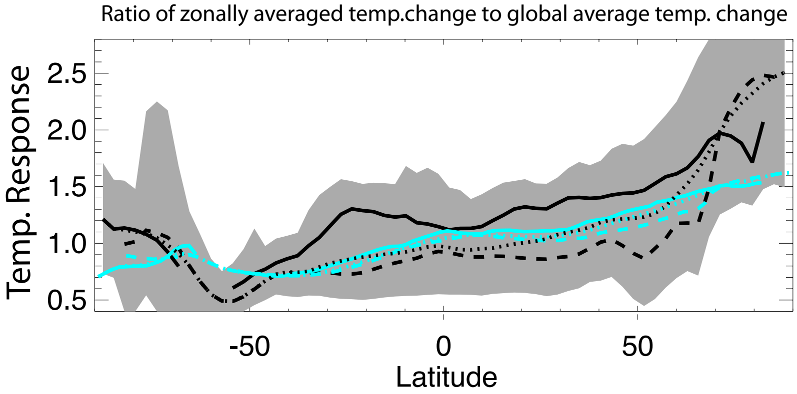


Figure 4.5: Zonally averaged temperature response (as a function of latitude) relative to global average temperature change at year 2100 following SRES A2 scenario in UVic ESCM and IPCC AR4 models. The figure indicates latitudes where warming is enhanced / reduced compared to the global average. The response of the UVic model is shown in the blue curves; black curves shows the mean response of IPCC AR4 models; grey shading represents the range in response of the AR4 models. The solid line represents temperature response over land, dashed is the response over the ocean and dotted is the global mean response. Figure produced by Michael Eby.

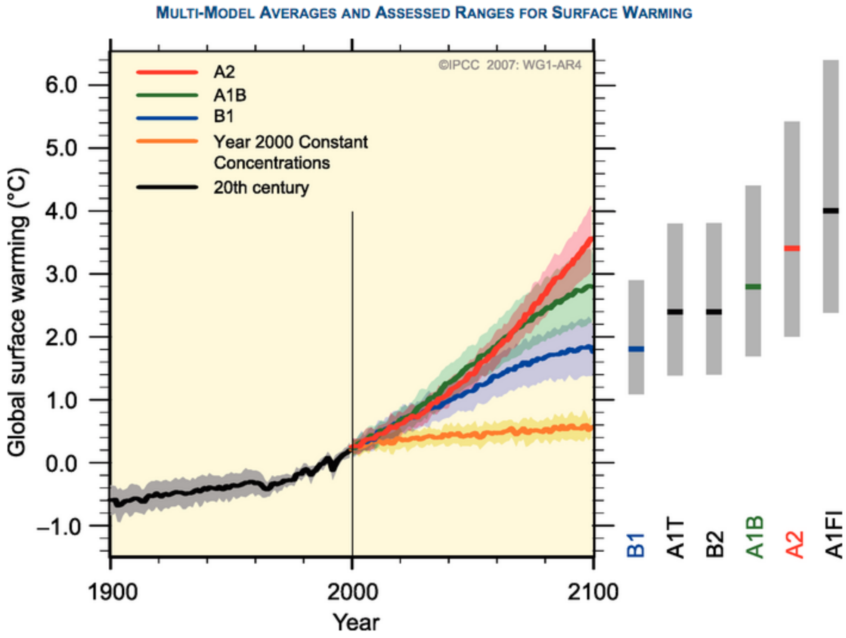


Figure 4.6: Time series of global mean surface air temperature changes from IPCC AOGCMs. Solid lines represent multi-model global averages relative to 1980-1999 and shading represents the ± 1 standard deviation range of individual model annual averages. The grey bars at the right indicate the best estimate (solid line) and likely range assessed for scenarios. Figure from IPCC (2007).

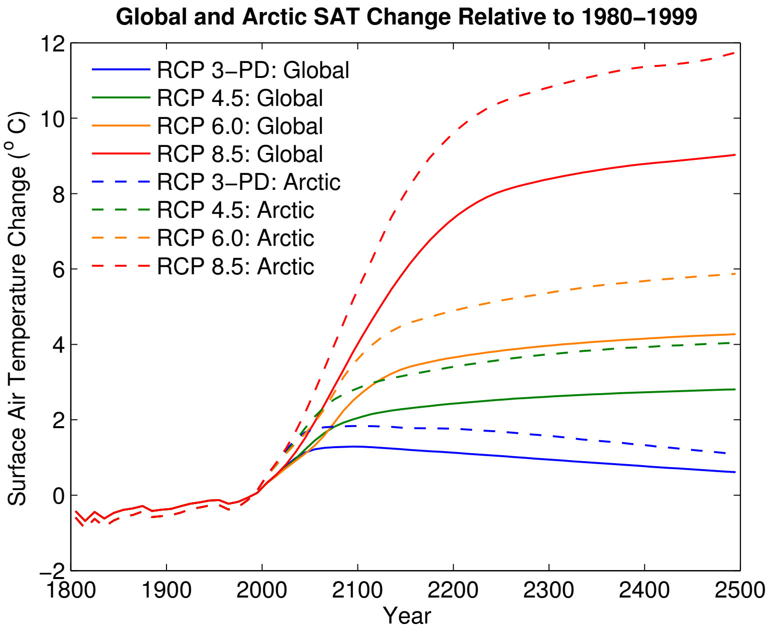


Figure 4.7: Timeseries of global and Arctic surface air temperature change in the UVic ESCM following four RCP scenarios. Temperature changes are relative to 1980-1999.

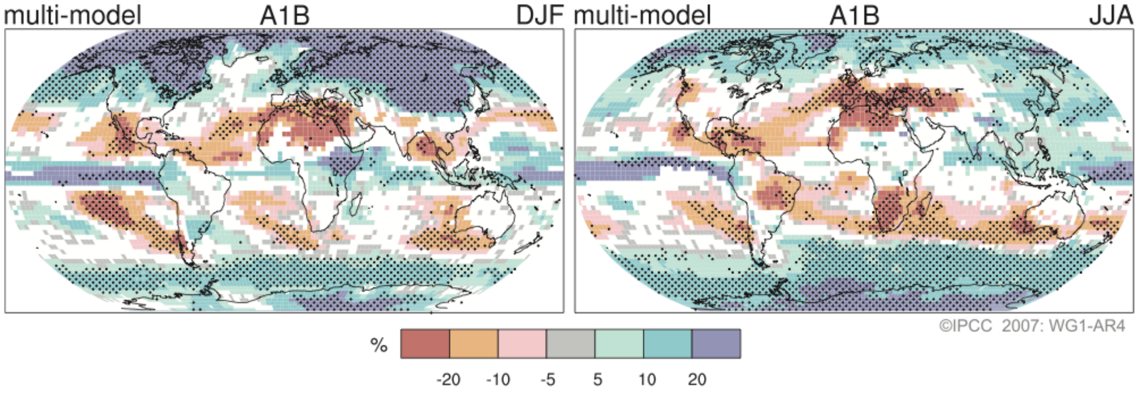


Figure 4.8: Projected changes in seasonal global precipitation rate for the late 21$^{st}$ century (2090-2099) relative to 1980-1999. The change in precipitation is expressed as a percentage change in precipitation relative to the reference period. The results presented are multi-model mean results for AOGCMs following the SRES A1B scenario. White areas represent regions where less than 66% of the models agree on the sign of the precipitation change; stippled areas are regions where more than 90 % of the models agree on the sign. Figure taken from IPCC (2007).

3.4 °C for the B1, A1B and A2 scenarios, respectively, with a likely range of end of century warming about this best estimate displayed by the grey bars on the right side of Figure 4.6. Using the UVic ESCM, I find an increase in global mean SAT over the same period of 1.29 °C (3-PD), 2.01 °C (4.5), 2.51 °C (6.0), 3.79 °C (8.5). The latter three are quite close to the multi-model means (and well within the likely range estimates) for the comparable SRES emission scenarios and suggests that the UVic model does a reasonable job of capturing global mean warming trends.

Figure 4.7 also shows projected Arctic warming in the ESCM, by plotting the mean surface air temperature at latitudes $\geq$ 60 °N. The plot illustrates that the projected rate of polar warming is clearly higher than the global average (by a factor of 1.3-1.4, in general). Again, though, the projected end of century Arctic warming in the ESCM (1.84 °C - RCP3-PD, 2.79 °C - RCP4.5, 3.47 °C - RCP6.0, 5.17 °C - RCP8.5) is lower than the multi-model ensemble mean Arctic warming in comparable scenarios (3.4 °C - B1, 4.9 °C - A1B, 5.9 °C - A2; Christensen et al. (2007)).

Models also predict an overall global increase in precipitation over the coming century although with pronounced seasonal and spatial variability. There is also a fair degree of inter-model regional disparity in terms of the sign and magnitude of precipitation change as indicated in Figure 4.8. Here, multi-model mean precipitation change at the end of the 21$^{st}$ Century (2090-2099) is shown for the A1B scenario for

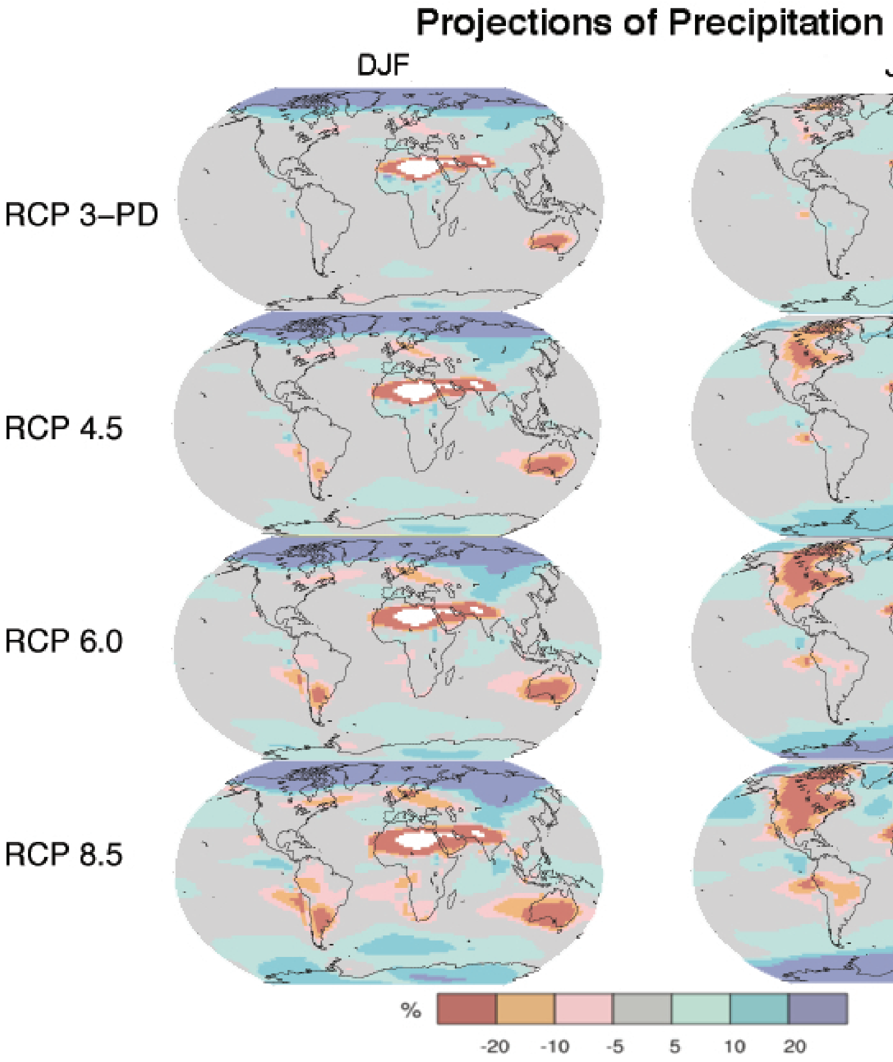


Figure 4.9: Patterns of DJF (NH winter) and JJA (NH summer) global precipitation change in 2090-2099 in the UVic ESCM. Precipitation changes are plotted as percentage changes relative to 1980-1999 change following the RCP 3-PD, 4.5, 6.0 and 8.5 scenarios. White areas denote regions where the ESCM simulates no precipitation in a particular season.

winter and summer. Areas in white indicate regions where less than 66 % of models agree on the sign of the precipitation change, while the stippled regions indicate areas where more than 90 % of models agree on the sign of the precipitation change. As there is a wide global range in the magnitude of precipitation rates (and change in these rates), changes are expressed as percentage changes in precipitation relative to the reference period. Broadly speaking, precipitation is projected to increase in the tropics and high latitudes as a consequence of increased atmospheric water vapour concentrations, with the latter owing to a higher level of water vapour transport from lower latitudes (IPCC (2007)). Conversely, there is a tendency towards a reduction in precipitation in the subtropics as a result of the increased poleward transport of water vapour out of this region.

The ESCM too simulates an overall global increase in precipitation, again with a complicated spatial pattern (Figure 4.9). Here, the RCP 6.0 output should be compared against the A1B plot shown in Figure 4.8. The model generally captures the broad spatial patten of increased precipitation in high latitudes and in the tropics, and reduced precipitation in the subtropics, but there are notable regional differences between the ESCM and IPCC multimodel ensemble mean changes. Many areas of precipitation change are less pronounced in the ESCM than in the IPCC plot, in part because the model's advective wind feedback option was not enabled. Consequently, there is no evolution in the surface wind fields in the ESCM and any changes in precipitation associated with changes in circulation in a warming world will not be captured by the model.

The most notable discrepancy between the ESCM and the IPCC precipitation response occurs during the Northern Hemisphere (NH) summer months where the model simulates a strong reduction in precipitation from the NH midlatitudes poleward, over land. Here, the model simulates a precipitation change of opposite sign to the IPCC plot in a regions where there is typically very strong agreement between the AOGCMs on the sign of the change. As discussed in the previous chapter, while the ESCM does a reasonable job of simulating the total annual precipitation in the mid-to-high latitudes, it poorly represents the seasonality of precipitation, simulating peak precipitation in the NH winter and minimum precipitation in the NH summer - the opposite of what is generally observed. I suggest that whatever factor accounts for the poor representation of summer precipitation in theses regions in the ESCM might also explain the different summer precipitation response. This might be a result of too low moisture transport into the sub-Arctic and Arctic during the summer or perhaps the model's inability to represent summer convective rainfall (which would be expected to increase in summer months in a warmer climate). As well, if vegetation transpiration / land evaporation rates were too low, this might result in anomalous summer precipitation owing to too little precipitation from moisture recycling over land. At any rate, this anomalous NH summer high latitude precipitation signal merits future investigation as it may affect other aspects of the climate, including the amount of soil moisture storage and the resulting distribution of Arctic vegetation.

The projected change in SAT and precipitation rates has a strong seasonal dependence as shown in Figure 4.10. Here, the change in both variables is shown following the A1B scenario for the IPCC models and the suite of RCP scenarios for the UVic ESCM (again, RCP 6.0 should be compared against the A1B scenario). In both the IPCC model simulations and the UVic ESCM, the projected warming (right panels) is greatest in the NH fall and winter months as the air temperature over the Arctic Ocean responds to the reduced sea ice cover; conversely, temperature change in the summer months is comparably low as temperatures over the Arctic ocean remain close to the freezing point owing to melting of sea ice (Kattsov and Kallen, 2005). Compared to the IPCC simulations, the Arctic summer warming in the ESCM is comparable to the multimodel mean, but winter warming is substantially lower than the mean and the UVic model response falls outside of the total range (light grey) of the IPCC models in January and February.

The change in precipitation (left columns) is also greatest in the NH winter and fall. As discussed, the projected increase in Arctic precipitation is thought to be due to enhanced moisture transport from lower latitudes. Bengtsson et al. (2011) have demonstrated that the increase moisture transport into the Arctic is particularly strong in the fall and winter and this may be one reason for the relatively strong enhancement of precipitation in these seasons. Kaplan and New (2006) also note that the effect of an overall enhancement in moisture transport on precipitation would be expected to be greatest in the fall and winter when local sources of moisture from evapotranspiration are lowest and hence increases in moisture will have the largest relative effect. The change in precipitation in the ESCM agrees well with the IPCC multimodel mean in the NH winter and early spring, but the model response strongly diverges during the late spring through to the early fall owing to the anomalous ESCM precipitation change over land that is most pronounced in the summer months. Here, during the late spring and summer, the model shows a reduction in precipitation over the Arctic region, as opposed to the IPCC simulations which show a mild increase in

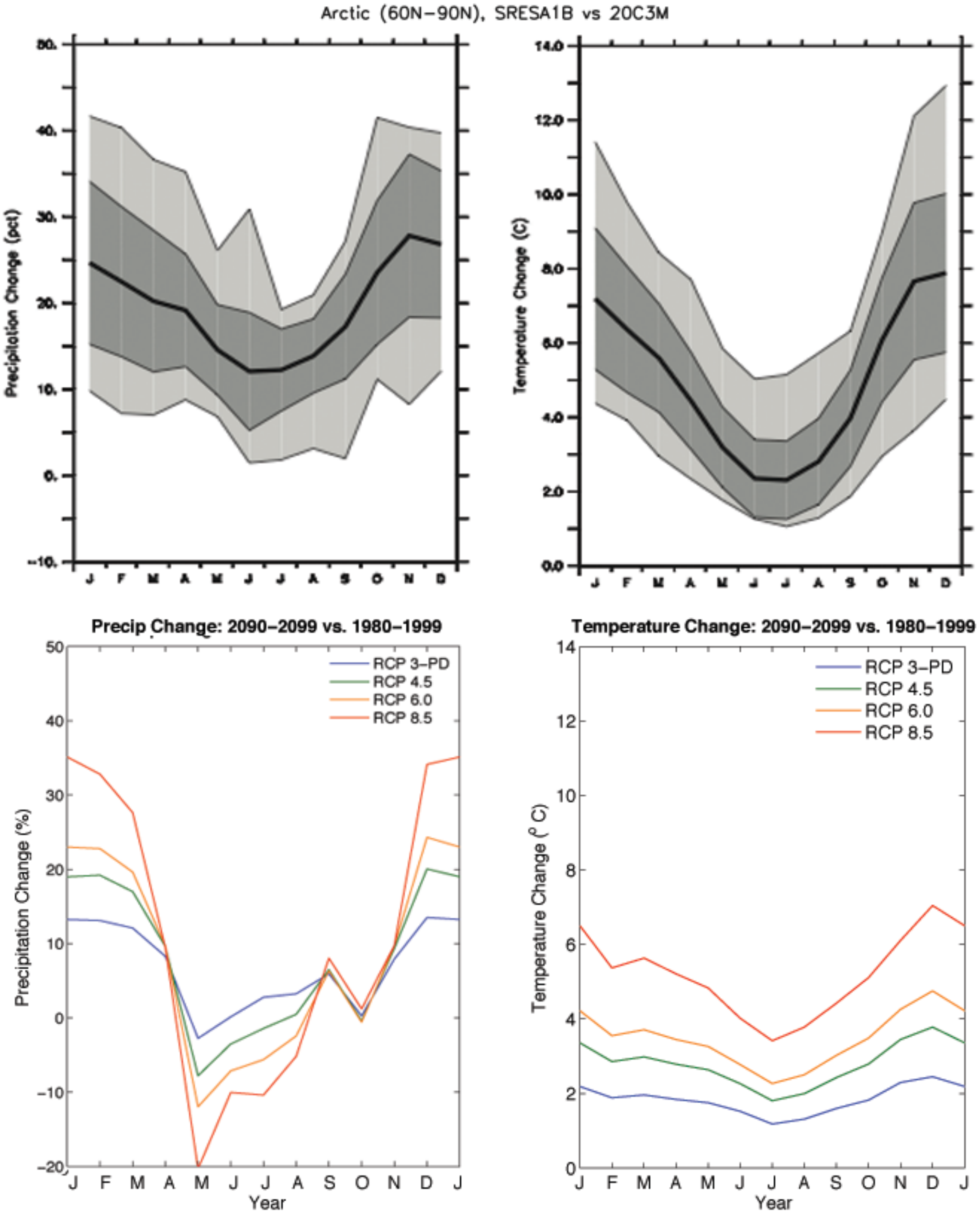


Figure 4.10: Annual cycle of Arctic (averaged over the area north of 60° N) mean percentage precipitation change (left column) and mean SAT change (right column). The top row is taken from Figure 11.19 in Christensen et al. (2007) and shows changes in 2080-2099 relative to 1980-1999 for models following the SRES A1B scenario. The solid line represents the ensemble mean of AOGCM models used in the simulation. The dark grey area represents the 25 and 75 % quartile values among these models, while the light grey area shows the total range of these models. The bottom panels shows equivalent timeseries for the UVic ESCM in 2090-2099 relative to 1980-1999 following the RCP 3-PD, 4.5 6.0 and 8.5 scenarios. Note that the scale in the IPCC precipitation plot is slightly different than that for the ESCM.

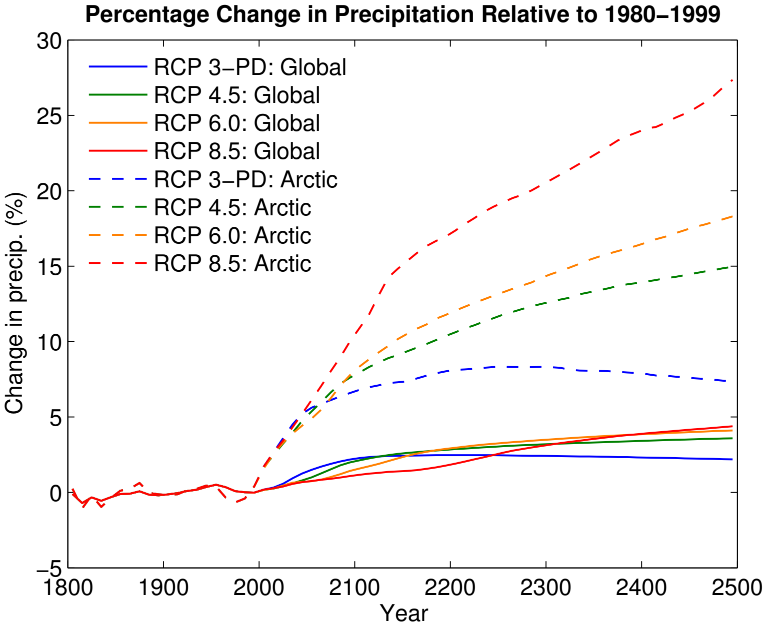


Figure 4.11: Timeseries of global and Arctic percentage precipitation change in the UVic Earth System Climate Model. Precipitation change following the RCP 3-PD, 4.5, 6.0 and 8.5 scenarios are shown relative to 1980-1999.

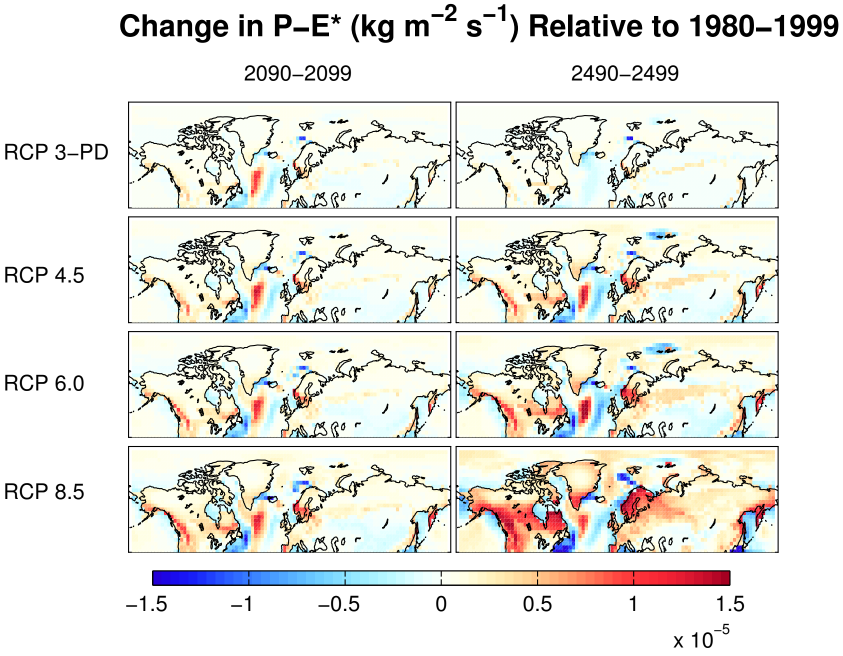


Figure 4.12: Changes in P - E* under RCP scenarios in the UVic ESCM. E* represents the sum of evapotranspiration, sublimation from bare soil and sublimation from snow. Changes are shown for the decades 2090-2099 and 2490-2499 relative to 1980-1999.

precipitation.

Although seasonal precipitation trends may have differing signs in the ESCM results, the model reveals a clear annual mean trend in Arctic precipitation following the RCP scenarios as shown in Figure 4.11 which shows the percentage change in global mean precipitation and Arctic precipitation relative to 1980-1999. There is a slow, gradual increase in the global precipitation rate, but the percentage increase in Arctic precipitation shows a much stronger positive trend in all of the RCP scenarios.

A warmer climate should also tend to lead to increases in evapotranspiration and sublimation, which can oppose the tendency towards a wetter land surface associated with heightened precipitation rates. Figure 4.12 shows the changes in P-E*[1] in the Northern Hemisphere mid-to-high latitudes. Nearly all land points show increases in P-E*, with a few showing very slight decreases at the end of the 21st Century (left panels) with these changes generally more pronounced by the end of the 25th Century at the conclusion of the simulation (with the exception of RCP-3PD). One might expect such an increase in P-E* to tend to lead to wetter land surface conditions, but as shown in the subsequent chapter, degradation of permafrost can lead to a shift in the vertical distribution of soil moisture and drying of near surface soil layers.

### 4.2.2 Vegetation Change

It is generally thought that projected high latitude warming and wetting will lead to shifts in vegetation distribution as these climate changes will tend to favour taller, denser vegetation. Consequently, it is expected that, over the coming centuries, tundra will tend to expand into existing polar deserts and forests will tend to expand northward into regions presently occupied by tundra (ACIA, 2004). A limited number of simulations of changes in Arctic ecosystems under warming scenarios have been conducted using regional- and global-scale models with interactive vegetation components. Their results are generally consistent with this hypothesis, showing northward shifts in both tundra and forest ecosystems, generally with the result that there is an overall increase in forest extent with a reduction in tundra, although model results may be highly dependent on simulation of rates of migration and succession of species (Epstein et al., 2007).

Figure 4.13 illustrates the response of the $C_3$ grass, shrub and tree plant functional

---

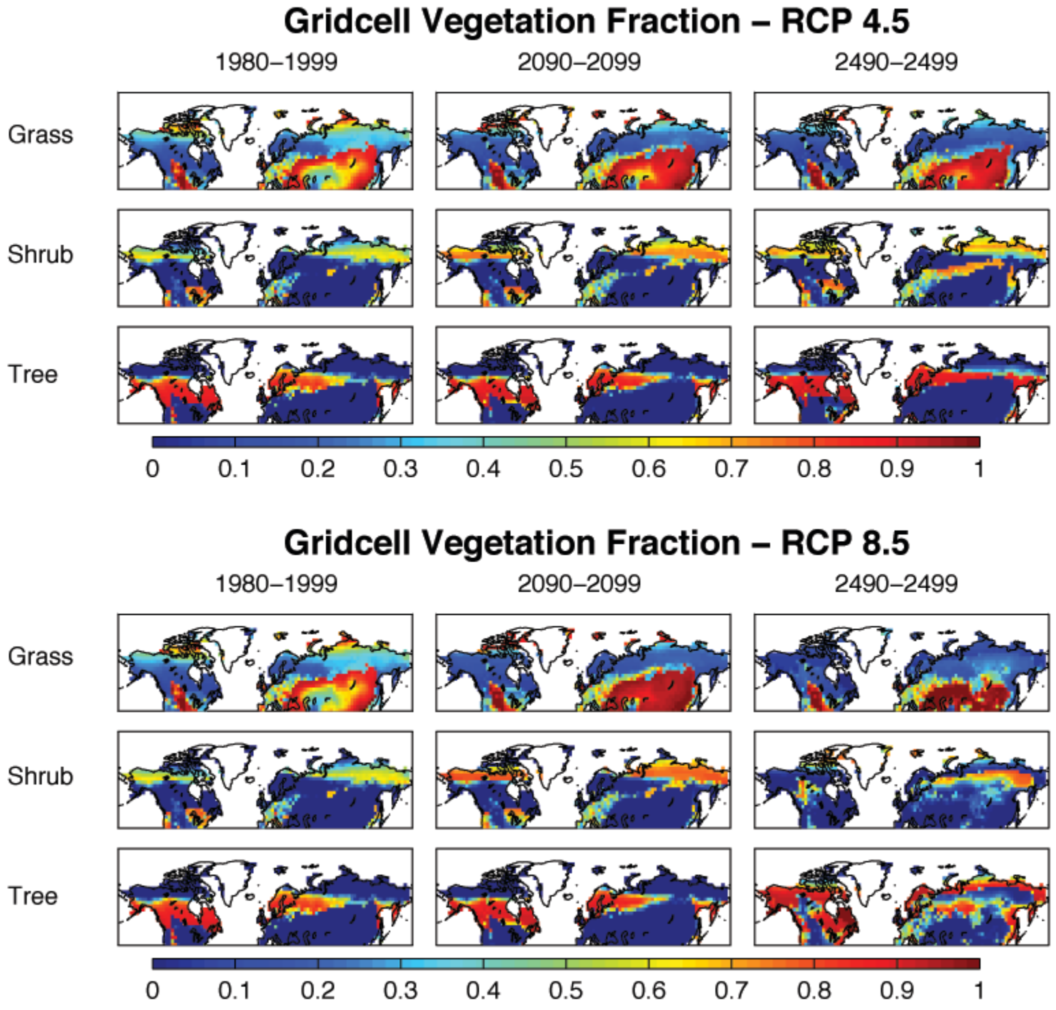
[1] E* is a term I used to represent all upwards losses of moisture from the land surface in high latitudes due to evapotranspiration, sublimation from soil and sublimation from snow. Thus: P-E* = precipitation - (evaporation + sublimation from soil + sublimation from snow)

types (the dominant PFTs at high latitudes) in the ESCM following the RCP4.5 and RCP8.5 scenarios. The left column in the figure represents the reference 1980-1999 vegetation state. It should be recalled when considering the vegetation warming response that this baseline vegetation distribution shows some discrepancies with observations. As discussed in the previous chapter (Figure 3.21), the model tends to simulate an overabundance of grasses at high latitudes, where bare soil should dominate, and an underabundance of trees in Central Siberia.

Vegetation changes at the end of the $21^{st}$ Century are shown in the central columns of the figure and are nearly identical in RCP4.5 and 8.5. There is a reduction in the areal extent of grasses in the Arctic and Subarctic, save at very high latitudes, where some bare soil regions are replaced by grasses. The loss of $C_3$ grass is associated with northwards expansion of shrubs into grasslands. There is also some expansion of shrubs along the southernmost limit of the needleleaf distribution, accompanied by loss of trees; this is particularly evident in a line extending from north-western Europe extending into central Siberia. At more northern latitudes, there is a decrease in the areal extent of shrubs accompanied by northwards expansion of trees.

Vegetation is noticeably different in the two scenarios at the end of the $25^{th}$ Century (Figure 4.13, right column). In the case of RCP4.5, the vegetation changes observed over the $21^{st}$ Century are slightly more pronounced, but otherwise the $21^{st}$ and $25^{th}$ Century vegetation fields are quite similar. In the case of RCP8.5, however, vegetation changes by 2500 are considerably more radical. Here, needleleaf trees are the dominant PFT nearly everywhere in North America, with the exception of a band extending along the Rockies and into the North-West territories where grasses (at the southern extent) and shrubs (at the northern extent) dominate. In Eurasia, trees also tend to be the dominant PFT with the exception of the highest latitudes and a broad band in central Siberia where shrubs are the dominant PFT. In fact, in some regions within this band, there is loss of needleleaf trees in regions where they are present in the baseline climate period

What prevents the establishment of needleleaf trees in these regions in RCP8.5 appears to be a reduction in unfrozen moisture at the top of the soil column (not shown). The band where needleleaf trees do not establish themselves lie in regions where the P-E* remains roughly constant or increases slightly, indicating that warming-induced drying is not responsible for the sharp reduction in moisture near the surface. Rather, it is the melting of near surface frozen soil moisture associated with permafrost thaw that is the cause. As permafrost degrades, the mean hydraulic conductivity of the soil


Gridcell Vegetation Fraction – RCP 4.5

Gridcell Vegetation Fraction – RCP 8.5

Figure 4.13: Changes in the distribution of grass, shrub and tree plant functional types simulated by the UVic ESCM following the RCP4.5 and 8.5 climate scenarios.

increases, reducing the soil moisture gradient between the surface and deeper layers. As a consequence, near surface soil moisture content decreases as the permafrost table deepens. This does not necessarily lead to dessication of the soil column as a whole, but rather a redistribution of soil moisture with depth

The root density in the model (which governs moisture availability to plants) specifies that root density decays exponentially with depth. The root density distribution does not evolve in response to a changing climate. As such, even though the moisture content of deep soil layers may increase, overall, plants experience more moisture stress if near surface moisture content strongly decreases. This appears to be the reason for the loss of needleleaf trees in certain regions of the present-day boreal forest. The tree PFT class may be overly sensitive to moisture in the model. I suspect that part of the reason that the model does a poor job at representing the present-day area of the boreal forest is in part due to the fact that the tree PFT is quite moisture sensitive and the presence of frozen moisture limits the range of the trees in permafrost regions to regions where there is a sufficient amount of unfrozen moisture available in summer months for the trees to thrive. Nonetheless, the potential limitation of tree expansion following warming and permafrost thaw is interesting in that it points to the possibility of a potential permafrost-vegetation feedback.

## 4.3 Response of Frozen Ground

I now investigate the warming response of the cryosphere in exposed land regions of the Northern Hemisphere, focusing on changes in the extent of permafrost (PF), seasonally frozen ground (SFG) and the maximum annual snow covered area (SCA). Figure 4.14 and Table 4.1 summarize these changes over the historical period and through to year 2500 following the four RCP scenarios. As in the previous section, I have chosen the period from 1980-1999 to serve as a reference period.

There are modest decreases in PF, SFG and SCA over the historical period, ranging between 1-2%, with each of these quantities decreasing nearly linearly up to the end of the $20^{th}$ Century. Short term interannual variability over the historical period, particularly apparent in the seasonally frozen ground plot, is predominantly associated with historic volcanic eruptions; this was determined by performing a separate experiment with the model with the volcanic forcing turned off. No assumptions are made about the rate of future volcanic eruptions and this short term variability consequently vanishes in the $21^{st}$ Century.

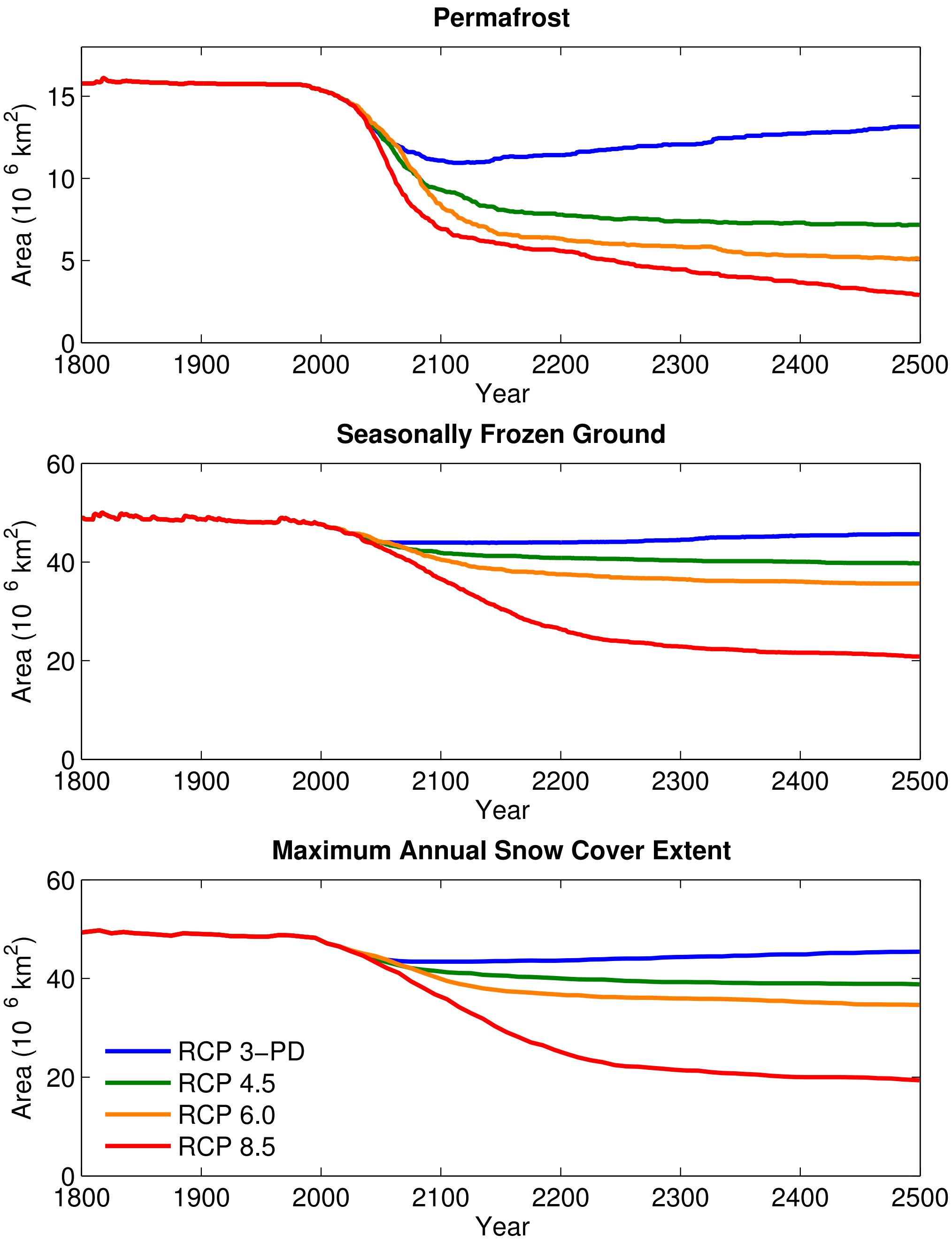


Figure 4.14: Changes in the total Northern Hemisphere areal extent of permafrost (top panel), seasonally frozen ground (middle panel) and annual maximum snow covered area (bottom panel) following Representative Concentration Pathways.

| | **Permafrost Area** ($10^6$ km$^2$) / **Percentage Change** (%) | | | |
|---|---|---|---|---|
| **Year** | **RCP3-PD** | **RCP4.5** | **RCP6.0** | **RCP8.5** |
| 1800-1809 | 15.78 (+1.35 %) | 15.78 (+1.35 %) | 15.78 (+1.35 %) | 15.78 (+1.35 %) |
| 1980-1999 | 15.57 | 15.57 | 15.57 | 15.57 |
| 2040-2049 | 12.98 (-16.65 %) | 13.09 (-15.92 %) | 13.27 (-14.77 %) | 12.48 (-19.85 %) |
| 2090-2099 | 11.13 (-28.51 %) | 9.41 (-39.56 %) | 8.71 (-44.05 %) | 7.19 (-53.82 %) |
| 2190-2199 | 11.42 (- 26.65 %) | 7.84 (-49.65 %) | 6.37 (-59.09 %) | 5.64 (-63.78 %) |
| 2290-2299 | 12.06 (-22.54 %) | 7.40 (-52.47 %) | 5.87 (-62.30 %) | 4.47 (-71.29 %) |
| 2390-2399 | 12.73 (- 18.24 %) | 7.30 (-53.11 %) | 5.31 (-65.90 %) | 3.73 (-76.04 %) |
| 2490-2499 | 13.15 (-15.54 %) | 7.18 (-53.89 %) | 5.11 (-67.18 %) | 2.95 (-81.05 %) |

| | **Season. Frozen Ground Area** ($10^6$ km$^2$) / **Percentage Change** (%) | | | |
|---|---|---|---|---|
| **Year** | **RCP3-PD** | **RCP4.5** | **RCP6.0** | **RCP8.5** |
| 1800-1809 | 48.72 (+1.67 %) | 48.72 (+1.67 %) | 48.72 (+1.67 %) | 48.72 (+1.67 %) |
| 1980-1999 | 47.92 | 47.92 | 47.92 | 47.92 |
| 2040-2049 | 44.21 (-7.74 %) | 44.15 (-7.87 %) | 44.55 (-7.03 %) | 43.49 (-9.24 %) |
| 2090-2099 | 43.93 (-8.32 %) | 42.10 (-12.15 %) | 40.83 (-14.86 %) | 37.20 (-22.37 %) |
| 2190-2199 | 44.00 (-8.18 %) | 40.84 (-14.77 %) | 37.57 (-21.60 %) | 26.73 (-44.22 %) |
| 2290-2299 | 44.44 (-7.26 %) | 40.36 (-15.78 %) | 36.57 (-23.69 %) | 22.93 (-52.15 %) |
| 2390-2399 | 45.32 (-5.43 %) | 40.08 (-16.36 %) | 36.05 (-24.77 %) | 21.63 (-54.86 %) |
| 2490-2499 | 45.64 (-4.76 %) | 39.76 (-17.03 %) | 35.64 (-25.63 %) | 20.86 (-56.47 %) |

| | **Max. Snow Covered Area** ($10^6$ km$^2$) / **Percentage Change** (%) | | | |
|---|---|---|---|---|
| **Year** | **RCP3-PD** | **RCP4.5** | **RCP6.0** | **RCP8.5** |
| 1800-1809 | 49.35 (+2.07 %) | 49.35 (+2.07 %) | 49.35 (+2.07 %) | 49.35 (+2.07 %) |
| 1980-1999 | 48.35 | 48.35 | 48.35 | 48.35 |
| 2040-2049 | 44.00 (-9.00 %) | 44.00 (-9.00 %) | 44.56 (-7.84 %) | 43.35 (-10.34 %) |
| 2090-2099 | 43.40 (-10.24 %) | 41.54 (-14.08 %) | 40.33 (-16.59 %) | 36.78 (-23.93 %) |
| 2190-2199 | 43.59 (-9.84 %) | 40.09 (-17.08 %) | 36.79 (-23.91 %) | 25.52 (-47.22 %) |
| 2290-2299 | 44.32 (-8.34 %) | 39.29 (-18.74 %) | 35.96 (-25.63 %) | 21.48 (-55.57 %) |
| 2390-2399 | 44.88 (-7.34 %) | 39.00 (-19.34 %) | 35.29 (-27.01 %) | 20.03 (-58.57 %) |
| 2490-2499 | 45.43 (-6.04 %) | 38.79 (-19.77 %) | 34.63 (-28.38 %) | 19.36 (-59.93 %) |

Table 4.1: Changes in the extent of permafrost, seasonally frozen ground and maximum annual snow covered area in the Northern Hemisphere, following the four RCP scenarios to year 2500.

The rate of decrease of PF, SFG and SCA sharply increases in the early part of the $21^{st}$ Century, coincident with a period where there is a sharp rise in both Arctic and Global surface air temperature (Figure 4.7). The exposed cryosphere responds similarly along the four RCPs over the first half of the $21^{st}$ Century, with PF, SFG and SCA decreasing by between $\sim$ 15-20 %, 7-9 % and 8-10%, respectively, by mid-Century relative to the 1980-1999 reference period. At this point, the trajectories of the RCP scenarios begin to diverge. At year 2100, PF has decreased by between 28.51 % (RCP 3-PD) and 53.82 % (RCP 8.5), SFG has decreased by between 8.32 % and 22.37 % and SCA has decreased by between 10.24 and 23.95 %.

The permafrost results presented in Figure 4.14 and Table 4.1 relate changes in the area containing permafrost at all depths. Impacts of permafrost thaw on surficial processes and characteristics are likely more closely related to degradation of permafrost near the surface which will precede the final degradation of the deepest permafrost layers. I define near surface permafrost as being permafrost that is entirely contained within the top four subsurface layers, or top 1.06 m of the ground. Figure 4.15 and Table 4.2 compare the different responses in the total and near surface permafrost area following the four RCP scenarios. The response of the two permafrost areas is very similar until the mid-$21^{st}$ Century, due to the fact that thin permafrost in the southernmost permafrost regions is the first to degrade. Most of this permafrost lies either entirely within the near surface zone or just beneath the near surface. Around the mid-$21^{st}$ Century, near surface permafrost in the thicker permafrost zone begins to vanish and the timeseries accordingly begin to diverge, with the most pronounced divergence apparent in the higher RCP pathways: for example, by year 2100 in RCP8.5, 64.9 % of near surface permafrost has vanished as opposed to 53.8 % of total permafrost. In a subsequent section, I carry out a sensitivity analysis to determine how perturbations to the land surface scheme affects permafrost degradation; where the different model configurations impact the degradation of near surface and total permafrost differently, I plot both the total and near surface permafrost response, otherwise, I simply plot the total permafrost response.

Changes in permafrost over the historical period and $21^{st}$ Century have been investigated by a number of other modelling groups and results from the literature are summarized in Table 4.3. I have restricted this table to consider groups investigating changes in pan-Arctic permafrost extent within coupled climate models. An elaborate comparison of the results from the UVic ESCM to any of these models would be non-trivial as the models vary both in terms of overall model complexity and in the

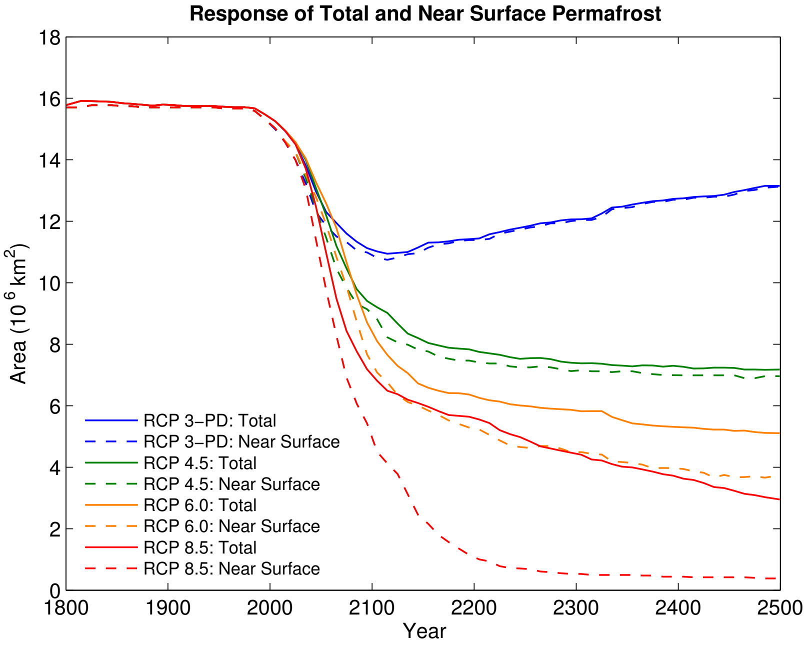


Figure 4.15: Changes in the total Northern Hemisphere areal extent of total permafrost (area containing permafrost at any depth) and near surface permafrost (area containing permafrost within top 1.06 m of ground).

| | **Permafrost Area** ($10^6$ km$^2$) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | **RCP3-PD** | | **RCP4.5** | | **RCP6.0** | | **RCP8.5** | |
| Year | Total | N. S. | Total | N. S. | Total | N. S. | Total | N. S. |
| 1800-1809 | 15.78 | 15.71 | 15.78 | 15.71 | 15.78 | 15.71 | 15.78 | 15.71 |
| 1980-1999 | 15.57 | 15.45 | 15.57 | 15.45 | 15.57 | 15.45 | 15.57 | 15.45 |
| 2040-2049 | 12.98 | 12.30 | 13.09 | 12.41 | 13.27 | 12.80 | 12.48 | 11.41 |
| 2090-2099 | 11.13 | 10.99 | 9.41 | 9.14 | 8.71 | 7.67 | 7.19 | 5.42 |
| 2190-2199 | 11.42 | 11.39 | 7.84 | 7.47 | 6.37 | 5.29 | 5.64 | 1.15 |
| 2290-2299 | 12.06 | 12.01 | 7.40 | 7.13 | 5.87 | 4.49 | 4.47 | 0.53 |
| 2390-2399 | 12.73 | 12.70 | 7.30 | 6.99 | 5.31 | 3.97 | 3.73 | 0.44 |
| 2490-2499 | 13.15 | 13.13 | 7.18 | 6.96 | 5.11 | 3.74 | 2.95 | 0.38 |

Table 4.2: Changes in the total Northern Hemisphere areal extent of total permafrost (area containing permafrost at any depth) and near surface permafrost (N.S. - area containing permafrost within top 1.06 m of ground).

sophistication of their land surface scheme (e.g. Demchenko et al. (2001) and Saito et al. (2007) both do not resolve soil freeze-thaw processes, as an example). Results may also differ significantly depending on the degree to which the model is developed to better represent high-latitude land surface processes as evidenced by the different simulated permafrost distributions of Lawrence and Slater (2005) and Lawrence et al. (2011). There is a considerable spread both in terms of the simulated late $20^{th}$ Century permafrost distribution and rate of $21^{st}$ Century permafrost degradation along comparable forcing scenarios which points to the relative infancy of representations of permafrost within coupled climate models and the need for further model development work to better constrain estimates of future permafrost degradation.

A cursory comparison of the permafrost distribution and response in the ESCM compared with the results from this table does however illustrate two points. Firstly, the present-day distribution of permafrost is simulated fairly well (using the observed total permafrost area as a metric) by the ESCM as it is in the NCAR CCSM model (Lawrence et al., 2011) and secondly, the rate of permafrost degradation in the ESCM lies in the mid range of this presented spectrum of models. The relatively slow degradation of permafrost from 1800-2000, followed by a rapid increase in the rate of degradation beginning in the early $21^{st}$ Century is a feature that is shown by other modelling groups (Koven et al., 2011; Lawrence et al., 2011), but the cumulative $21^{st}$ Century degradation in the ESCM is relatively low, resulting in a decrease of at most 49 % by the end of the Century due, at least in part, to the aforementioned weak polar amplification of the model. As well, the ESCM has a deeper ground column than any of the other models surveyed and consequently the thermal inertia of deeper soil layers would be expected to slow the rate of warming of layers closer to the surface. Finally, given that the loss of near surface permafrost tends to occur well in advance of very deep permafrost, it is important to keep in mind different groups' definitions of permafrost (and depth of their ground columns) when comparing results.

Figure 4.16 shows changes in the spatial distribution of PF, SFG (and intermittently frozen ground) and SCA in the ESCM at the end of the $21^{st}$ Century. Unsurprisingly, permafrost loss is most extensive at its southern limits, where permafrost is thinnest (Figure 3.17). The region of most extensive permafrost loss is in Central-Southern Siberia as future warming along the southern limit of permafrost tends to be greatest along these longitudes (Figure 4.4). There is also substantial degradation of permafrost in the Tibetan plateau region in all scenarios, although some permafrost persists here even in the high concentration RCP 8.5 scenario at

| **Authors** | **Ref. PF Area** | **Scenario** | **Future PF / % Change** |
|---|---|---|---|
| Lawrence and Slater (2005) | 11.9 x $10^6$ km$^2$ (1970-1989) | SRES B1, 2080-2099 | 4.6 x $10^6$ km$^2$ (-61 %) |
| | | SRES A1B, 2080-2099 | 2.7 x $10^6$ km$^2$ (-77 %) |
| | | SRES A2, 2080-2099 | 2.1 x $10^6$ km$^2$ (-82 %) |
| Lawrence et al. (2011) | 12.5 x $10^6$ km$^2$ (1970-1989) | RCP 3PD, 2080-2099 | 8.4 x $10^6$ km$^2$ (-33 %) |
| | | RCP 4.5, 2080-2099 | 6.3 x $10^6$ km$^2$ (-49 %) |
| | | RCP 6.0, 2080-2099 | 4.8 x $10^6$ km$^2$ (-62 %) |
| | | RCP 8.5, 2080-2099 | 3.5 x $10^6$ km$^2$ (-72 %) |
| Koven et al. (2011)* | ~ 13.5 x $10^6$ km$^2$ (2000) | SRES A2, 2100 | ~ 10 x $10^6$ km$^2$ (-30 %) |
| Saito et al. (2007)** | 18.1 x $10^6$ km$^2$ (1980-1999) | SRES, A1B, 2080-2099 | 7.3 x $10^6$ km$^2$ (-60 %) |
| Demchenko et al. (2001)** | 25.7-26.2 x $10^6$ km$^2$ (1961-1990) | IS92a, 2100 | 17-18 x $10^6$ km$^2$ (-34 %) |

Table 4.3: Summary of changes in permafrost over the $21^{st}$ Century (Future PF) as reported by other modelling groups. The table summarizes the late $20^{th}$ Century permafrost distribution (Ref. PF Area), the scenario for future climate change / time at which the future permafrost distribution is determined along with the future permafrost extent. *Results are for permafrost North of 60 °N. ** Models do not explicitly resolve soil freeze-thaw processes.

year 2100. Permafrost that was present in the reference period in Northern Europe (principally in the Scandinavian countries) has also completely vanished by 2100 in RCP 6.0 and 8.5.

Changes in SFG and SCA by 2100 are less dramatic than changes in permafrost as warming at the southern limits of seasonal ground freeze and snow cover is less pronounced than in permafrost regions. All regions of permafrost loss have transitioned to seasonally frozen ground at 2100. At this time, seasonally frozen ground retreats in the region between the Black and Caspian seas (replaced in part by intermittently frozen ground only in RCP-3PD), and retreats in the Southern Rockies with the rate of retreat particularly pronounced in the higher concentration pathways. I simulate an overall reduction in SFG of between 8.32 % (RCP3-PD) and 22.37 % (RCP8.5) by 2100 relative to the reference period. This agrees reasonably well with the few published projections of changes in SFG. Lawrence et al. (2011) reports a reduction in SFG by the end of the $21^{st}$ Century ranging between 4 % (RCP3-PD) and 15 % (RCP8.5), while Saito et al. (2007) reports a reduction in SFG of 15.3 % by 2099 following the SRES A1B scenario.

There is very strong agreement between the distribution of seasonally frozen ground and the maximum snow covered area both in terms of spatial distribution and overall area (eg. the nearly identical SFG and SCA time series in Figure 4.14). This result is perhaps unsurprising as both are strongly controlled by surface air temperature and respond to surface warming with a comparable timescale. Zhang

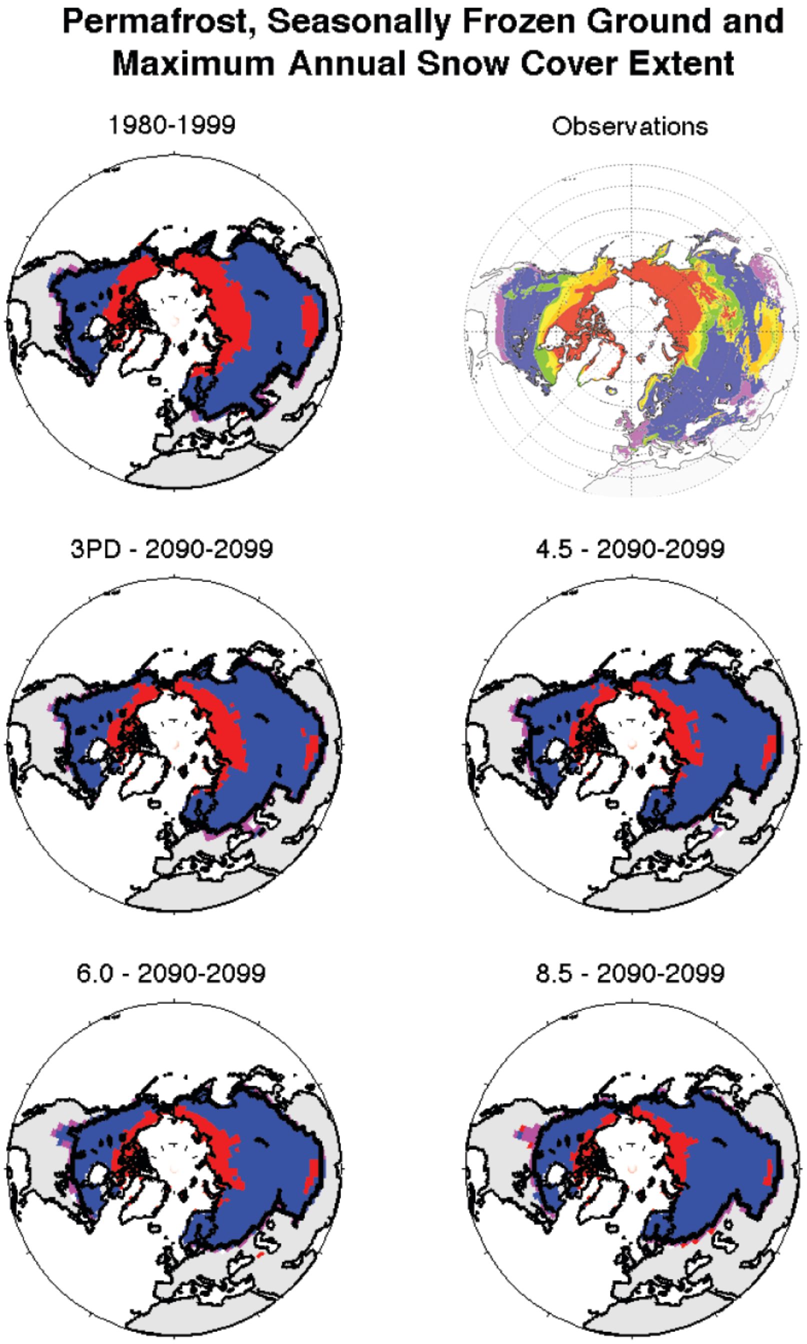


Figure 4.16: Changes in the extent of permafrost, seasonally and intermittently frozen ground and the maximum annual extent of lying snow. The top left panel shows these quantities simulated by the model over the period 1980-1999 with the maximum annual snow cover extent shown by the bold black line. The top right panel shows PF, SFG and IFG from observations over approximately the same period. The bottom four panels show the state of the land cryosphere over the period from 2090-2099 following the four RCP scenarios. The state of the cryosphere over the period 2490-2499 is shown in the plots in Figure 4.18 along with a legend.

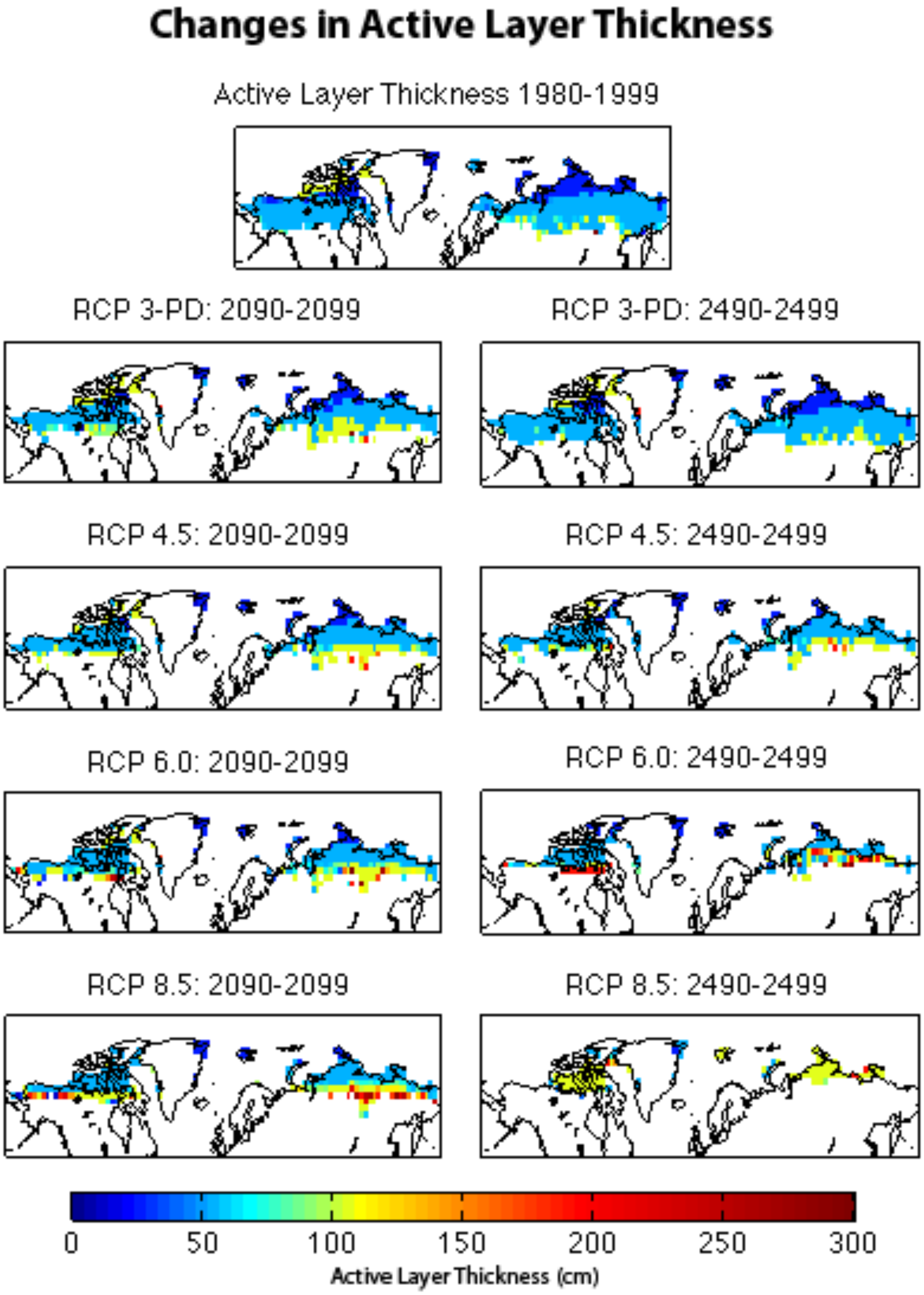


Figure 4.17: Changes in active layer thickness (ALT) following the Representative Concentration Pathways. The average ALT for 1980-1999 is displayed along with the average ALT for 2090-2099 and 2490-2499 for the four RCP scenarios. The region of the Tibetan Plateau is excluded from these plots in order to better show changes at higher latitudes.

et al. (2003), however showed that, while maximum SCA and the extent of seasonally frozen ground are closely aligned, the southern limit of the snow covered area is generally further north than the southern limit of seasonally frozen ground and SCA is generally about 11 % less extensive than the area of SFG. The fact that the two regions coincide so closely in the model may again be an artefact of the model's limited vertically resolution as a too-thick upper soil layer may prevent the ground from freezing seasonally. Indeed, in the resolution sensitivity study investigated in the previous chapter, the SFG was shown to increase by ∼6 % when model resolution was quadrupled.

Deepening of the active layer too is apparent by year 2100 (left columns, Figure 4.17), though changes in active layer thickness are generally restricted to the southern limits of permafrost; changes in active layer thickness in high latitude permafrost regions tend to be considerably less pronounced, even under the high RCP scenarios. However, as has been shown in the previous chapter, the coarse vertical resolution of the model limits the active layer thickness to certain discrete values which is likely why an expected deepening of the active layer is not shown at all sites. The active layer thickness also generally remains shallower than 300 cm, even at sites in the lower limits of permafrost with thawing occuring. This is because, by definition, the active layer must undergo seasonal freezing and thawing. So, a deepening of the active layer as a result of the annual warming front penetrating deeper into the soil must also be accompanied by a deepening of the cooling front in the winter in order for soil layers to remain within the active layer. The thermal diffusivity of the ground column may therefore limit the extent to which active layer deepening can occur.

Once the penetration depth of the summer warming front begins to extend below the depth of the winter cooling front, sustained warming of deeper layers begins to occur. Eventually, this sustained warming can lead to the formation of a *talik*, or a region of year-round unfrozen ground, within the soil column between the active layer and the permafrost table. Taliks can occur within permafrost regions when the permafrost is in equilibrium in with climate; for example when permafrost is present adjacent to a large body of water, the heat capacity of the water may prevent the ground above permafrost from fully freezing, yielding a talik. However, the presence of a talik between the active layer and the permafrost table is an indication of permafrost in a state of climatic disequilibrium. Once taliks form, they generally strongly resist refreezing associated with interannual temperature fluctuations as soils within taliks typically have a high moisture content and thus require a substantial loss of latent heat in order to refreeze (Schuur et al., 2008).

In the RCP-3PD scenario, the peak radiative forcing occurs at ~2050, and the mean surface air temperature in the Arctic begins to slowly decline in the latter half of the 21st century. Following this pathway, permafrost begins to recover in the early 22nd Century, reaching a minimum of 10.39 x $10^6$ $km^2$ in 2110 and slowly increasing thereafter. SFG responds earlier, holding nearly steady at a value of 43.93 x $10^6$ $km^2$ beginning about 2060, but not showing a continuous increase until 2172. The SCA ceases to decrease in 2075, and begins to increase in 2132. Extensive loss of permafrost and seasonally frozen ground is therefore averted in this pathway as radiative forcing begins to decrease before permafrost loss becomes too extensive.

The radiative forcing fields in RCP4.5 and 6.0 stabilize in 2150 and in 2250 for RCP 8.5 (Figure 4.1); at this point, there is a noticeable decrease in the rate of change of global (and Arctic) SATs, yet both continue to increase for the duration of the simulation (Figure 4.7) and consequently, a reduction in the extent of PF, SFG and SCA continues until the end of the simulation, although these quantities appear to have nearly stabilized at 2500 in the RCP4.5 and 6.0 scenarios.

The spatial distribution of PF, SFG and SCA at year 2500 in the four scenarios is shown in Figure 4.18. Relative to its year 2100 state, in RCP3-PD, permafrost extends slight more to the south in most places in North America and Siberia and there is also some recovery of permafrost in the vicinity of the Tibetan plateau. Shoaling of the active layer has occurred in Southern permafrost locations in this scenario too and simulated ALTs more closely resemble their distribution in 1980-1999 than in 2090-2099 (Figure 4.17), supporting the observation that dramatic changes to the future state of permafrost may be avoid if appropriate efforts to mitigate climate change are undertaken.

Continued permafrost loss occurs in each of the remaining RCP scenarios, though the amount of permafrost lost between 2100 and 2500 is less than the amount lost between 1980-1999 and 2100. Figure 4.14 indicates that, by year 2500, the exposed Northern Hemisphere cryosphere is close to a stabilized state following RCP4.5 and 6.0; if this is the case, then Figure 4.18 provide an estimate of what the cryosphere might look like under sustained +4.0 $Wm^{-2}$ and +6.0 $Wm^{-2}$ radiative forcings. In RCP4.5, a small amount of permafrost remains in the Tibetan plateau region and in Northern Europe and there has been continued northwards migration of the southern limits of permafrost, seasonally frozen ground and snowcover. This is true too of RCP6.0, though here no permafrost remains in Northern Europe nor the Tibetan

Plateau and the region experiencing seasonally frozen ground in the Southern Rockies is gone. Deepening of the active layer progresses too by year 2500 (Figure 4.17) and by this time all but the most northernly permafrost sites show deepening of the active layer, but again, the active layer is seldom deeper than 300 cm.

The Northern Hemisphere cryosphere is radically different by year 2500 under RCP8.5's sustained +12 $\mathrm{Wm}^{-2}$ radiative forcing and Figure 4.14 indicates that permafrost is not yet in equilibrium at the end of this simulation. In this scenario, only 2.95 x $10^6$ $\mathrm{km}^2$ of permafrost remains and is restricted to to the high Arctic. As shall be shown, much of this remaining permafrost is in the process of degradation and would likely be lost if the radiative forcing were to be maintained at this level beyond the end of the simulation. Snowcover and seasonally frozen ground occurs only in the Arctic and in the vicinity of the Tibetan Plateau in this dramatically warmer world.

Results from this latter simulation are also interesting as they highlight the relative vulnerability of different permafrost regions. To this end, I produced a plot (Figure 4.19) illustrating the year at which permafrost is lost both near the surface (*NS*, top 1.06 m of ground) and at all depths (*total*) following the RCP8.5 scenario. Near surface permafrost (top panel) vanishes by year 2200 nearly everywhere save the northern margins of Greenland and the high Arctic islands. As expected, there is a latitudinal gradient in the year at which permafrost vanishes. Despite temperature change becoming more pronounced at higher latitudes, temperatures are warmer near the southern limits of permafrost and as well, in this region permafrost tends to be quite thin (less than a metre in some areas, vs. 10s to 100s of metres in higher latitudes). Consequently, permafrost tends to vanish most rapidly in its southern limits.

The plot of year of degradation of total permafrost (bottom panel) shows an interesting division between the southern limits of permafrost where all permafrost largely vanishes by 2100 and regions where permafrost has not yet vanished in year 2500, with only a narrow latitudinal transition zone between the regions. This transition zone corresponds neatly with the boundary between very thin and very thick permafrost (Figure 3.17) that appears both in the model and in observations. While this result is perhaps unsurprising, the model indicates that thick permafrost will degrade at a substantially slower rate than thin permafrost and points to the importance of distinguishing between loss of near surface and deep permafrost. As well, if one is concerned about the loss of such deep permafrost, then these results also point to the importance of establishing a model with a very deep bottom boundary as it is clearly

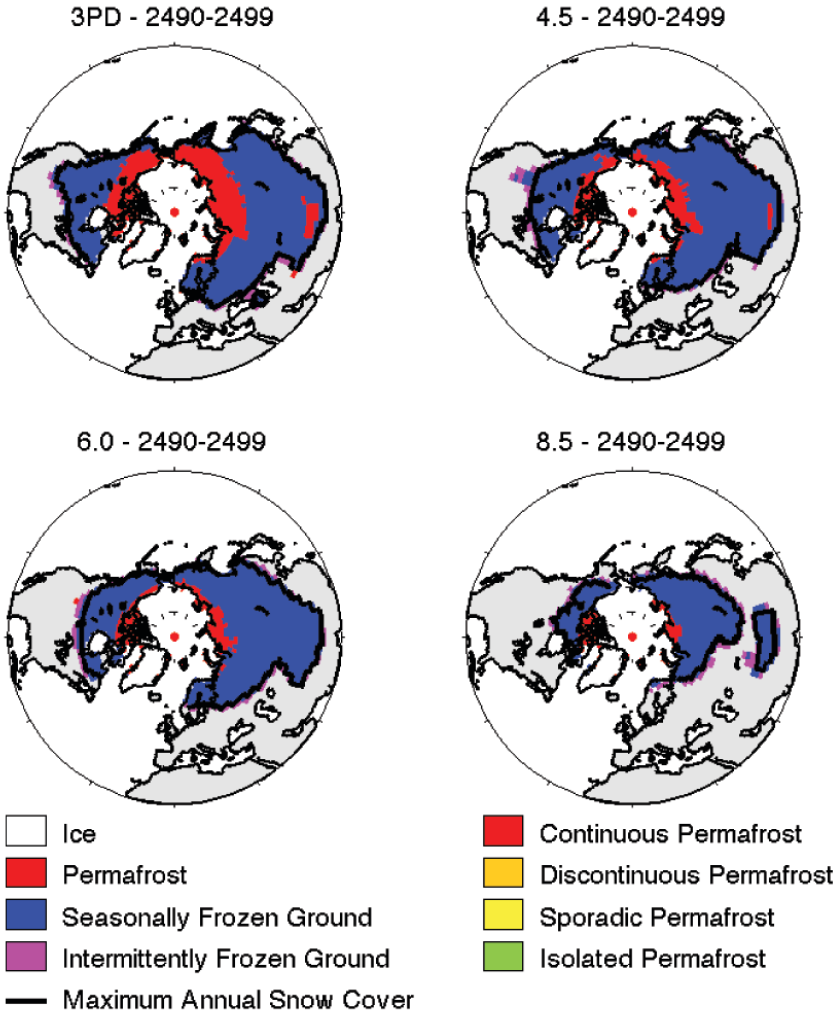


Figure 4.18: Extent of permafrost, seasonally and intermittently frozen ground and the maximum annual extent of lying snow over the period 2490-2499.

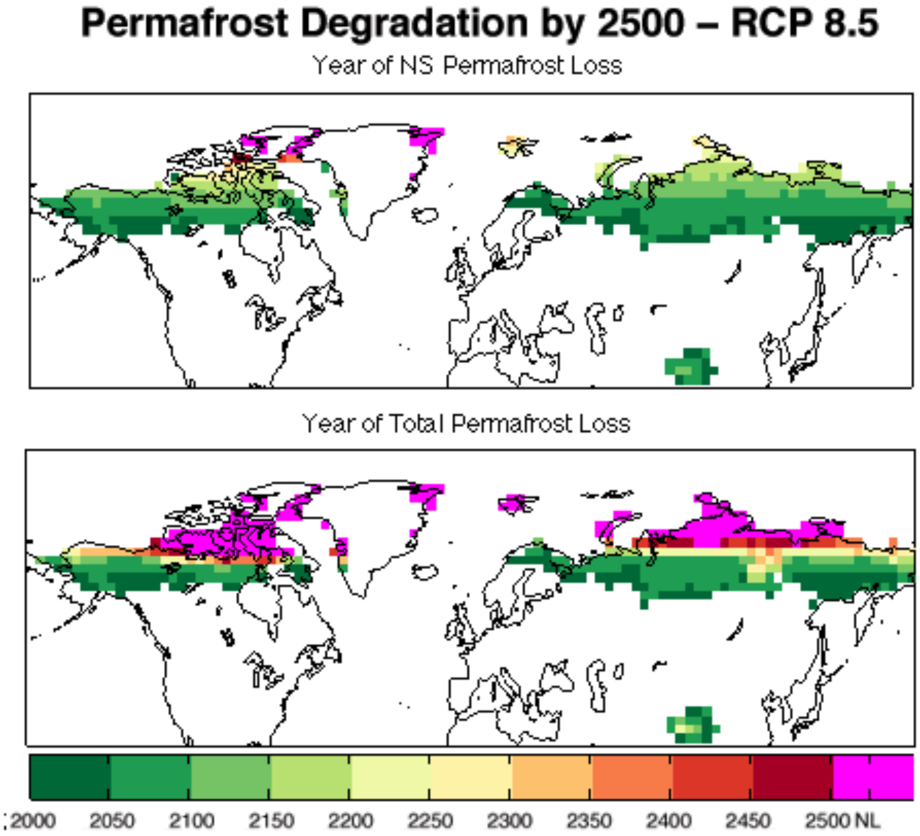


Figure 4.19: Degradation of permafrost following the RCP8.5 scenario up to year 2500. Year at which Near Surface (NS, depth $\leq$ 1.06 m) permafrost vanishes (top panel) and year at which total permafrost (at all depth) vanishes (bottom panel). NL =“No Loss”, meaning permafrost is still present (at some depth) at year 2500.

invalid to assume that such deep layer will respond anywhere nearly as quickly as the surface layers.

Permafrost degradation begins with a progressive deepening of the active layer followed by the formation of a talik. In regions of thick permafrost, one might then look to the loss of near surface permafrost or the formation of a talik as a warning sign that thaw is underway and that deeper layers of permafrost will be subsequently lost if warming progresses. Figure 4.20 (top panel) shows the time lag between the loss of near surface permafrost and the loss of total permafrost following the RCP8.5 scenario. Within the thin permafrost layers, all existent permafrost is close to the surface, so the loss of near surface permafrost coincides with the loss of total permafrost. In thicker permafrost regions, the loss of near surface permafrost preceeds the total loss of permafrost by between 50 and 350 years, thickness dependent.

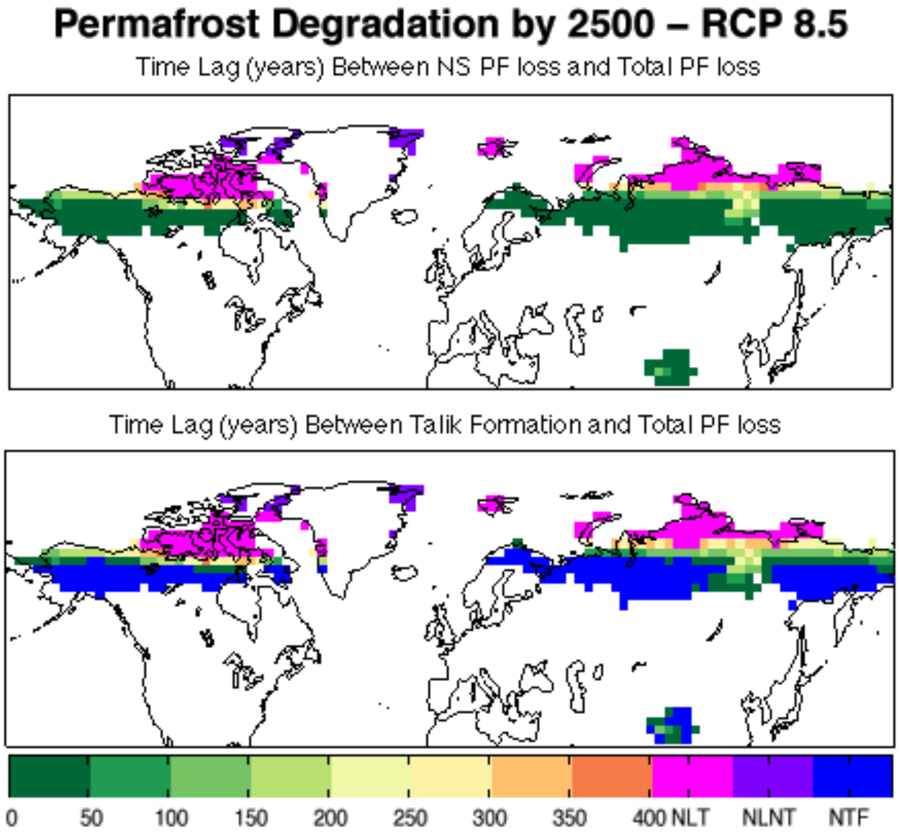


Figure 4.20: Temporal indicators of total permafrost loss, based on permafrost loss by year 2500. The top panel shows the time lag between the loss of near surface permafrost and the loss of total permafrost. The bottom panel shows the time lag between the development of a talik and the loss of total permafrost. NLT = No Loss (PF present at some depth at year 2500), but Talik Present; NLNT = No Loss, Talik not present; NTF = Permafrost degradation occurred without Talik formation.

I identify the presence of taliks within the model if any soil layers within a grid cell remain above the freezing point for two consecutive years. Figure 4.20 (bottom panel) shows the time lag between the formation of a talik and the loss of total permafrost with the time lag essentially identical to that in the top panel in the transition zone between thin and thick permafrost supporting the observation that the loss of near surface permafrost also generally coincides with the formation of taliks within the model. Lawrence et al. (2008b) also investigated talik formation in the context of permafrost loss and found that the onset of talik formation allowed for enhanced ground heat accumulation in part because the talik's ability to resist refreezing allows soil heat to accumulate at depths and also because layers below the talik tended to remain isothermal close to 0 °C, thereby reducing heat loss from the to cooler layer below. Thus we may envision the formation of taliks as a harbinger of the onset of enhanced permafrost thaw within a gridcell.

Permafrost in the thin zone generally degrades wholly through deepening of the active layer, without the formation of a talik. This however may be an artefact of the coarse vertical resolution of the model - it is possible that the model may not be able to resolve the formation of very thin talik layers that preceed the degradation of this thin permafrost. At year 2500 some permafrost remains in the high Arctic, but taliks are present nearly everywhere but the northernmost margins of Greenland and a few high Arctic islands, indicating that thaw of this remaining permafrost is underway. Consequently, it is anticipated that there will be near total loss of Northern Hemisphere permafrost if this warm state were to persist.

## 4.4 Sensitivity Experiments

In this section, I attempt to illuminate some of the uncertainties associated with permafrost degradation through a sensitivity analysis of different model configurations.

### 4.4.1 Excess Ice Experiment

In some permafrost regions, significant amounts of ice are present in volumes in excess of the soil pore spaces. This *excess ice* occurs in thin films and wedges. The thawing of ice-rich ground can lead to ground subsidence, which is one of the most tangible impacts of permafrost degradation, particularly in regions of human habitation or where structures such as pipelines may be damaged by thaw. For this reason, identification of regions which are most likely to be affected by future subsidence is of great importance.

In its standard configuration, the model is capable of simulating ground ice that can occupy at most, a volume equal to the soil pore spaces. Excess ice has a variety of forms and formation mechanisms and many of these mechanisms would be challenging to simulate within the ESCM's land surface scheme. However, the presence of excess ice may play an important role in slowing the rate of permafrost thaw on the basis of the additional latent heat required to thaw the additional ice. While ground ice takes on a variety of forms, it may be broadly grouped into two categories: *structure forming ice* which consists of ice cemented within the mineral components of the ground and *underground ice bodies* consisting of large slabs of essentially pure ice (Heginbottom, 2002).

### Model Development

To include excess ice within the ESCM, I developed a new model configuration in which each soil layer is allowed to include a segment of ice in addition to the ice present in soil pore spaces and the thickness of the layer is increased by an amount equal to the thickness of excess ice in the layer. The layer's heat capacity is augmented by an additional heat capacity for the excess ice and the thermal and hydraulic conductivity of the layer is determined from the conductivities of the soil and excess ice (with the hydraulic conductivity set to 0 for the ice), weighted by the volume of the soil and ice. Both the soil and excess ice have the same temperature; when the temperature exceeds 0 °C, it is reset to this value and the extra energy is used to melt the excess ice. Figure 4.21 demonstrates the two different excess ice configurations and how the model layers' thicknesses are augmented by the presence of this ice. The figure also demonstrates how ground subsidence is measured after excess ice has melted and the soil layer thicknesses have returned to their nominal thicknesses.

I do not simulate the formation of excess ice in the model, but rather specify its distribution as an initial condition during the model's spin up process where it is driven by constant year 1800 conditions until it equilibrates. To determine this initial condition, I used the ground ice distribution from the digital version of the International Permafrost Association (IPA) Circumarctic Map of Permafrost and Ground Ice Conditions compiled and edited by Brown et al. (1998). The map identifies regions of low ($< 10$ %), medium ( 10 - 20 %) and high ($> 20$ %) ground ice content representing a percentage by volume of visible ice in the upper 10-20 m of the ground and including segregation, injection ice, reticulate ice veins, ice crystals and ice coatings on particles (Heginbottom and Dubreuil, 1993). Thus, the ground ice content provided by the IPA map is not synonymous with excess ice as the latter two forms of ice likely include some of the pore space ice simulated within our model's standard permafrost scheme.

As it is unclear how to separate the visible pore space ice from the excess ice in these data, I have opted to interpret the IPA ground ice content as representing solely excess ice. Furthermore, I have chosen to take the upper limits for the ice percentages choosing 10 %, 20 % and 30 % as the excess ice content by volume in regions of low,

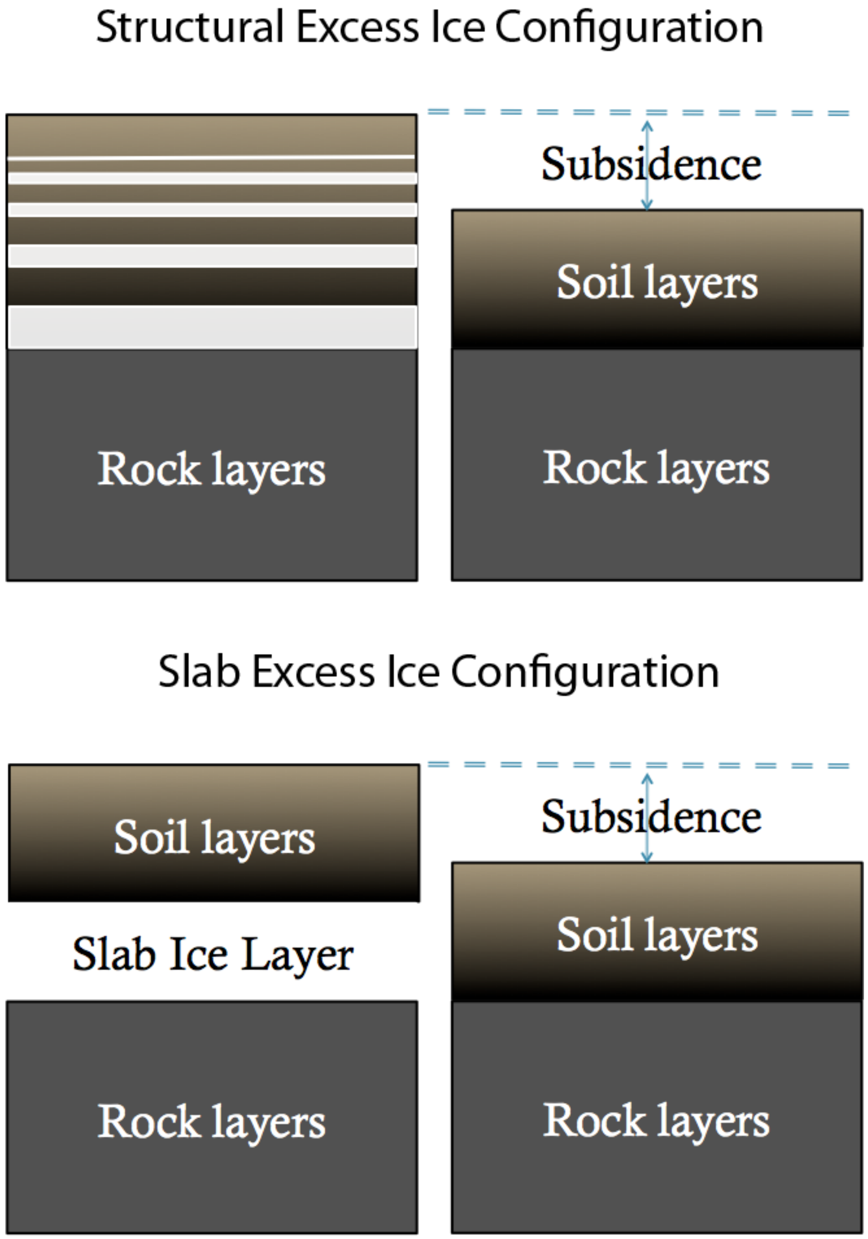


Figure 4.21: Excess ice model configurations. This figure demonstrates how the thicknesses of soil layers was increased to allow for the presence of ice in excess of soil pore spaces. Top panel: structural excess ice, representing ice that forms an integral part of the soil matrix. Bottom panel, slab excess ice in which ice takes the form of a massive, nearly pure layer of ice at the base of the bottommost soil layer. Ground subsidence is determined by the change in height of the ground surface as excess ice melts and soil layers return to their nominal thicknesses.

medium and high ground ice content, respectively. The justification for this decision is as follows: as the standard permafrost scheme, representing solely pore space ice, would tend to underestimate the volume of ground ice in permafrost regions, I have chosen to take a reasonable upper limit for the amount of ground ice with our excess ice scheme models. In this way, by comparing the transient runs of the standard and excess ice versions of the model, I will have an upper and lower limit for the rate and extent of permafrost thaw in warming scenarios. I do note however that the (non-digital) form of the IPA permafrost map identifies the presence of massive ice bodies within certain regions of the map although their presence and characteristics are poorly constrained (eg. they are listed as either present or absent). In the absence of more detailed information on the pan-Arctic distribution of these bodies, their influence must remain an unquantified uncertainty in these simulations and the lower limit to the rate of permafrost thaw could conceivably be lower still in regions where significant amounts of massive ice is present.

A second point of difficulty in specifying ground ice content as an initial condition to the model lay in determining the form(s) that the excess ice should take within a grid cell. As such, I conducted two sets of excess ice simulations with the model, one in which the excess ice is solely in the form of structural excess ice and one in which the excess ice was solely in the form of slab excess ice. The volume of excess ice within a particular grid cell was the same for both simulations. In the case of the structural excess ice, I initialized the excess ice in the lowest 5 layers of the soil scheme, corresponding to depths below 0.56 m as this was below the year 1800 active layer depth throughout most of the permafrost region. To simulate excess ice in a slab form, I assumed that the excess ice consisted of a single large mass of ice in the bottommost soil layer (starting at a depth of 5.81 m and with a nominal thickness (i.e. with no excess ice) of 4.19 m). In both schemes, as the excess ice thaws, the total thickness of the soil column decreases towards its nominal value (namely, 10 m). By tracking changes in the the thickness of the soil column, I am also able diagnose the projected ground subsidence that occurs as the excess ice melts.

The ground ice content distribution used as an initial condition in the model is shown in Figure 4.22, along with the actual ground ice content in the slab and excess ice configurations after equilibration to year-1800 conditions. The plots differ from the ground ice content from the IPA permafrost map as the map indicates the presence of excess ice in regions of sporadic and isolated permafrost where model soil temperatures are too warm for permafrost to occur. Consequently, any ground ice

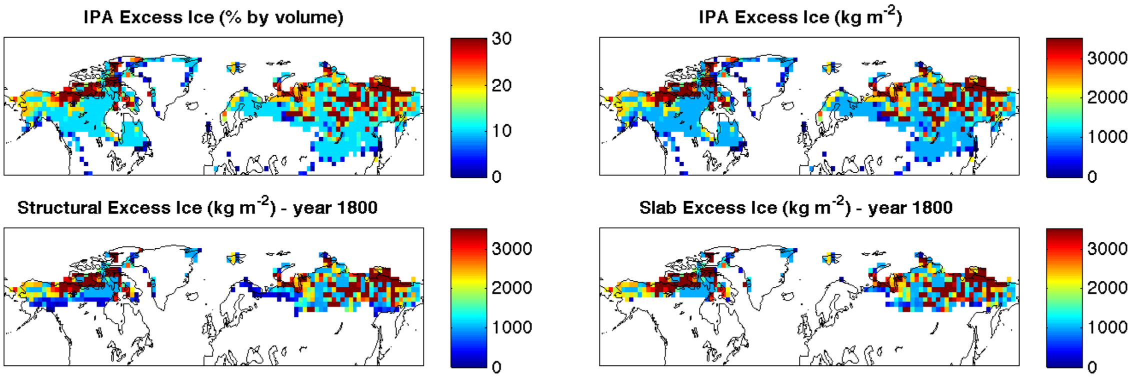


Figure 4.22: Excess ice data supplied to the UVic ESCM. The top left panel shows the excess ice expressed as a percentage of ice by volume of ground and the top right panel shows the resulting areal density of excess ice under the assumption that excess ice is restricted to the soil column (with a nominal thickness of 10 m). The bottom panels shows the total excess ice content in the model after model spinup for the structural (left panel) and slab (right panel) schemes.

that was initiated in such cells thaws during the spin-up phase of the model.

After model spinup, I calculated the Northern Hemisphere ground ice volume at year 1800 to be 25.19 x $10^3$ $\text{km}^3$ in the slab model and 26.60 x $10^3$ $\text{km}^3$ in the structural model. Zhang et al. (2000) calculates the volume of NH ground ice to range between 10.80 and 35.46 x $10^3$ $\text{km}^3$. My intention was to produce an upper limit to the extent of ground ice in permafrost in the model, so it may seem puzzling that I do not simulate ground ice content closer to Zhang et al. (2000)'s upper limit. I suggest that the discrepancy is a consequence of the fact that Zhang et al. (2000) allows for varying thicknesses of ground overburden, with regions of relatively thick overburden taken to have a thickness of 20 m, and thin overburden set at 10 m. Zhang et al. (2000) provides a breakdown of their calculations in tabular form and I repeated their calculation, taking overburden thickness to be 10 m everywhere, yielding a ground ice upper limit of 25.27 x $10^3$ $\text{km}^3$ - very close to the values for the UVic model which is not unexpected given that the ESCM simulated a total permafrost area very close to that given by Zhang et al. (2000).

### Experiment Results

I now compare the differing response of the cryosphere in the model configurations with excess ice to the standard version of the model (with only pore space ground ice). Changes in SFG and SCA were found to be nearly identical in the excess ice and standard configurations, so I consider only variations in the rate of permafrost change in these simulations. As the UVic ESCM is, at present, the only model with a pan-Arctic domain to include excess ice in simulations, my results give a useful first estimate of the extent to which large concentrations of ground ice may impact permafrost degradation.

Figure 4.23 presents a time series showing changes in the total North American permafrost extent for the two ice configurations and the standard model. All three models started with the same initial permafrost distribution and the response of the permafrost up to the beginning of the $21^{st}$ century is essentially identical. The models begin to diverge early in the $21^{st}$ Century when the loss of thin permafrost begins to occur. The three configurations are most different over the $21^{st}$ Century, which is perhaps to be expected as this is the time period when the most rapid loss of permafrost occurs in all RCP simulations (associated with the degradation of the aforementioned thin permafrost).

The divergence between the three configurations is largest in the two lower RCP pathways - 3-PD and 4.5 but the order of permafrost decay from fastest configuration to slowest are consistent across all pathways in the $21^{st}$ Century. As expected, the standard configuration of the model shows the most rapid loss of permafrost owing to the absence of additional heat required to melt the extra excess ice. The average rate of permafrost degradation over the $21^{st}$ Century is between 2.6 and 6.7 % slower in the slab configuration than in the standard configuration, but the structural configuration is considerably slower, showing between 23.3 % and 32.8 % slower rate of permafrost loss.

Recall that in the case of the structural ice model, ground ice was initiated throughout the soil column from 0.56 m to the base of the soil column while ice in the slab model was initialized everywhere at the bottommost soil layer. As a consequence of this setup, there is a greater volume of ground ice present in the structural configuration than in the slab configuration (26.60 x $10^3$ km$^3$ vs. 25.19 x $10^3$ km$^3$ at year 1800). A comparison of the equilibrium distribution of ground ice in these models (bottom panels, Figure 4.22) reveals that the difference in ice volume occurs at the southern limits of permafrost. Here, the geothermal temperature gradient is such that the temperature of the bottommost soil layer remains above the freezing point year round. As such, any ice that is initiated at the base of the permafrost would ablate during the model spinup, while some ice initiated higher up in the soil column remains once the model has equilibrated. Consequently, as there is less ground ice

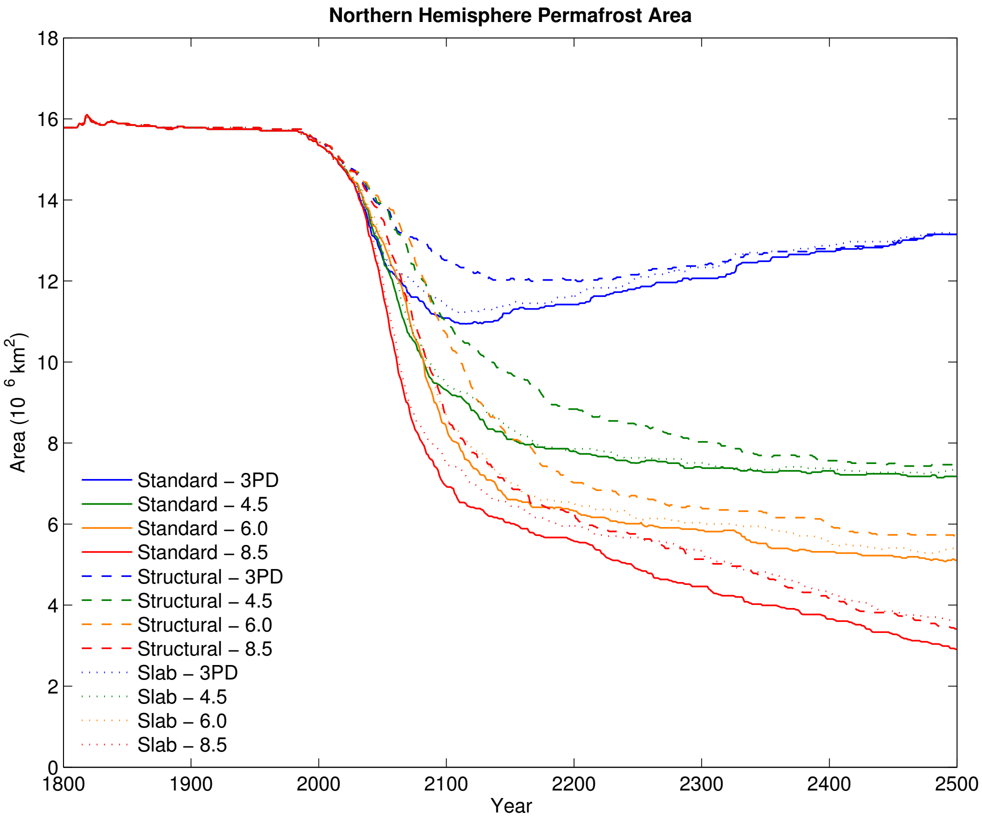


Figure 4.23: Time series showing change in Northern Hemisphere permafrost extent including varying forms of excess ground ice following four RCP concentration pathways. Solid line: standard configuration of the model (no excess ground ice), dashed line: structural configuration of the model with excess ice distributed throughout the soil column, dotted line: slab configuration of the model with excess ice present as a single slab of ice at the base of the soil column.

present in the southernmost permafrost regions (which are the first to thaw), this leads to a faster rate of PF thaw in the slab configuration. As has been discussed, permafrost thaw is initiated from above, first through deepening of the active layer and then through talik formation. As such, even if the same total amount of ground ice were present everywhere in the slab and structural models one might expect the rate of permafrost thaw to be higher in the slab model as the structural model would provide the greater resistance to active layer deepening and talik formation.

Interestingly, in the one simulation where there is partial recovery of permafrost (RCP3-PD), the slab excess ice configuration shows a faster rate of recovery than the structural configuration. While permafrost degradation occurs from above, permafrost aggradation occurs from below. In cells where a significant amount of thick, deep ice (slab configuration) recently degraded, deep soil temperatures are closer to the freezing point than equivalent cells in the structural configuration. Consequently, recovery of permafrost occurs at a faster rate in the slab model such that the slab permafrost area is actually slightly higher than the structural permafrost area by year 2500.

The degree of divergence between the permafrost extent in three model configurations generally decreases in RCP4.5, 6.0 and 8.5 after 2100. In the case of RCP8.5 at about 2250, the rate of decay of permafrost in the structural ice simulation overtakes the slab simulation. This is because, at this point, it is the thick permafrost that is in the process of thawing. Near surface permafrost in these regions (and layers of excess ice) have nearly completely degraded at this point and consequently, the greater amount of ice present in the deepest soil layer of the slab configuration allows this permafrost to be slightly more resilient than the structural excess ice configuration.

I draw two key observations from Figure 4.23. Firstly, the vertical distribution and not simply the volume of ground ice is an important determinant of the rate of permafrost change and more detailed cartographic information on the volume, form and distribution of ground ice in permafrost regions would therefore be useful to improve projections of permafrost thaw. Secondly, the presence of ground ice, even in the structural form does not radically change the rate of permafrost thaw, by which I mean the rate of thaw including and disregarding ground ice tends to be of the same order of magnitude. Accordingly, models that do not include a representation of excess ice may yet produce reasonable estimates of permafrost degradation, all other factors being equal.

The fact that the presence of excess ice (at least if the order of magnitude of the amount ground ice represented in the model is correct) does not radically change the rate of permafrost degradation may be understood from energetic principles. Focussing again on the period of most rapid permafrost loss, I examined the regions where permafrost completely vanishes by year 2100 in all four RCP simulations. I determined the amount of heat absorbed by the ground in regions where permafrost is present at year 2000 and vanishes by 2100 to range between 6.51 x $10^{17}$ J (RCP3-PD) and 1.38 x $10^{18}$ J (RCP8.5) with the larger RCP8.5 value due to both a larger region that experiences thaw and a higher degree of warming within this region. I then calculated the amount of energy that was required to melt all of the (structural) excess ice present in these regions, including the energy to raise the ice to the melting point and the energy of fusion; this was determined to range between 2.02 x $10^{16}$ J (RCP3-PD) and 4.30 x $10^{16}$ J (RCP8.5). Thus, in all cases the energy involved in melting excess ice is at least an order of magnitude lower than the total amount of energy absorbed by the ground in all cases. This both neatly explains why the presence of excess ice shouldn't radically slow the rate of thaw and why the discrepancy between the rate of structural thaw and standard thaw is greatest in the case of the RCP 3-PD simulation. Here, the amount of energy required to melt the excess ice relative to the amount of energy absorbed by the ground is greater than in any of the scenarios. This calculation also provides a simple rule of thumb for assessing the importance of ground ice in permafrost thaw simulations: the higher the degree of ground warming relative to the amount of heat needed to melt excess ice that may be present in a region, the lower the importance of excess ice in restricting thaw.

### Ground Subsidence

The explicit inclusion of excess ice within the LSS allows for a direct estimate of the potential impact of a hazard associated with permafrost thaw: ground subsidence following the degradation of excess ice. To obtain a measurement of subsidence, I assume that the ground immediately settles whenever any excess ice is lost, with liquid water resulting from ice melt added as a downwards water flux to deeper ground layers. Thus, the degree of ground subsidence that occurs as permafrost degrades is equal to the thickness of excess ice that is lost by that time relative to the starting state. Ground subsidence is known to be a highly heterogeneous process and results in the formation of uneven *thermokarst terrain* owing to strong spatial variations in ground ice content and the degree of ground warming. As the ground ice content (and ground warming) is uniform in a given gridcell, there is no representation of the formation of thermokarst in the model; I therefore suggest that the model simulates what may be thought of as large-scale average subsidence in a region. As the model does not simulate the formation of excess ice, subsidence is a monotonic process - there is no "recovery" once the ground has subsided even in simulations where there is overall recovery of permafrost as in the RCP3-PD scenario. Finally, I note that the maximum degree of possible subsidence in a given region is set by the equilibrium excess ground ice content in that region (as shown in Figure 4.22) and related to the ground ice content through the fixed density of ice (917 kg $m^{-3}$).

Figure 4.24 shows projections of ground subsidence over the next century and next 500 years for the two extremes of the RCP pathways both in the case of structural and slab excess ice. The plots clearly suggest that it is the southern limits of permafrost that are most vulnerable to permafrost subsidence in the immediate future with the degree of subsidence dependent on the position of the ground ice. The structural ice simulation, with ice distributed throughout the soil column and thus with more ice experiencing thaw in a given scenario than the slab simulation shows enhanced subsidence over the coming century in both RCPs. Ground subsidence in the case of the structural ice is generally less than 1 m by 2100 in the case of RCP 3-PD and may be up to ∼2.5 m in some locations in the case of RCP 8.5. Although year 2100 marks the approximate time at which the total permafrost areal extent ceases to change and begins to recover in the 3-PD scenario a comparison of the year 2500 plots following this scenario with the year 2100 plots indicates that substantial subsidence can yet occur after 2100 in the vulnerable southern permafrost region. This again points to the importance of extending permafrost simulation beyond merely the $21^{st}$ Century, especially if one is concerned with longterm planning for settlements and infrastructure in permafrost regions.

By year 2500 the RCP8.5 subsidence maps for the structural and slab ice resemble each other strongly with the only anomalies being the southern limits of permafrost (where there was less slab ice) and the high Arctic where permafrost near the top of the soil column has already thawed and only the deep permafrost layers remain. Consequently, more total column excess ice has thawed in the structural scheme than in the slab scheme in the high Arctic and the degree of subsidence is greater. Outside of the high Arctic, the spatial patterns of subsidence are the same as patterns in ground ice (bottom panels, Figure 4.22) as all of the ground ice has melted. This results in subsidence as high as 4.00 m in regions and the regions showing the most radical degree of change then coincide with those having the greatest extent of ground ice. Gridcell mean subsidence values of up to 4.00 m may seem extreme, but an examination of the literature indicates that this is not the case. For instance Linell (1973) reports on a controlled thaw experiment near Fairbanks, Alaska that resulting in up to 6.7 m of subsidence over a period of 26 years, while regions of Siberia reveal the presence of thermokarst depressions 20-40 m deep and covering areas in excess of 25 km$^2$ (Nelson et al., 2002).

The subsidence risk from thawing permafrost over the $21^{st}$ Century has been previously considered by Nelson et al. (2001, 2002). The authors defined a *settlement*

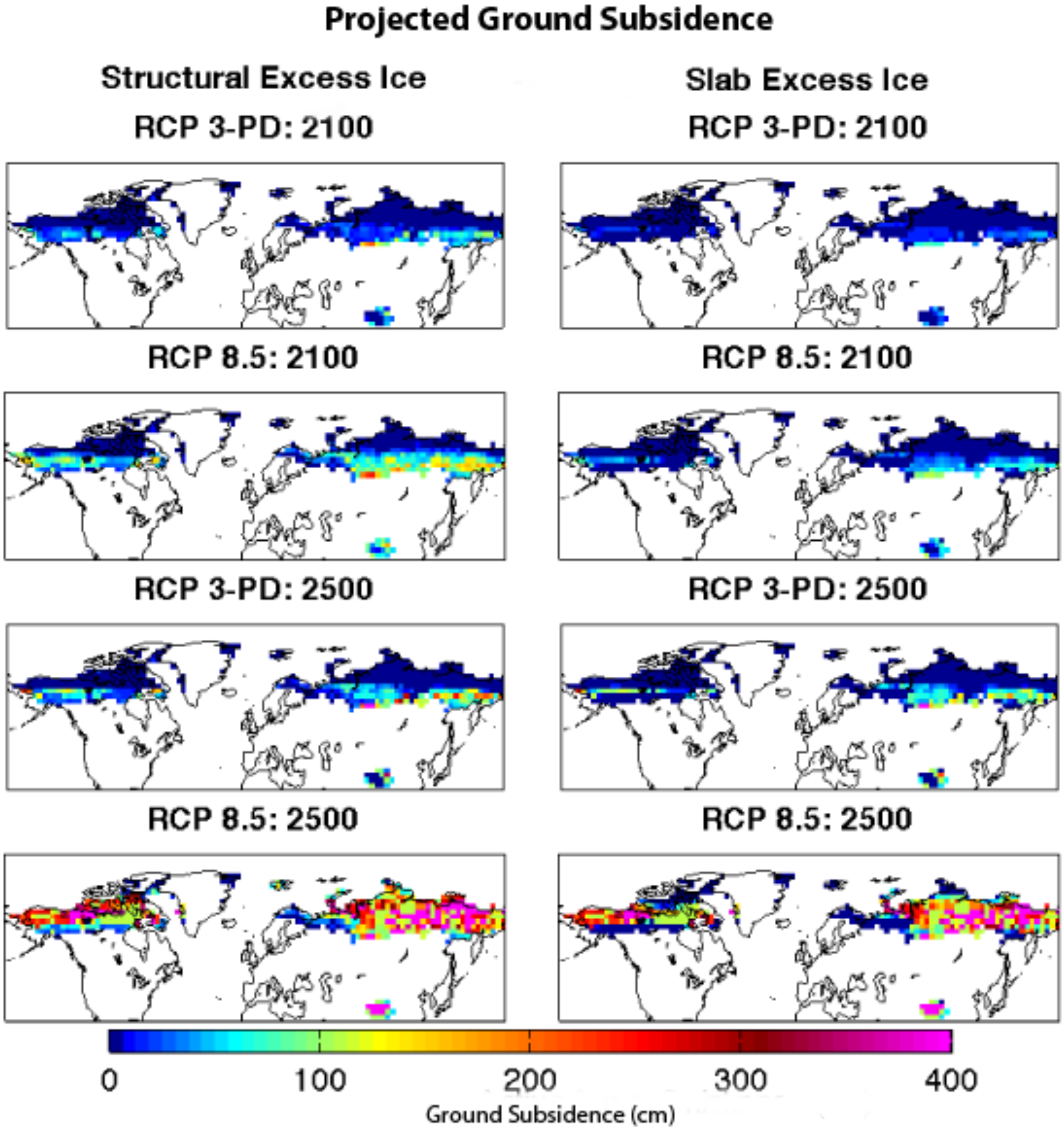


Figure 4.24: Projections of ground subsidence relative to year 2000 at 2100 and 2500. Subsidence projections are shown for the low-warming (3-PD) and high-warming (8.5) end members of the RCP simulations in the case of structural and slab excess ice.

*index* ($I_s$) to assess subsidence risk defined as the product of projected change in active layer thickness ($\Delta z_{al}$) and the volume of ground ice present ($V_{ice}$) in a region.

$$I_s = \Delta z_{al} \cdot V_{ice} \tag{4.1}$$

Changes in active layer thickness were determined using a numerical routine that relates active layer thickness to air temperature, snow cover and vegetation and snow characteristics (Anisomov et al., 1997). The authors calculated the change in active layer thickness using climatic data from the ECHAM1-A and UKTR (UKMO) climate models runs. The ground ice volume data is derived from the IPA Permafrost map, as in my work. The climate model output used in their study come from a much older generation of climate models and the models were forced following different carbon dioxide concentration scenarios. The model output used by Nelson et al. (2002) appears to be a harmonized version of the output from the original models with the harmonization performed by combining spatial predictions of change from the three dimensional models with a simple one-dimensional climate model following the IPCC IS92a emissions scenario[2]. I therefore think that this scenario may be thought of as akin to the RCP 8.5 scenario in terms of radiative forcing. The authors considered the climatic state from the two models in the mid-21$^{st}$ Century, and calculated the settlement index using this state for the models. The authors partitioned the logarithm of the thaw index into three thaw hazard categories using a "nested means" procedure (Nelson et al., 2002).

The hazard maps produced by the authors are displayed alongside my own subsidence maps for the RCP 3-PD and 8.5 scenarios in Figure 4.25. The two types of maps are not directly comparable for a number of reasons. For instance, Nelson et al. (2002) simulate a much more extensive permafrost region (as they capture intermittent and sporadic permafrost), not to mention that it is difficult to know how to directly compare hazard zonation against simulated ground subsidence. Nonetheless, I suggest that one may compare regions of relatively low / high hazard zonation and subsidence in the models and elucidate some reasons for the differences. The most striking difference between Nelson et al. (2002)'s maps and my own lies in the regions where each simulate the highest degree of potential subsidence. Nelson et al. (2002) tend to show the highest degree of subsidence in regions of relatively high ground ice content. My own results indicate that eventual subsidence (at year 2500, Figure

---

[2] The exact details of the harmonization elude me as I could not locate the original paper (Greco et al., 1994) which describes the harmonization process

4.24) may be greatest there as well if a high RCP is followed, but the spatial pattern of projected $21^{st}$ Century warming clearly shows that subsidence is expected to be greatest in the southern permafrost limits with localized subsidence enhancements in regions of high ground ice content within this band. The area where there is uniformly the strongest agreement between the two types of maps lies in the permafrost surrounding the Tibetan plateau where both models show relatively little subsidence in the central plateau, yet increasing (and potentially high degrees of subsidence) in the margins of the plateau.

There are a number of possible reasons for the differences between the hazard maps. Firstly, it should be recalled that Nelson et al. (2002)'s work was conducted using a much earlier generation of climate simulations and they do not directly simulate the ground climate as part of a coupled simulation within these models, although they do obtain a reasonable map of the present-day permafrost distribution from their work. Secondly, they use an empirical equation which relates subsidence to the product of ground ice content and changes in the active layer thickness. I suggest that this may be an overly simplistic way to assess subsidence. Presumably, subsidence should only occur if excess ground ice lies in the region that is thawed as the active layer deepens; I have already demonstrated that the position of ground ice within the soil column can have a strong impact on the degree of subsidence and these simulation differences again point to the need for more detailed knowledge as to the vertical structure of ground ice in the Arctic. A third reason for the differences has to do with the response of the active layer, which is a key determinant of Nelson et al. (2002)'s thaw index. The authors show the strongest increase in active layer thickness in the high Arctic and in the Southern limits of isolated and sporadic permafrost which is quite different from the regions where the UVic ESCM simulates the strongest increase in ALT. This points to either a poor representation of the transient response of the active layer in the ESCM or a limitation in using changes in air temperature and precipitation to track future changes in active layer thickness[3].

As there has been so little work done in assessing the hazards associated with thaw subsidence in permafrost, I do not feel that I can state that either the work of Nelson et al. (2002) or my own work is necessarily more likely to be correct. What the simulation discrepancies do point towards is the need for more such simulations

---

[3] Certainly the UVic ESCM shows strong changes in air temperature in the high Arctic but this does not appear to translate into strong changes in ALT, at least in the standard resolution configuration of the model

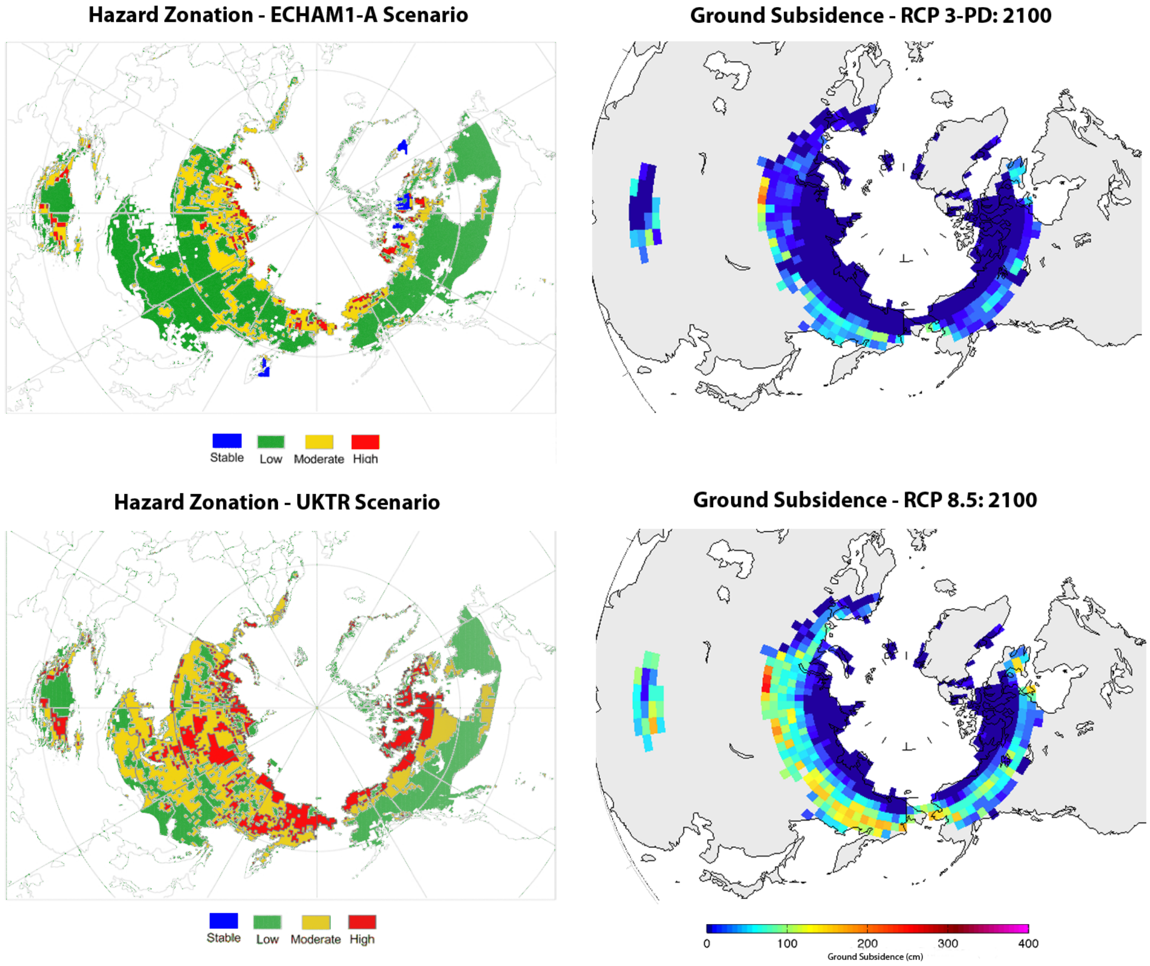


Figure 4.25: A comparison of the hazards associated with thawing permafrost. Left columns show hazard zonation maps for ECHAM1-A and UKTR climate simulations around 2050, which are related to the logarithm of the settlement index described in the text; coloration to the plots added by the author. Right columns are ground subsidence maps at year 2100 for the RCP 3-PD and RCP 8.5 scenarios with structural excess ice - the same maps as in Figure 4.24 but with a different projection.

by additional modelling groups. As coupled permafrost schemes are now actively being used by a number of groups and the inclusion of excess ice may be performed in a fairly straightforward manner as I have outlined, such simulations ought to be fairly easy to perform and may potentially be of great interest for those in northern communities.

### 4.4.2 Vegetation Dynamics Experiment

In the second sensitivity experiment, I examine the influence of vegetation dynamics on permafrost thaw. Here, I compare the standard configuration of the model, which includes vegetation dynamics (*dynamic configuration*) with a second configuration of the model with the vegetation distribution and characteristics fixed (*non-dynamic configuration*). Vegetation changes in the dynamic vegetation configuration of the model include changes in the vegetation gridcell fractional coverage, height and leaf area index; changes in these characteristics are determined via the TRIFFID dynamic vegetation component. The non-dynamic vegetation configuration of the model was setup using the methodology as was used for the vegetation in the configuration of the model forced by ECMWF data as described in detail in the previous chapter. To briefly recapitulate, monthly mean fields of vegetation fractional cover, height and leaf area index were derived from the standard configuration of the model in its year-1800 equilibrium state. These were supplied as data fields to the model. Vegetation retains a limited degree of interactivity in the non-dynamic configuration through vegetation control on evapotranspiration.

Representative vegetation changes in the RCP scenarios were shown in Figure 4.13. Very broadly, as warming occurs, latitudinal bands of dominant plant functional types in permafrost regions tend to shift northward. Grasses expand into very high latitude bare soil regions, shrubs expand into high latitude regions occupied by grasses and trees expand northwards into regions occupied in the present climate by shrubs. There is also loss of trees in certain regions along the southern limit of the present-day boreal forest; here, the trees are replaced by grasses and shrubs. As Figure 4.13 also indicates, the extremity of vegetation change varies considerably depending on the RCP scenario.

Figure 4.26 compares changes in the total Northern Hemisphere permafrost extent following the RCP scenarios in the dynamic vegetation (solid line) and non-dynamic (dashed line) configurations of the model. The dynamic and non-dynamic timeseries

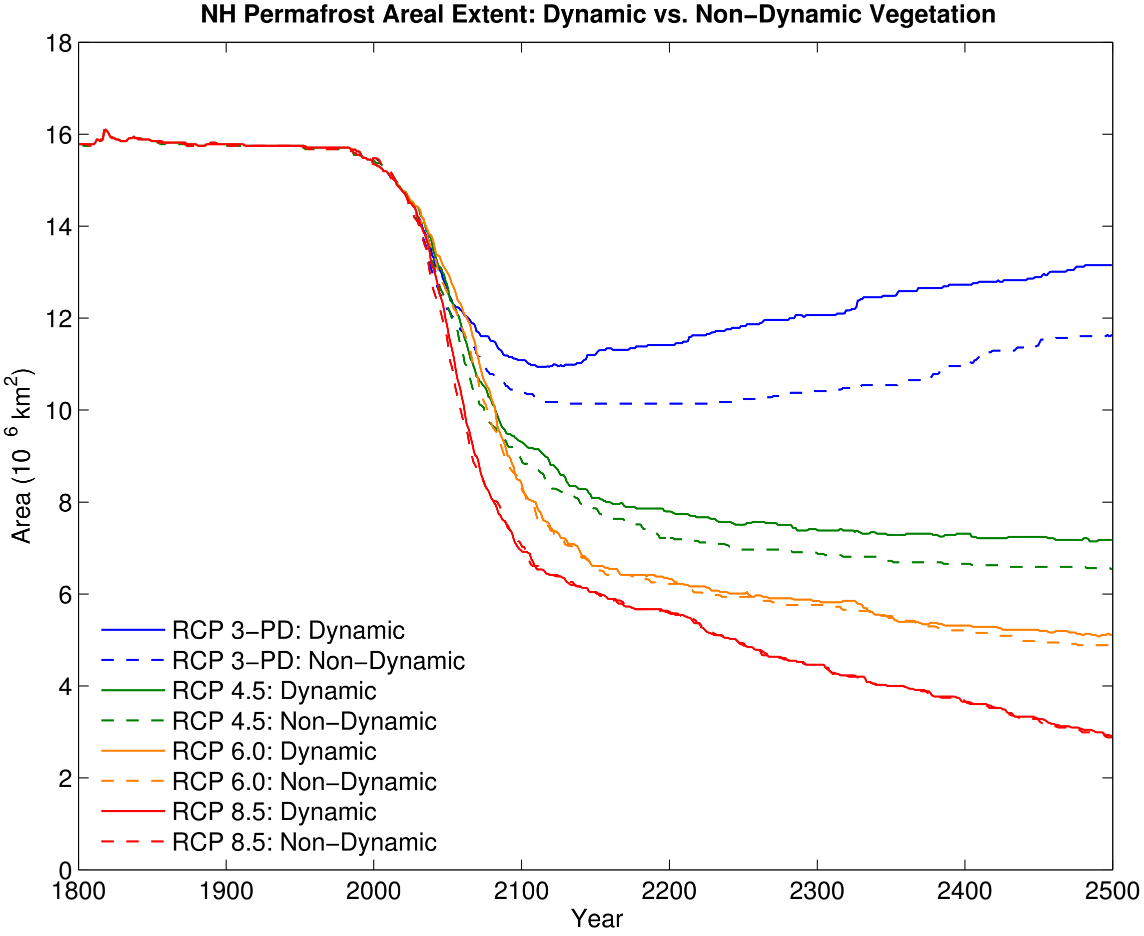


Figure 4.26: Rates of permafrost degradation in the UVic ESCM following the four RCP scenarios with vegetation dynamics enabled (*Dynamic* - solid line) and disabled (*Non-Dynamic* - dashed line)

are essentially identical from the starting point to the mid-21$^{st}$ Century when the RCP scenarios begin to diverge. The non-dynamic vegetation configuration of the model generally shows a faster overall rate of permafrost degradation than the dynamic configuration. Interestingly, although the degree of vegetation change tends to increase from RCP3-PD through to RCP8.5, the difference between the dynamic and non-dynamic vegetation configurations decreases from RCP3-PD to RCP8.5.

In order to better understand the differences between the dynamic and non-dynamic configurations, I focus on RCP3-PD which shows the greatest degree of disparity between the two configurations. The overall permafrost area decreases most rapidly and is slowest to recover in the non-dynamic configuration. However, when spatial patterns of permafrost loss are plotted, permafrost does not uniformly de-

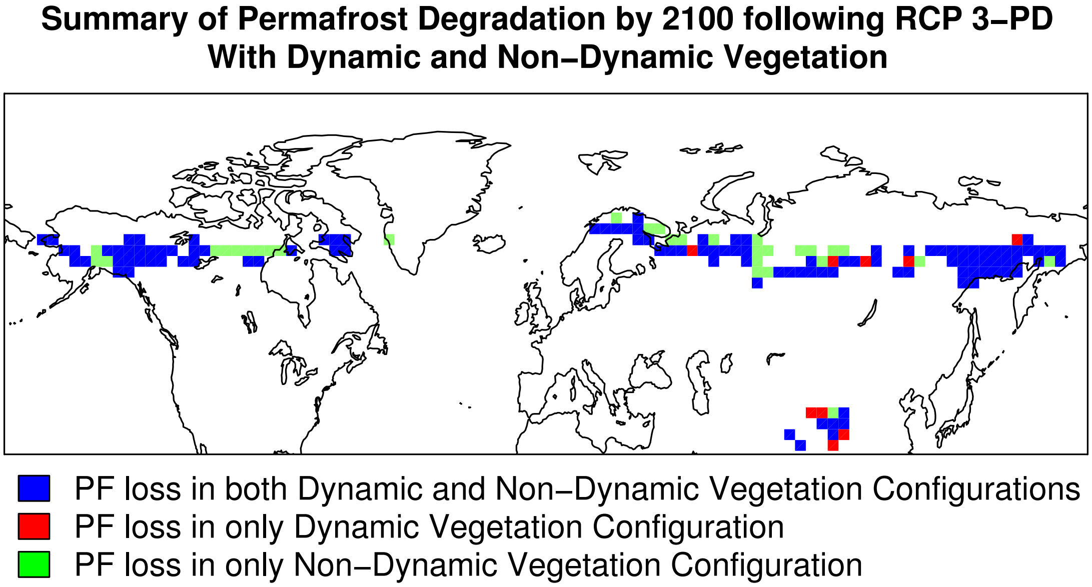


Figure 4.27: Areas of permafrost loss with and without dynamic vegetation. This plot summarizes the differences in permafrost degradation at year 2100 in the RCP 3-PD scenario which showed the most dramatic differences between the dynamic and non-dynamic vegetation.

cay faster everywhere in the non-dynamic configuration. Rather, in certain regions, permafrost is lost only in the dynamic configuration and in others, it is lost only in the non-dynamic configuration with the total areal extent of the latter regions being greater. This is illustrated in Figure 4.27 which summarizes the loss of permafrost by year 2100 (the approximate time of minimum RCP3-PD permafrost extent) in this scenario. Areas plotted in blue represent cells where permafrost is lost in both the dynamic and the non-dynamic configuration, areas plotted in red are lost by 2100 in only the dynamic configuration and areas plotted in green are areas in which permafrost is lost in only the non-dynamic configuration. To identify the role of vegetation dynamics in resisting or enhancing permafrost thaw, I isolated the cells where permafrost thaw only occurs in the dynamic configuration (green cells in Figure 4.27) and cells where permafrost occurs only in the dynamic configuration (red cells in Figure 4.27) and plotted the vegetation areal cover change in only those cells; these changes are displayed in Figure 4.28.

The left hand column of Figure 4.28 correspond to cells where permafrost is lost by 2100 in the non-dynamic vegetation configuration but not the dynamic configuration,

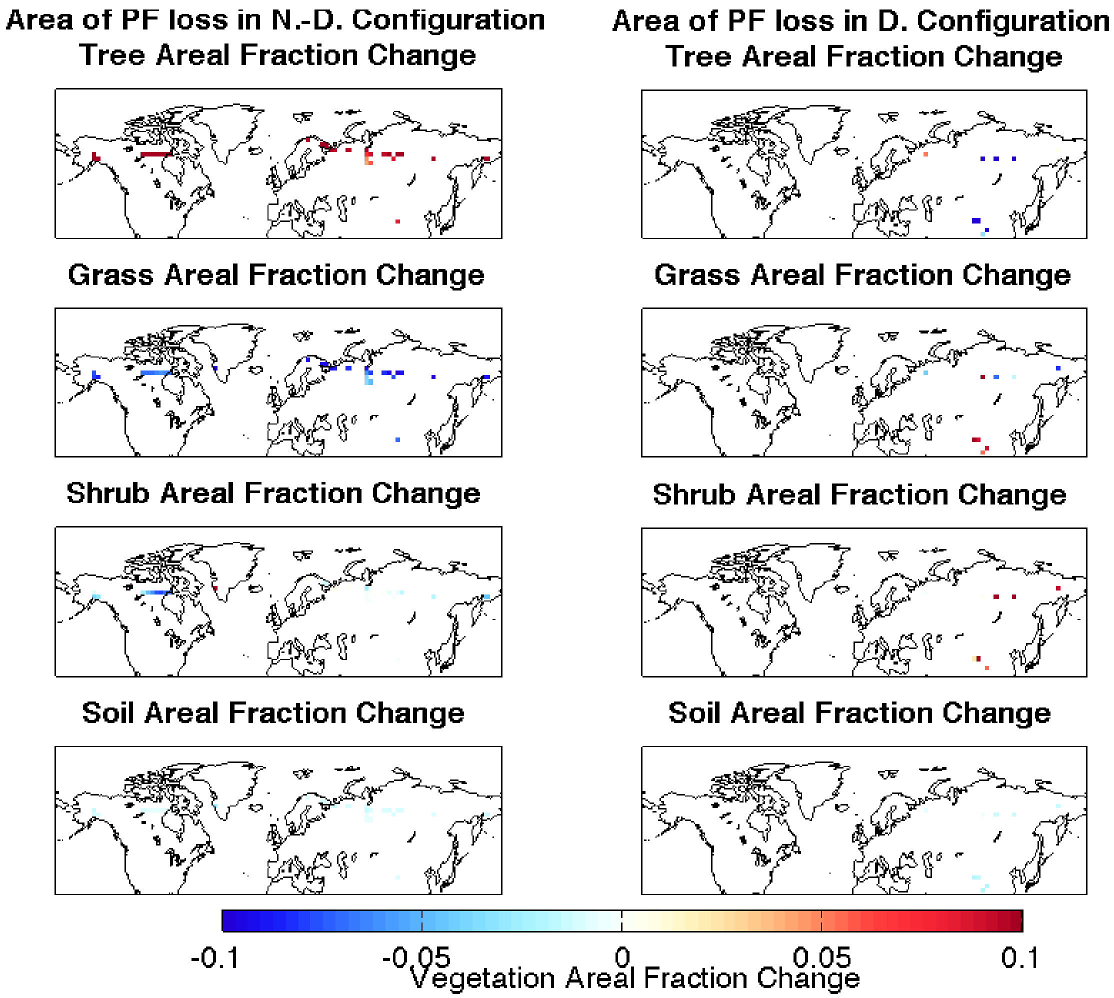


Figure 4.28: Changes in vegetation areal fractions in the dynamic vegetation configuration. Left column: vegetation changes in regions where permafrost is lost in the non-dynamic configuration (but not the dynamic configuration). This column corresponds to cells shaded in green in Figure 4.27. Right column: vegetation changes in regions where permafrost is lost in the dynamic configuration (but not the non-dynamic configuration). This column corresponds to cells shaded in red in Figure 4.27.

indicating that vegetation changes that occur in standard configuration of the model allow permafrost to be preserved in these cells (at least in RCP3-PD). These cells correspond to regions where there is a relatively strong increase in the fractional cover of needleleaf trees at the expense of shrub areal cover (and to a lesser extent, cover of grass and bare soil). The right hand column of Figure 4.28 represents cells where permafrost is lost by 2100 in only the dynamic configuration, indicating that vegetation changes that occur in the standard configuration tend to lead to enhanced loss of permafrost. These cells tend to experience fairly strong loss of tree cover (with one exceptional cell) and a reduction of bare soil areal cover and the lost cover is replaced in some regions by grasses, and in others, shrubs.

These results are consistent with the findings of the vegetation-cover sensitivity experiments discussed in the previous chapter. In that chapter, the vegetation cover was varied in the offline configuration of the model driven by ECMWF data. I investigated the permafrost distribution in model configurations where the land surface was completely covered by needleleaf trees, $C_3$ grasses, shrubs and bare soil and interpreted the results in terms of changes in surface albedo and surface roughness. I found that the annual mean soil temperatures tended to be cooler (and permafrost area more extensive) in the all tree configuration compared with the standard vegetation configuration (with vegetation cover as in the non-dynamic configuration of this section). Trees tend to have lower albedo than other vegetation types which should have a warming influence on the soil. However, this warming influence was found to be counteracted by the impact of the greater aerodynamic roughness of trees (as compared with other terrain). Typically, the land skin temperature is warmer than the air temperature and is most strongly warmer than the air temperature in the summer months. Consequently, when non-treed land surfaces are replaced with trees, turbulent heat fluxes tend to increase owing to the greater aerodynamic roughness, causing the skin (and soil) temperature to become closer to the air temperature. As the disparity between the skin and air temperature tends to be greatest in the summer, the net result is a cooling of the soil.

Consequently, in areas where other vegetation is replaced by trees (left column, Figure 4.28), one would expect the rate of permafrost loss to be slower due to the cooling influence of the trees on the soil. Conversely, where trees are replaced by other vegetation types (right column, Figure 4.28), permafrost loss may occur more rapidly owing to warming of the soil as trees are lost. As the areal extent of trees generally increases overall in permafrost regions (though certainly not everywhere),

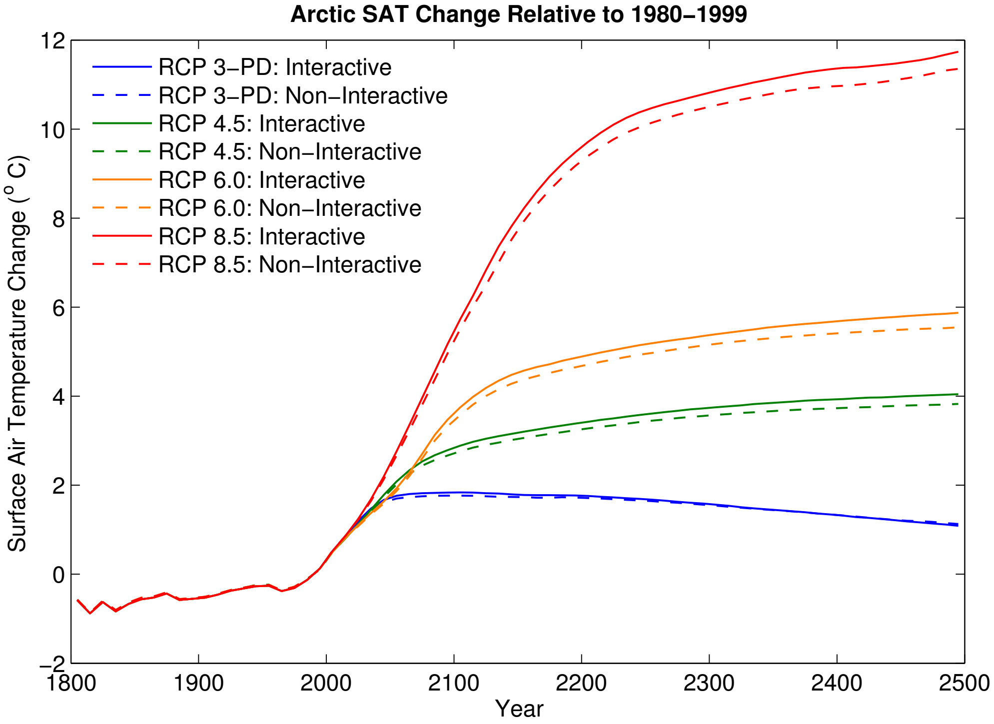


Figure 4.29: Changes in average Arctic surface air temperature (relative to 1980-1999) in dynamic and non-dynamic vegetation configurations of the model.

I suggest that this explains why the non-dynamic vegetation simulations show faster loss of permafrost as compared with the dynamic vegetation simulations.

It should be emphasized that the above analysis describes changes in skin / soil temperature, not air temperature. As revealed in Figure 4.29, the average surface air temperature in the Arctic increases more rapidly under the dynamic vegetation simulation than under the non-dynamic vegetation simulation as a result of the general decrease in surface albedo (and enhanced cooling of the surface/soil through turbulent heat fluxes) associated with the shift towards taller vegetation.

I am as yet, not perfectly certain why the disparity between the dynamic and non-dynamic simulations is greatest in the case of RCP3-PD and least in the case of RCP8.5 which respectively show the least and the greatest vegetation changes. One possible explanation may be that the vegetation-induced soil heating/cooling becomes progressively smaller relative to the soil heating resulting from surface air temperature change as one progresses from RCP3-PD to RCP8.5. Another possibility is that rapid vegetation change may only occur once permafrost thaw is underway (if, for instance, increased availability of moisture or strong surface air temperature warming is a prerequisite for vegetation shifts) and this may limit the ability of vegetation to influence thaw, especially if thaw is already fairly advance once vegetation shifts happen.

### 4.4.3 Geothermal Heat

The third sensitivity experiment investigates the importance of the inclusion of the geothermal heat flux as a bottom boundary condition in land surface schemes. The geothermal heat flux is relatively small compared with surface energy fluxes, typically on the order of 100 $mWm^{-2}$ or less (Figure 2.4) and, consequently, is often neglected in permafrost models. To perform this experiment, I integrated a second version of U3.0M to equilibrium with the geothermal heat flux set to zero so that the bottom boundary of the deepest bedrock layer has a no flux boundary condition. The year 1800 equilibrium Northern Hemisphere Permafrost area was nearly identical in both configurations, with the standard (with geothermal heat flux - with GT) configuration having an area of 15.78 x $10^6$ $km^2$ and the no geothermal heat flux configuration (no GT) having an area of 16.12 x $10^6$ $km^2$. The equilibrium seasonally frozen ground area and maximum snow covered area were identical in the two configurations and their transient responses were virtually the same and, as such, these quantities are not discussed further.

Figure 4.30 shows the difference in annual mean equilibrium temperature between the GT and no GT configurations. The temperature difference between the two configurations is greatest in the deepest layer (bottom panel) with the GT configuration of the model warmer by up to 10 °C, with warm spots coinciding with regions of peak geothermal heat flux (Figure 2.4). The near surface (top 1 m, top panel) mean temperatures typically differ by less than 0.1 °C in the two configurations, with mean temperatures warmer in the GT configuration, save for the region of permafrost and seasonally frozen ground where there are extensive phase changes in the top of the soil column on an annual basis; here, the pattern is more complicated with some points cooler and others warmer in the GT configuration. This might seem to be a paradoxical result, as one would expect temperatures to be uniformly warmer throughout the soil configuration in the GT configuration. The explanation has to do with differing thermal diffusivities in the top of the soil column; in the GT configuration, deeper soil layers undergo a more extensive annual temperature cycle than in the non-GT configuration with the result that the timeseries of the effective thermal diffusivity and its correlation with the skin temperature is altered slightly. The net result of such changes appears to be a slight cooling of the surface layers in regions of extensive phase changes.

While the geothermal heat flux has relatively little bearing on conditions at the top of permafrost, it is a crucial determinant of the characteristics at deeper layers; as Figure 1.1 indicates, the gradual increase in temperature with depth associated with the geothermal heat flux is what determines the bottom boundary of permafrost. Figure 4.31 displays the depth to the base of permafrost in the $20^{th}$ Century in the two model configurations; the North American portion of this plot for the GT configuration of the model was previously plotted in Figure 3.17 (middle panel). While permafrost extends all the way to the base of the soil column in higher latitudes in both configurations; the permafrost characteristics of the warmer permafrost region is dramatically different in the two versions. While the GT configuration was shown to do a good job at capturing the transition between thick and thin permafrost (Figure 3.17), that transition is entirely absent in the no GT configuration as permafrost extends to 250 m, the base of the ground column, everywhere that it is present. Consequently, the inclusion of a geothermal heat flux term is of critical importance when modelling very deep permafrost.

Figure 4.32 shows the transient response of the two configurations following the RCPs. The top panel shows the response of the total permafrost area - the area containing permafrost at *any* depth, while the bottom panel shows the response of the near surface permafrost area. The total permafrost area is comparable in both model configurations over the historical period when both models show a gradual decline, but the $21^{st}$ Century response of the two is very different, with the onset of the rapid decline of total permafrost offset in the no GT configuration by over 100 years. The reason for the discrepancy is apparent from Figure 4.31: the permafrost that is initially and rapidly lost in the GT configuration is essentially all fairly thin permafrost, while this region is occupied by thick permafrost in the no GT configuration.

A plot of a timeseries of the extent of near surface permafrost (permafrost that is present within the top 1 m of the soil - Figure 4.32, bottom panel) shows much stronger agreement between the two configurations. Here, the no GT configuration shows a slight lag relative to the GT configuration until $\sim$ 2200 for all of the RCPs. This is presumably because of a greater heat flux between the near surface and deeper soil layers owing to a strong surface-to-depth temperature gradient in this configuration which results in slightly slower warming of near surface layers. Around 2250, the

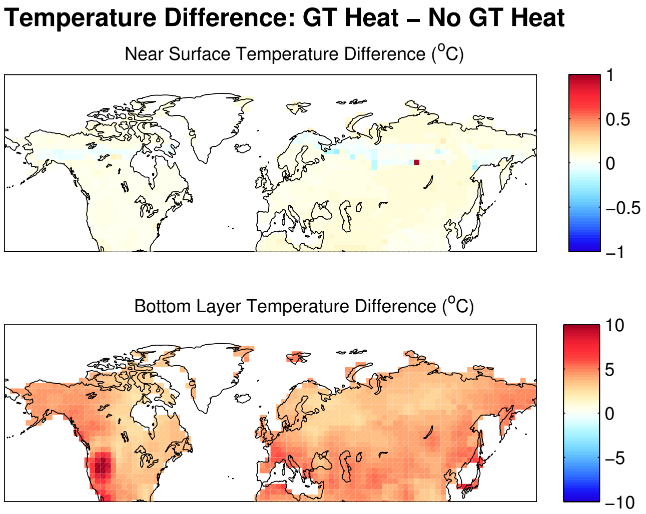


Figure 4.30: Difference in annual average ground temperature in model configurations with and without geothermal heat flux. Top panel: near surface (top 1 m) ground temperature difference; bottom panel: temperature difference in bottommost ground layer.

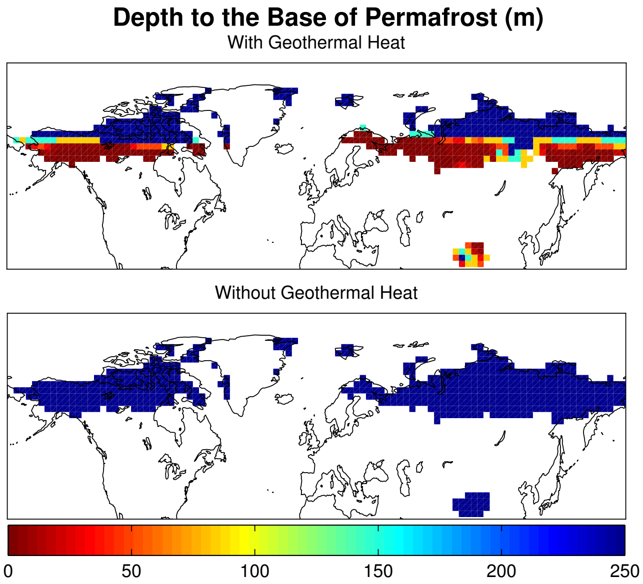


Figure 4.31: Depth to base of permafrost. Top panel: model configuration with geothermal heat flux; bottom panel: model configuration without geothermal heat flux.

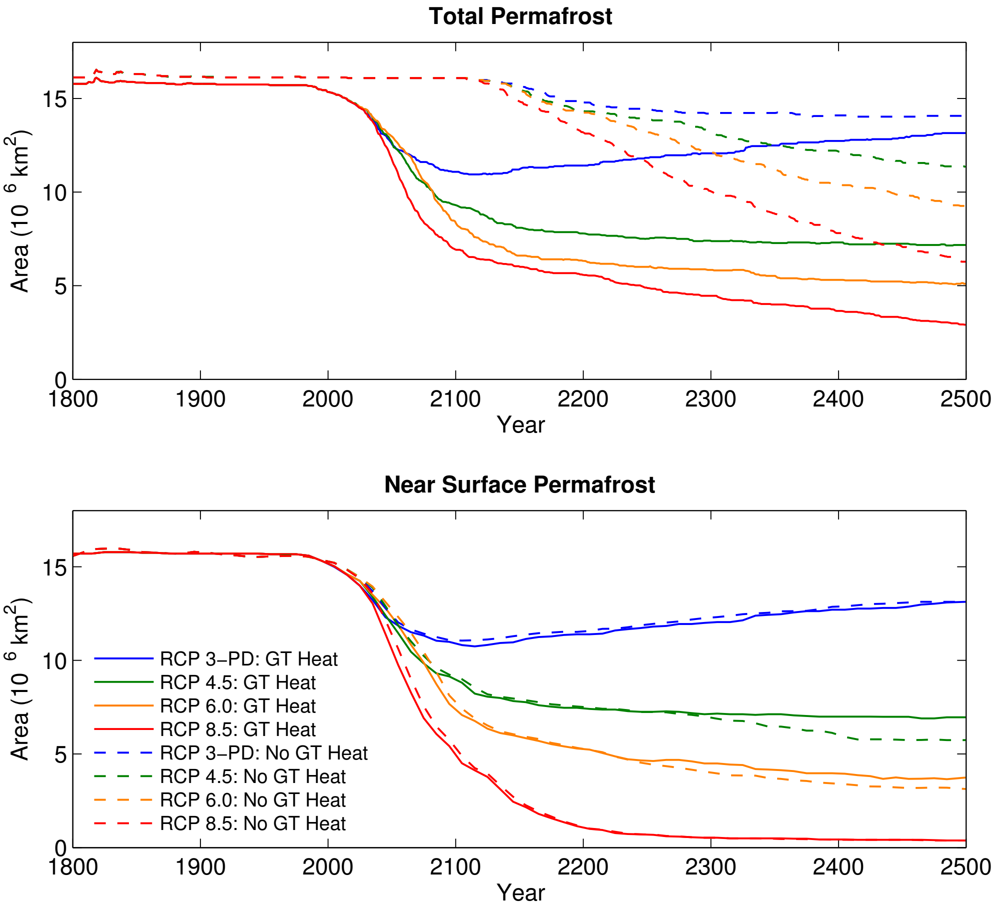


Figure 4.32: Timeseries showing response of permafrost to RCP forcing in models with and without geothermal heat flux included. Top panel: timeseries of total permafrost area; bottom panel: timeseries of near surface permafrost area.

no GT configuration overtakes the GT configuration with a slightly greater rate of permafrost loss in RCP4.5 and 6.0. This appears to be due to slightly different vegetation response owing to different responses in near surface soil moisture content associated with different warming rates of deep soil layers.

Overall, the above plots suggest that if one is strictly concerned about modelling the response of near surface permafrost, then inclusion of a geothermal heat flux term in models is not significant. Conversely, this term is extremely important when modelling the characteristics of permafrost at substantial depths below the surface. Here, accurate measurements of this flux as well as bedrock thermal diffusivity (which, along with flux, determines the steepness of the geothermal gradient and the depth to the base of permafrost) are essential.

### 4.4.4 Soil Thickness Experiment

In the previous chapter, I performed several sensitivity analyses using the offline U3.0E configuration of the model. Of the analyses I performed, I found that the model was most sensitive to changes in the soil thickness (or, depth-to-bedrock). In particular, when thin soil (0.27 m, 1.06 m) configurations of the model were compared against the basic configuration (soil thickness of 10.00 m), it was found that the total permafrost area tended to be lower and the active layer deeper, with the changes particularly pronounced when the soil thickness was 0.27 m. I argued that the lower column moisture content (and associated latent heat associated with melting / freezing) in the thin soil experiments resulted in a greater range of soil temperature and active layer thickness in areas that experience a large amount of freezing and thawing over the course of a year. I now consider how different soil thicknesses (or, depths to bedrock) impact the rate of permafrost degradation. In this section, I have used the same soil thickness configurations as in the previous chapter: 0.27 m (two soil layers), 1.06 m (four soil layers), 5.81 m (seven soil layers) and 29.32 m (ten soil layers). In all cases, the total depth (250 m) and number (14) of ground layers were kept constant. Here, simulations are conducted using the standard (U3.0M) version of the model.

Figure 4.33 compares the year-1800 equilibrium permafrost distribution and simulated active layer thicknesses in the different soil configurations. The same key impacts that were seen in U3.0E (Figure 3.35) are also observed in U3.0M: the very thin soil configurations (0.27 m and 1.06 m) have substantially deeper active layer thicknesses (left column, Figure 4.33) in the southern permafrost zone, while active layer thicknesses are comparable to the standard configuration in the 5.81 m and 29.32 m configuration. The total permafrost areas (blue + red shaded area right column, Figure 4.33) are also reduced in the thin soil versions (14.75 x $10^6$ km$^2$ - 0.27 m, 15.63 x $10^6$ km$^2$ - 1.06 m) compared with the standard configuration (15.74 x $10^6$ km$^2$ - 10.00 m), although the difference in total permafrost area between configurations is not as great as in the uncoupled model. More dramatic is the difference in the near surface permafrost area (blue shaded area, right column, Figure 4.33) which is considerably reduced in the thin soil configurations (2.61 x $10^6$ km$^2$ - 0.27 m, 12.85 x $10^6$ km$^2$ - 1.06 m) relative to the standard configuration (15.70 x $10^6$ km$^2$ - 10.00 m) on account of the deeper active layer thicknesses of the thin soil configurations that

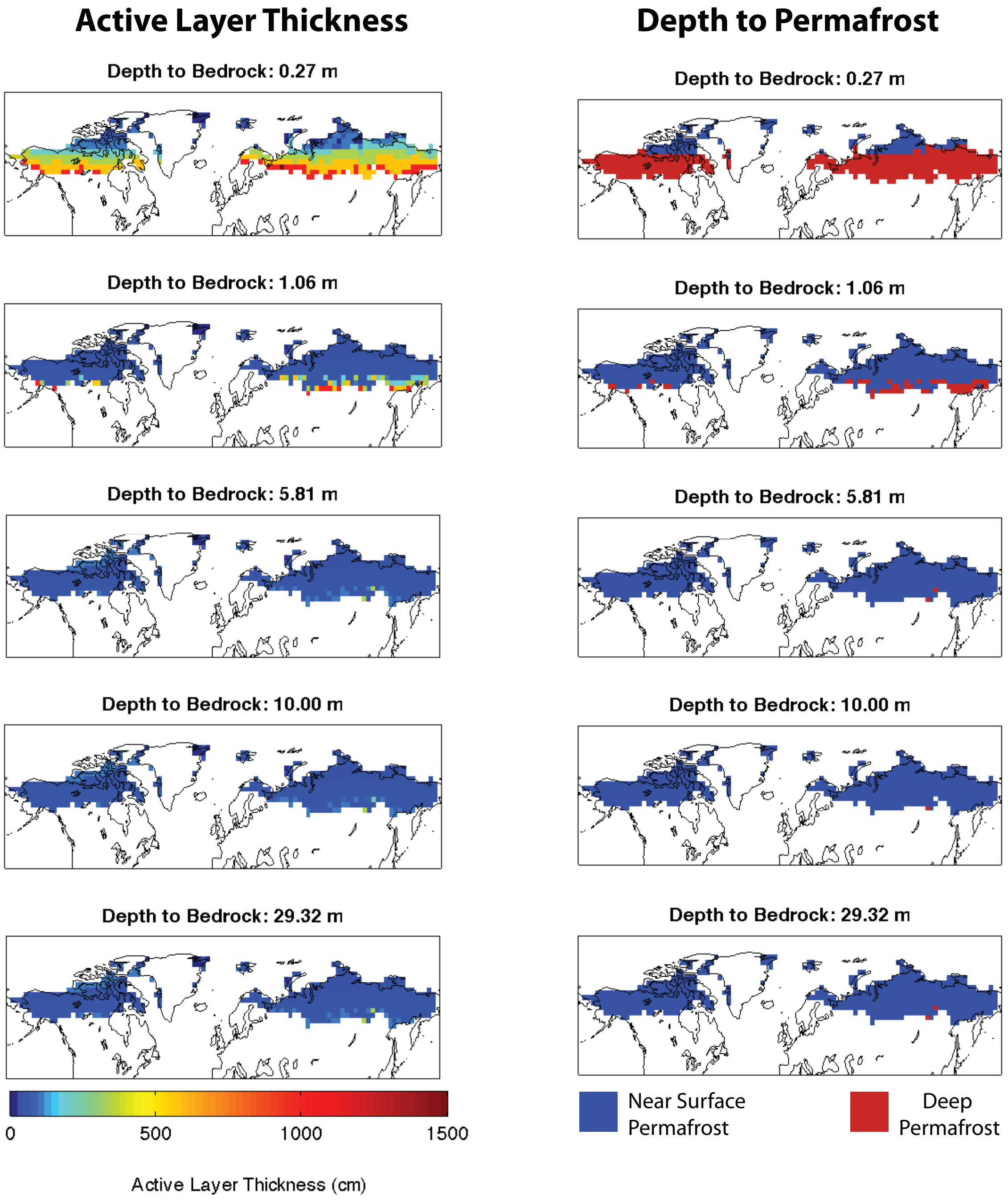


Figure 4.33: Distribution of permafrost and active layer thickness in model configurations with different depth to bedrock. Left column: active layer thicknesses (cm); right column: depth to permafrost. Here, two depth categories are used. The Near Surface Permafrost category indicates that the permafrost table is within the top 1 m of the ground; otherwise, permafrost is categorized as Deep Permafrost.

prevent permafrost from being present in the top soil layers.

Next, I consider the response of permafrost to RCP forcing. My initial supposition was that thin soil configurations of the model should show the most rapid permafrost decay as a consequence of the greater average thermal conductivity and greatly reduced average heat capacity when phase changes occur (owing to lower soil moisture storage and associated latent heat of fusion) in the top of the ground column. This supposition certainly holds true when one considers the response of near surface (N.S.) permafrost as shown in Figure 4.34. Here, I plot the response of N.S. permafrost in RCP3-PD and 8.5 only for space considerations. The top row of this figure shows the response of the near surface permafrost area in each soil depth configuration, clearly showing the reduced N.S. permafrost area in the thin soil configurations; it is also evident that the N.S. permafrost response is essentially the same in the three deep soil configurations.

To demonstrate that N.S. permafrost degrades more rapidly in the thin soil configurations, I identified the area where all of the soil configurations simulate N.S. permafrost at year 1800 - this is plotted as the ‘common N.S. Permafrost Area’ in the bottom row of Figure 4.34. This area generally incorporates only permafrost at very high latitudes (common blue area in the right column of Figure 4.33) as this is the only region where the 0.27 m configuration simulates permafrost at year 1800. This very thin soil configuration shows dramatic changes in the near surface permafrost area over even the historical period, while the four thicker soil configurations show no change in this area in RCP3-PD. The rate of loss of near surface permafrost over the historical period in the 0.27 m configuration is considerably greater than observed, demonstrating, as in the previous chapter, that assigning such a shallow depth to bedrock over the entire permafrost domain results in an unrealistic representation of permafrost characteristics. However, in some regions of the Arctic, bedrock certainly is this shallow and the figure demonstrates that N.S. permafrost decay will proceed at a markedly different rate in such regions relative to deeper bedrock areas.

In RCP8.5 configuration, N.S. permafrost degrades faster in the two thinnest soil configurations than in the thicker soil configurations. However, the rate of decay is substantially faster in the 0.27 m configuration, while the decay rate is relatively similar in the 1.06 m configurations to the deep soil configurations. As in Chapter 3, these results suggest that the response of permafrost may differ markedly when the depth to bedrock is very shallow - within the first metre of the ground column - but when the depth to bedrock is deeper than this, the response of near surface

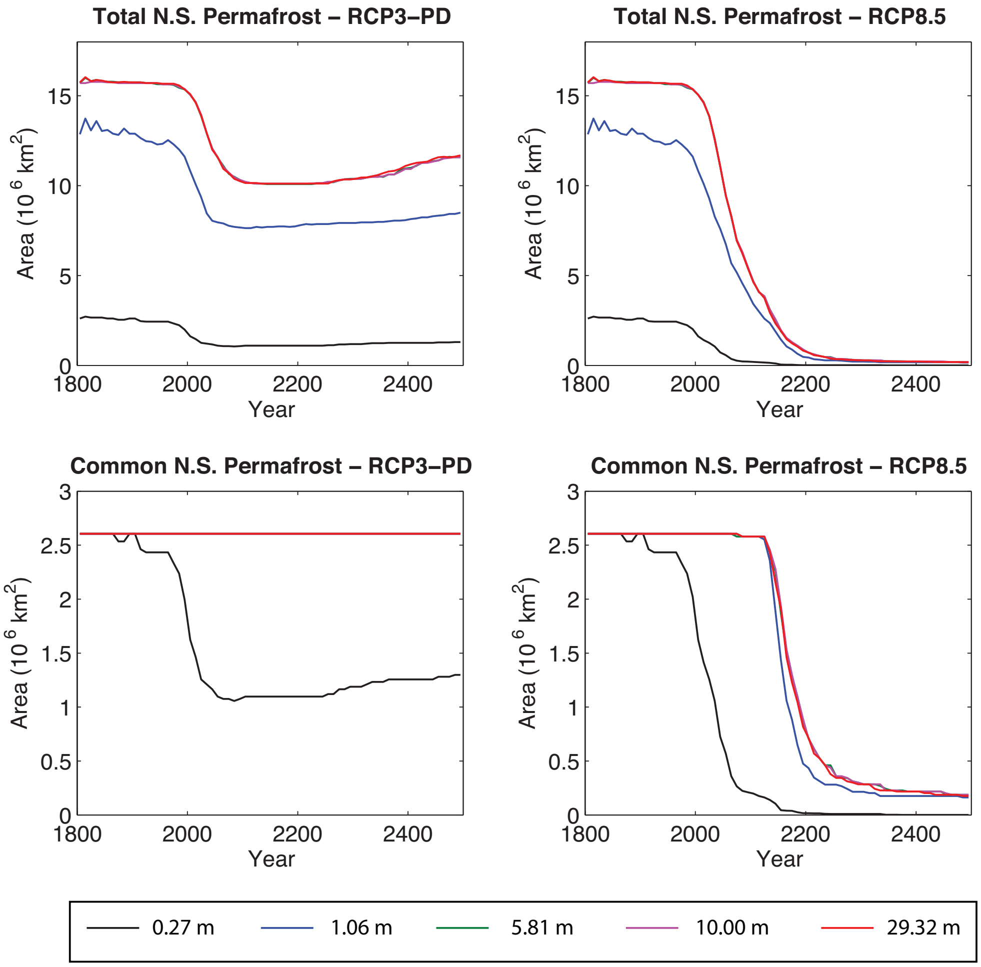


Figure 4.34: Response of regions of near surface permafrost to Representative Concentration Pathway forcing in models with different depth to bedrock. Left column: response following RCP3-PD; right column: response following RCP8.5. Top row: response of near surface permafrost area in each model configuration; bottom row: response of common near surface permafrost area at equilibrium.

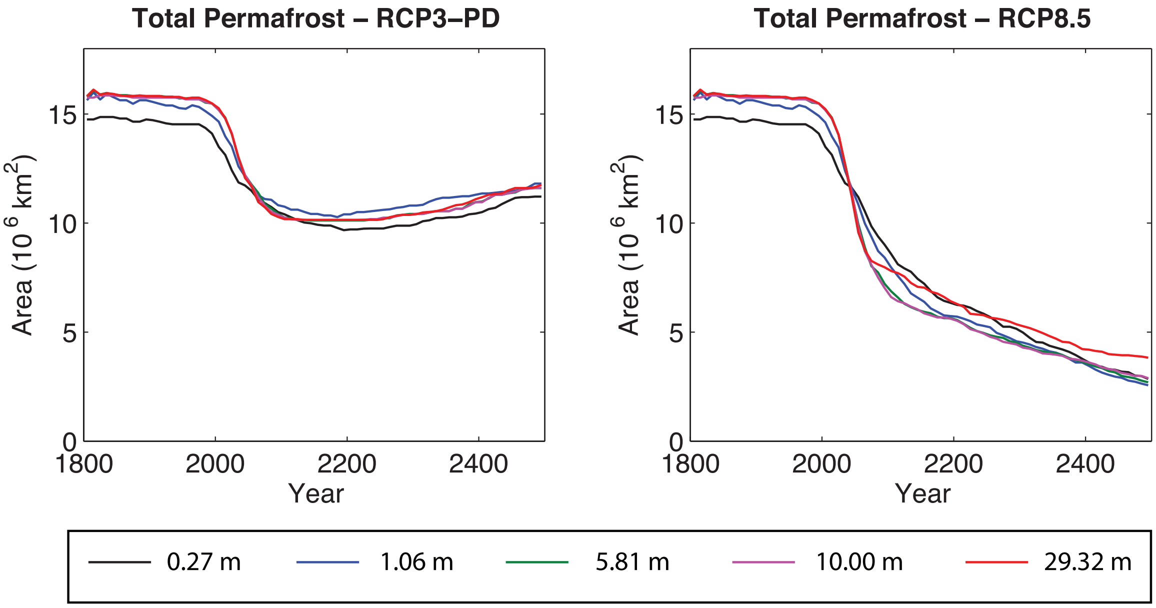


Figure 4.35: Response of total permafrost area to Representative Concentration Pathway forcing in models with different depth to bedrock. Left column: response following RCP3-PD; right column: response following RCP8.5. Top row: response of total permafrost area in each model configuration; bottom row: response of common total permafrost area at equilibrium.

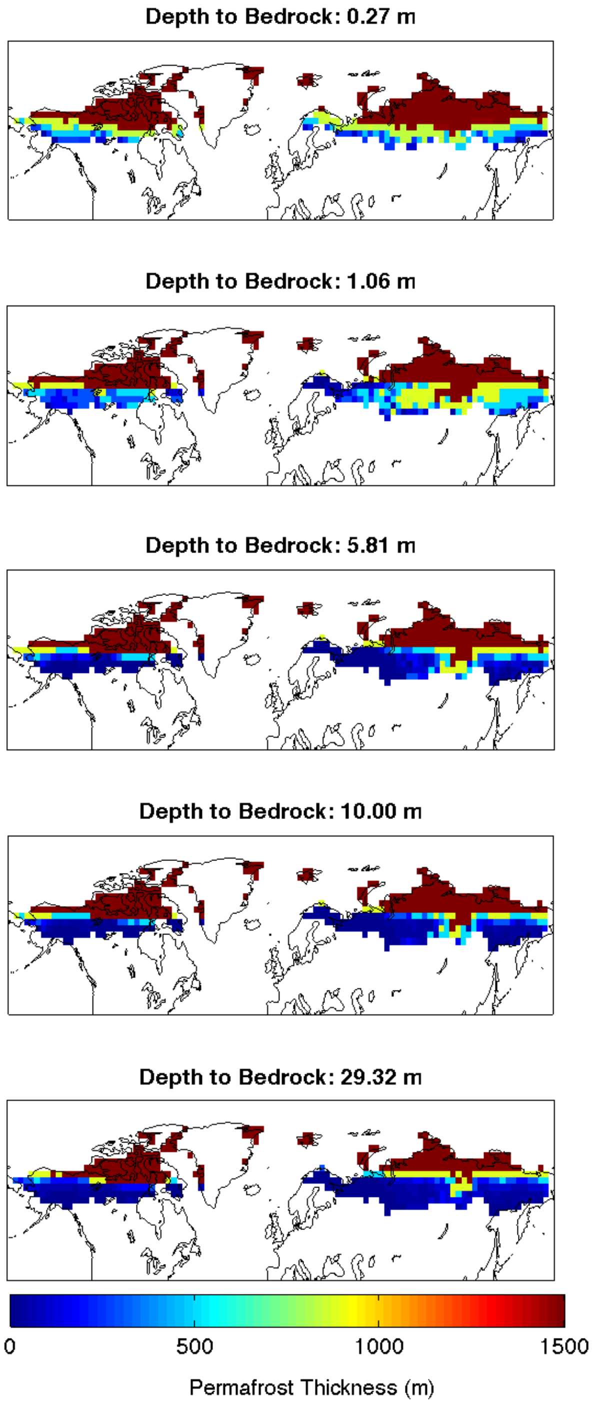
permafrost is essentially independent of bottom boundary of soil.

I next considered the response of the total permafrost area in the different soil depth configurations (Figure 4.35). Surprisingly, although there are substantial differences in the rate of degradation of near surface permafrost, the rate of loss of total permafrost is fairly similar between the five model configurations. The explanation for this apparent paradox relates to another pronounced difference between the depth-to-bedrock configurations: the total thickness of permafrost. Figure 4.36 is a plot of the thickness of permafrost in each of the model configurations; while all configurations simulate thin permafrost at the lower limits of the permafrost zone and permafrost extending all the way to the base of the ground column at high latitudes, the position of the transition zone between these regions differs between configurations. As the depth to bedrock increases, the transition line between thin and thick permafrost shifts northwards, and the average permafrost thickness and total volume of permafrost decreases. Thus, while the rate of warming of permafrost is greater in the thin soil configuration so too is the total volume of permafrost; these two countervailing effects appear to be roughly the same magnitude in terms of their impact on the degradation of total permafrost area and consequently the rate of degradation of total permafrost is similar in all bedrock depth configurations.

To understand what causes the difference in permafrost thickness in the different configurations, I produced a plot of the annual mean ground temperature profile (depth vs. mean temperature - Figure 4.37) for a gridcell where the position of the bottom permafrost boundary changes markedly as the depth to bedrock is varied. The plot is split to show key features in the thermal profile near the surface and at depth. The depth of the permafrost base occurs where the temperature at depth crosses 0 °C. Below $\sim$ 50 m, the temperature profiles are parallel due to the fact that the thermal conductivity and the geothermal heat flux is the same in each of the model configurations. To explain why the depth to permafrost varies, one must therefore examine processes that occur higher up in the soil column.

In the absence of soil moisture phase change effects, the ground temperature should increase with depth following the geothermal gradient. In such a no-phase change scenario, variations in ground thermal diffusivity manifest themselves as 'kinks' in what would be an otherwise straight soil temperature curve. The curves show two distinct 'kinked' regions - the first kink, near the surface results from treating the top soil layer as a combined soil/snow layer (as outlined in Chapter 2). When snow is present, the thermal conductivity of this layer is lowered from that of soil, with the result that the temperature gradient between the first two ground points is steeper than elsewhere in the soil profile. The second kinked regions occurs near the depth of bedrock where the thermal conductivity increases from soil to bedrock (recall that the thermal conductivity of bedrock is uniformly greater than that of soil, even when soil pores are fully ice-filled). This is reason for the shifted temperature curves of the three thicker soil configurations of the model; in these three configurations, the temperature at the base of the active layer is essentially identical and the reason for the shift in ground temperature is due to the different position of the boundary between bedrock and soil. Since the steepness of the geothermal gradient at depth is set by the bedrock thermal conductivity, this indicates that accurate representation of bedrock thermal characteristics is a region is key to realistic simulation of the bottom boundary of permafrost.

The strong temperature shift between the 0.27 m and 1.06 m depth to bedrock configurations and the deeper bedrock configurations is more complicated and has to do with soil phase changes; in each of the configurations the annual cycle in snow depth for this gridcell was nearly identical (not shown), ruling out snow insulation



Depth to Bedrock: 0.27 m

Depth to Bedrock: 1.06 m

Depth to Bedrock: 5.81 m

Depth to Bedrock: 10.00 m

Depth to Bedrock: 29.32 m

Figure 4.36: Permafrost thickness (m) in model configurations with different depths to bedrock.

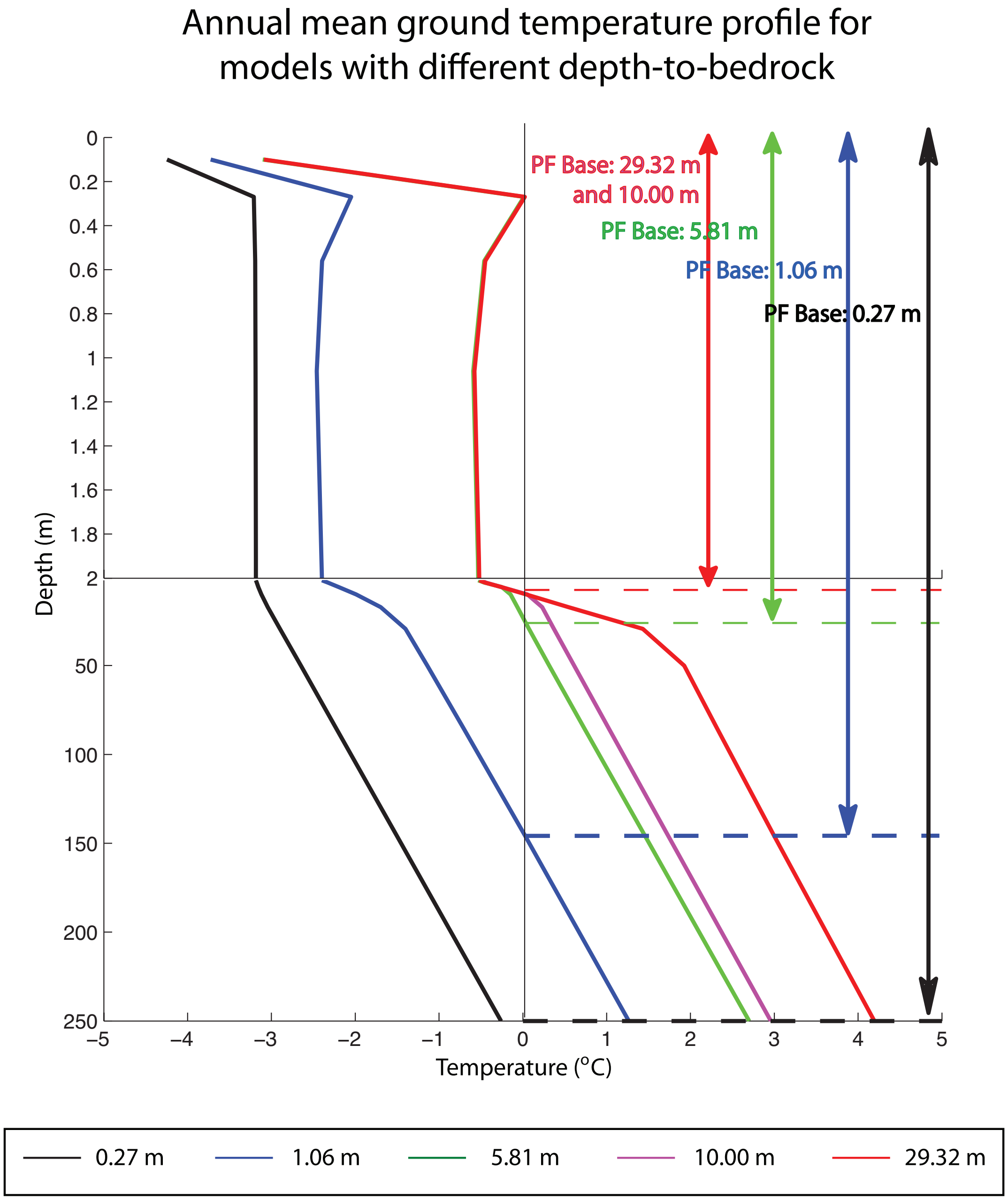


Figure 4.37: Annual mean ground temperature profile for a grid cell in model configurations with different depths to bedrock; also shown is the depth to the base of permafrost (PF base) determined where the deep ground temperature intersects the zero degree line. The profiles generally show two distinct kinks; the first is the so-called 'thermal offset' within the active layer due to phase change effects. The second kink occurs deeper in the ground profile at the depth to bedrock where the thermal diffusivity switches from that of soil to that of bedrock. Note that the y-axis is divided into two regions with different scales.

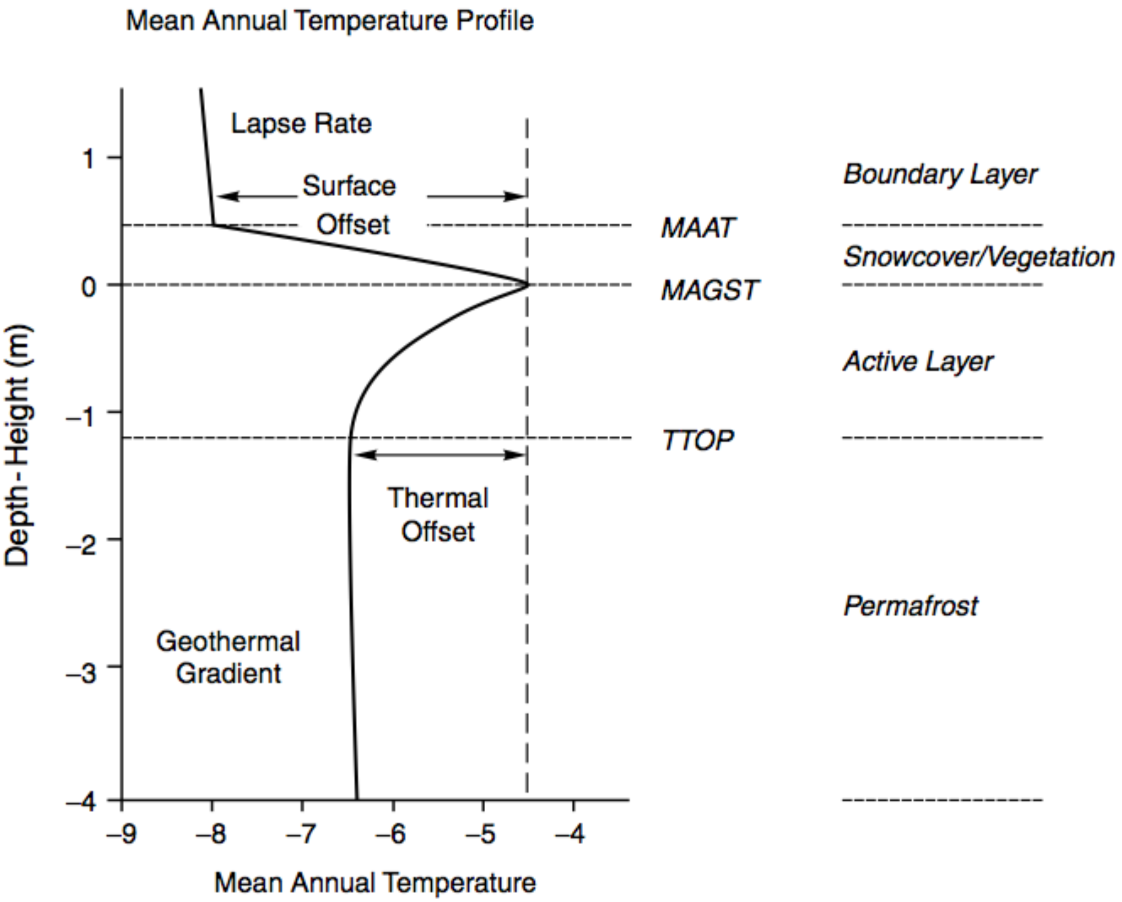


Figure 4.38: Representative annual mean air, surface and ground temperatures in permafrost. The diagram illustrates the mean annual air temperature (MAAT), mean annual ground surface temperature (MAGST) and temperature at the top of permafrost (TTOP). The figure also illustrates the surface offset: the increase in MAGST over MAAT due to the insulating effect of snow as well as the thermal offset: the increase in MAGST relative to TTOP due to the effect of annual phase changes in the soil. Figure taken from Smith and Riseborough (2002).

changes as the cause of the temperature profile differences. Here, it is helpful to consider Figure 4.38, taken from Smith and Riseborough (2002) which describes key features of a representative ground surface profile. In this diagram, the mean annual atmospheric temperature (MAAT), temperature at the top of permafrost (TTOP) and mean annual ground surface temperature (MAGST) are defined. The latter is thought of as a thermal boundary representing the depth at which conduction becomes the dominant energy transport processes and not necessarily strictly the very surface of the ground, per se. The surface offset corresponds to the increase in MAGST relative to MAAT owing to the thermal insulation effects of lying snow; each configuration of the model shows such an increase in temperature with the temperature of the top two layers of the model greater than the skin temperature (skin temperature not shown in figure 4.37).

Where ground phase changes are significant, the temperature profile shows a second offset, the thermal offset, between the MAGST and the TTOP, with the extent of the offset dependent on soil moisture content and thermal characteristics. When phase changes occur within soil, the temporal lag between the soil temperature timeseries and the skin temperature curve increases, as shown in Figure 4.39 which compares the annual cycle in temperature in the skin temperature and top six soil layers in the 0.27 m and 10.00 m depth to bedrock configurations. The occurrence of phase changes may be noted by the so-called 'zero-curtain' effect wherein the temperature of a soil layer remains nearly isothermal just below 0°C as phase changes occur. In the case of the 0.27 m configuration, there is very little soil moisture and the zero curtain effect is minimal, except for in the first two soil layers and the soil temperatures closely track the skin temperature. French (2007) notes that in solid bedrock where there is little or no moisture, the thermal offset is minimal or non-existant and the plot of Figure 4.37 with its absence of an offset for the 0.27 m configuration supports this observation.

Conversely, in the 10.00 m configuration, the zero curtain effect is pronounced in several layers and the period in which deeper layers undergo phase changes can cause layers near the surface to substantially lag behind the skin temperature in terms of the timing of minimum and maximum temperatures. The overall result is a temperature offset which results in the near surface layers having a greater annual mean temperature than layers at the top of permafrost. An earlier numerical study by Goodrich (1978) may be used to explain the results presented here. Goodrich (1978) found that when ground thermal conductivity was allowed to vary through-

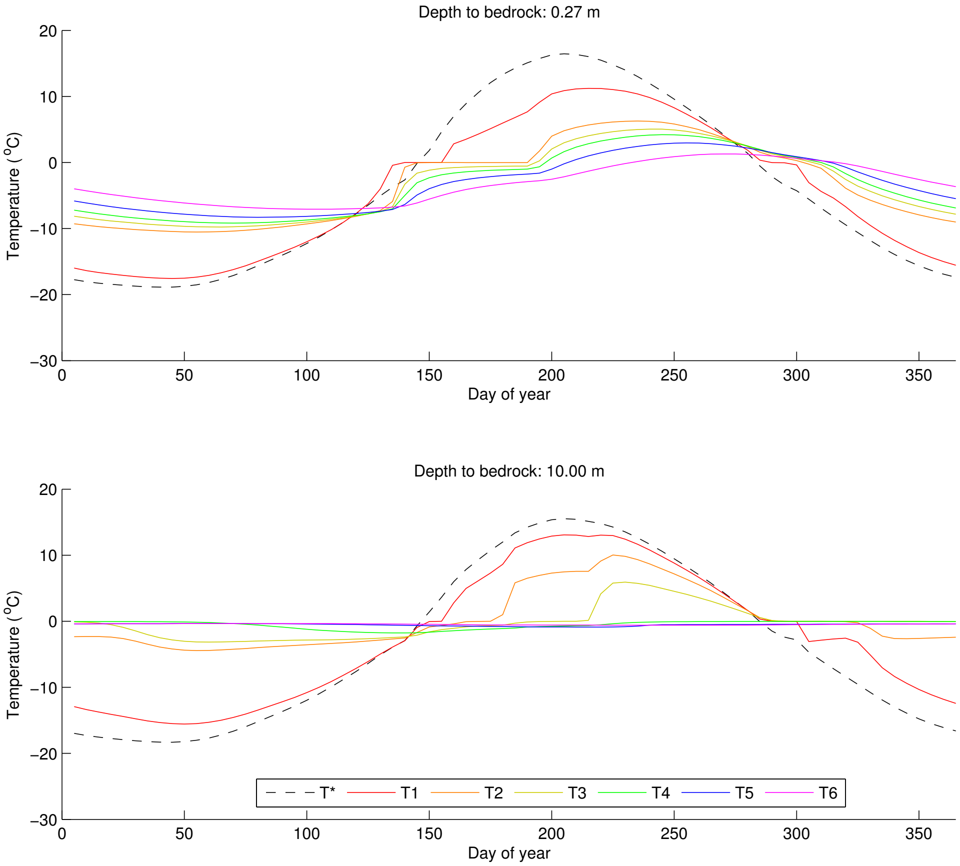


Figure 4.39: Annual cycle in ground temperature and skin temperature in a gridcell in permafrost. Top: gridcell in model configuration with depth to bedrock of 0.27 m; bottom: gridcell in model configuration with depth to bedrock of 10.00 m. Differences in the temperature cycles are attributable to the different number of layers able to undergo phase changes in each case.

out the year but the latent heat transfer accompanying freezing and thawing was not modelled that ground temperatures decreased with depth within the active layer strictly as a consequence of the variation in ground thermal conductivity throughout the year. When latent heat transfer was simulated with no variation in ground thermal conductivity, the heat source supplied from deeper thawed layers was sufficient to maintain the top layer ground temperatures at relatively high values during the freezing season, a phenomenon Goodrich (1978) termed 'pinning'. This resulted in an increase in annual mean ground surface temperature that was also reflected in a shift in the temperature profile at depth. Thus, the differences in the ground temperature profiles presented in Figure 4.37 may be understood as being due to differences in the extent of the ground that undergoes variations in thermal conductivity (resulting in a more extensive thermal offset) and is subject to phase changes (which results in a shift in the mean annual value of the entire temperature profile).

As Figure 4.37 indicates, the thermal offset can allow permafrost to be present in deeper layers, even though the mean temperature of layers close to the surface exceeds 0 °C. As well, as Figure 4.37 suggests, the mean annual temperature at the top of permafrost may also be substantially shifted if the ground column experiences phases changes throughout a greater depth. Deep layers that are subject to little-to-no annual cycle in temperature have a temperature that is set by the annual mean temperature of higher layers. Consequently, if higher layers' mean temperature is increased as phase changes occur throughout deeper sections of the ground column (transition from 0.27 m to 1.06 m and 1.06 m to 5.81 m configurations), then the temperature of deep layers may be shifted as well. Thus, the different permafrost thicknesses in the depth to bedrock configurations result both from shifts in mean temperature at depth determined by the extent to which the top layers of the soil undergo phase changes throughout the year as well as by the shifts in the thermal gradient at depth due to the different thermal conductivity of bedrock and soil.

To integrate and clarify these findings, I considered the thermal response of the total volume of permafrost (Figure 4.40), calculated by summing the product of permafrost thickness and gridcell areas over the permafrost domain. The plot clearly illustrates the differences in equilibrium permafrost thicknesses between the different soil depth versions. Despite having the smallest near surface permafrost area, the total volume of the shallow bedrock configuration is the greatest ($\sim 2.5 \times 10^6$ km$^3$) at year 1800, while the volume of the very deep bedrock configuration is the least ($\sim 1.5 \times 10^6$ km$^3$). However, the reduction in latent heat that must be supplied to thaw

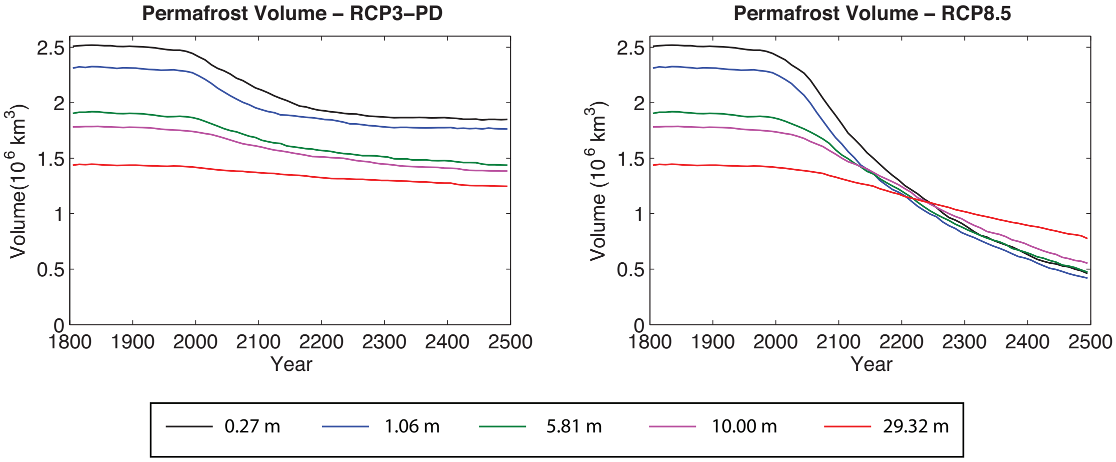


Figure 4.40: Response of permafrost volume ($10^6$ km$^3$ to RCP forcing in models with different depth to bedrock. Left panel: RCP3-PD, right panel: RCP8.5.

permafrost and the greater thermal conductivity of bedrock (compared to deep soil), means that the permafrost in the shallow bedrock configuration shows the most rapid response to the RCP forcings and the greatest rate of decay.

The complicated results from this sensitivity experiment illustrate two key points. Firstly, rates of permafrost degradation can occur very differently at different depths in the ground column, so when comparing permafrost results from different models, it is important that a common depth or thickness of permafrost be compared. Conversely, examining the rate of permafrost volume change is also a very useful metric in understanding differences in model configurations. Secondly, the UVic ESCM has been demonstrated to be very sensitive to bedrock characteristics, in terms of its total thickness and the thickness of the active layer, and its rate of degradation. Consequently, improved data on bedrock depth and thermal characteristics has the potential to greatly improve simulation of permafrost within climate models.

## 4.5 Summary

In this chapter, the fully coupled version 3.0M of the UVic ESCM was forced over the historical period through to year 2500 following four Representative Concentration Pathways in order to produce estimates of the response of frozen ground to future climate warming. I first compared the model's surface air temperature and precipitation response against results from a multi-model ensemble of GCMs from the IPCC

Fourth Assessment report following comparable SRES scenarios:

- **Surface Air Temperature Response**: Future warming is greatest over land and in high latitude regions. The UVic ESCM's global mean SAT response following RCP scenarios agrees well with IPCC ensemble mean SAT response. Relative to the mean IPCC results, the UVic ESCM shows lower land-ocean contrast in terms of warming and less intense warming at high latitudes. Within the Arctic, warming tends to be most pronounced in the fall and winter months.
- **Precipitation Response**: IPCC models project increased precipitation rates within the tropics and high-latitudes, with a reduction in precipitation in middle latitudes, with an overall global increase in precipitation. The UVic ESCM generally captures this pattern, though shows anomalous precipitation changes in certain regions; most significantly, the model shows a reduction in summer precipitation in the Northern Hemisphere mid-to-high latitudes where the IPCC models project increased precipitation. Like the IPCC results, the model shows a general increase in precipitation in the Arctic with precipitation rates most strongly increasing in the late fall through to the early spring. Overall, P-E* increases, or remains the same throughout the Arctic and Sub Arctic.

These climate changes generally result in a greening of the Arctic in the ESCM with northward expansion of shrubs into grasslands and expansion of needleleaf trees into shrublands. Vegetation changes are particularly extreme in RCP 8.5 where Needleleaf trees dominate nearly everywhere in the Arctic and Sub Arctic save in areas where decreases in near surface soil moisture storage associated with permafrost loss prevent trees from thriving. Next, I considered the simulated response of the exposed cryosphere in North America:

- There are only modest decreases in the extent of permafrost, seasonally frozen ground and the maximum annual snow covered area over the historical period. There is a sharp increase in the rate of loss of these quantities beginning in the early 21$^{st}$ Century. By year 2100, permafrost has extent has decreased by between 28.51 % (RCP3-PD) and 53.82 % (RCP8.5), Seasonally Frozen Ground (SFG) has decreased by between 8.32 % and 22.37 % and Snow Covered Area (SCA) has decreased by between 10.24 and 23.95 %, although one might expect that these rates of decrease might be lower than expected in the UVic ESCM given its relatively weak polar amplification. The UVic ESCM compares well with other modelling groups that simulate permafrost within coupled models in terms of its present day distribution of permafrost. The model also lies in the middle-to-low side of the range of projected rates of permafrost loss in these models. By 2100, there is a reduction in SFG of between 8.32 % (RCP3-PD) and 22.37 % (RCP8.5) by 2100 relative to the reference period, which also agrees well with results from other modelling groups.

- Permafrost loss occurs first at the southern limits of the permafrost distribution, where it is thinnest and warmest. The region of greatest loss in the $21^{st}$ Century is in Central-Southern Siberia where warming is greatest. Accompanying permafrost thaw, is a deepening of the active layer that is most apparent in the southernmost permafrost sites.

- Recovery of PF, SFG and SCA occurs in the RCP-3PD scenario where radiative forcing peaks in 2050 and begins to decline thereafter. Following this pathway, permafrost reaches a minimum of 10.39 x $10^6$ km$^2$ in 2110 and slowly increases after this point thereafter, while SFG and SCA begin to recover $\sim$ 50 years earlier. Radiative forcings stabilize in 2150 for RCP4.5 and 6.0, but a reduction in the extent of PF, SFG and SCA continues until the end of the simulation, although these quantities appear to have nearly stabilized at 2500 in these scenarios. Peak radiative forcing is reached at 2250 in RCP8.5, and the extent of permafrost continues to decrease in year 2500 at the end of this simulation; here, only 2.95 x $10^6$ km$^2$ of permafrost remains in the high Arctic and much of this remaining permafrost would likely be lost if the simulation were allowed to continue.

- Using the results of the RCP8.5 simulation, I demonstrated that permafrost loss occurs both by deepening of the active layer and by talik formation. Thin permafrost degrades strictly by deepening of the active layer, while the loss of thicker permafrost is initiated by deepening of the active layer followed by talik formation. Both talik formation and the loss of near surface permafrost can be thought of as early indicators of the total loss of permafrost in a given location. The presence of taliks nearly everywhere in the remaining permafrost zone in RCP8.5 further indicates that permafrost will be lost nearly everywhere in the Northern Hemisphere if this high radiative forcing state were to be maintained beyond the end of the simulation.

I then conducted a series of experiments to assess the sensitivity of the rate of permafrost degradation to the configuration of the land surface scheme. The key observations for each sensitivity experiment were as follows:

- **Excess Ice Experiment**: Two configurations of the model containing ground ice in excess of soil pore spaces were prepared, one in which ice is distributed throughout the soil column (structural ice) and one in which ice is concentrated in a single layer at the base of the column (slab). The rate of permafrost degradation is slower in the two excess ice configurations of the model, with the structural ice slowest and the greatest degree of divergence occuring over the $21^{st}$ Century in the two lower RCP scenarios. However, the overall rate of permafrost degradation beyond the $21^{st}$ Century is not radically different due to the presence of excess ice. A more significant consequence of the presence of excess ice is the potential for ground subsidence. The southern limits of permafrost are shown to be most vulnerable to ground subsidence with subsidence ranging between 1 to 2.5 m by the end of the $21^{st}$ Century. Ground subsidence can be as much as 4 m by year 2500, pointing to the importance of long term simulations when assessing hazards in regions of permafrost. My subsidence results differ from an earlier hazard analysis by Nelson et al. (2001) which generally showed highest ground settlement risk in the $21^{st}$ Century in regions of highest ground ice content.

- **Vegetation Dynamics Experiment**: A configuration of the model in which vegetation was allowed to evolve dynamically following changes in climate was compared with a configuration in which the vegetation distribution and characteristics were fixed. The non-dynamic vegetation configuration generally showed a greater rate of permafrost loss compared with the dynamic configuration, with the greatest difference between the two configurations in the lower RCP scenarios (despite the fact that vegetation changes were greatest in the higher scenarios). However, permafrost does not uniformly decay faster everywhere in the non-dynamic configuration of the model - the response depends on the nature of the vegetation change that occurs. Where expansion of trees occurs, the rate of permafrost loss is lower in the dynamic configuration and where trees are lost, the rate of permafrost loss is greater. These observations may be explained with reference to the vegetation sensitivity experiments of the previous chapter where it was found that expansion of trees allowed for a greater cooling of the soil through an enhancement of surface roughness.

- **Geothermal Heat Flux Experiment**: The response of the standard configuration of the model (which includes a geothermal heat flux term) was compared against a version of the model with a zero flux bottom boundary condition. Omitting the geothermal heat flux eliminates the geothermal gradient which produces a bottom boundary to permafrost; consequently, in this version of the model, where permafrost is present, it extends to the base of the ground column everywhere and there is no transition zone between thin and thick permafrost. Degradation of near surface permafrost occurs at a comparable rate in both situations, though slightly more slowly in the no geothermal heat flux configuration owing to enhanced heat transfer from surface layers to depth. The results suggest that a representation of the geothermal heat flux is not crucial in simulating the transient response of near surface permafrost, but is essential in simulating the characteristics of deep permafrost.

- **Depth to Bedrock Experiment**: The model showed the greatest sensitivity to changes in the depth to bedrock. When the depth to bedrock is shallow, the active layer is deeper and permafrost warms more rapidly, both as a consequence of a reduction in soil moisture storage and accompanying latent heat fluxes required to thaw ground ice. Consequently, the equilibrium distribution of near surface permafrost is considerably reduced in the shallow bedrock configurations, and the near surface permafrost that is present decays at an enhanced rate relative to deeper bedrock configurations. Somewhat counterintuitively, the shallow bedrock configuration shows a greater thickness of permafrost owing to a reduction in the thermal offset (warming) of near surface layers and a reduced depth at which the ground thermal conductivity switches from soil to bedrock; both factors result in an increase in the depth of the permafrost base in the shallow bedrock configurations.

The results from this section also point to the importance in precisely specifying the depth of permafrost under consideration as the thermal response of near surface and deep permafrost may differ in certain sensitivity experiments. An examination of changes in the volume of permafrost rather than strictly its areal extent is also helpful in understanding and elucidating differences between model configurations.

# Chapter 5

# Response of High-Latitude Wetlands to Permafrost Thaw

## 5.1 Introduction and Author's Note

In Appendix B of the thesis, I examine the impact of changes made to the land surface scheme of the UVic ESCM to enable it to model permafrost in terms of changes to the thermal and hydrological state of the ground. I found that one of the largest impacts of the inclusion of soil freeze/thaw processes was a substantial increase in the storage of soil moisture in the top layers of the soil column (Figure B.7) in permafrost regions which is attributable to the reduction in soil hydraulic conductivity when frozen moisture is present. As permafrost degrades, the hydraulic conductivity of the soil should increase, allowing for enhanced drainage of soil moisture, but, accompanying Arctic warming is an overall increase in precipitation (Figure 4.11) and an increase in P-E (Figure 4.12). Here, I demonstrate that the present-day distribution of global and high-latitude wetlands is well reproduced by a simple function of near surface, unfrozen moisture content and investigate the response of high-latitude wetlands to Arctic warming.

This chapter is derived with only minor modifications from an article published in Nature Geoscience (Avis et al. (2011)) by myself and my supervisors, Dr. Andrew Weaver and Dr. Katrin Meissner. After publication, I realized that I had inadvertently run the wetland simulations with the deep soil (29.32 m depth to bedrock) configuration of the model as opposed to the standard 10.00 m depth to bedrock configuration. As demonstrated in this chapter, changes in wetlands depend on the degradation of near surface permafrost, changes in which were essentially identical between the 10.00 and 29.32 m configurations of the model, as demonstrated in the previous chapter. Consequently, I have largely retained the article's original text and figures, rather than repeating the analysis.

Wetlands are vegetated regions that are inundated on a permanent, seasonal or intermittent basis (Wheeler, 1999). They play an important role in the carbon cycle through carbon uptake and storage in vegetation and soil and through carbon dioxide and methane release from bacterial decomposition of organic matter (Mitra et al., 2003). Over 50% of wetlands are located in high northern latitudes (Wania, 2007) where perennially frozen ground (permafrost), exerts strong controls on wetland hydrology (Woo and Winter, 1993). Observations have linked permafrost degradation to changes in Arctic lakes: from 1973-2004, the abundance of lakes increased in continuous permafrost zones, but decreased in other zones (Smith et al., 2005).

## 5.2 Model Description

The UVic Earth System Climate Model (ESCM, (Meissner et al., 2003; Weaver et al., 2001)) is an intermediate complexity climate model consisting of several coupled components: an energy-moisture balance atmospheric model, a three-dimensional ocean general circulation model, a dynamic-thermodynamic sea ice model and a land surface scheme which includes a representation of vegetation dynamics. All model components have a common horizontal resolution of 1.8° x 3.6°. The version of the model used in this work features a representation of soil freeze-thaw processes (Cox et al., 1999) and an increased number of subsurface layers in the land surface scheme. There are 14 subsurface layers whose thicknesses exponentially increase with depth with the surface layer having a thickness of 0.1 m and the bottom layer a thickness of 104.4 m. The total thickness of the subsurface layers is 250 m. The top ten layers (to a depth of 29.32 m) are soil layers; below this are bedrock layers having the thermal characteristics of granitic rock. Moisture undergoes free drainage from the base of the soil layers and the bedrock layers are hydrologically inactive.

I also allow for spatially varying mineral and organic soil content. Mineral soil thermal and hydrological parameters are determined from the relative abundance of clay, sand and silt sized particles as specified from the ISLSCP-II soil characteristics data set (Scholes and Brown de Colstoun, 2011), which gives particle abundance from 0-0.3 m and from 0.3-1.5 m. I use the latter dataset to specify mineral soil characteristics for all soil layers deeper than 0.3 m. In actuality, the abundance of coarse particles would likely increase below 1.5 m, possibly enhancing drainage from deep soil layers, but we are unaware of any global-scale datasets that could be used to specify such deep soil characteristics. In specifying soils organic characteristics, I follow the methodology of an earlier study (Lawrence and Slater, 2008). My implementation of organic soils differs slightly from the original study in that I use a single globally-average cumulative carbon distribution function and use a hydraulic conductivity of 0.02 kg $m^{-2}$ $s^{-1}$ (intermediate between values for hemic and fibric peat) for organic soil, a value an order of magnitude lower than the value of 0.1 kg $m^{-2}$ $s^{-1}$ for fibric peat used in the earlier study (Lawrence and Slater, 2008). This was found to be necessary to maintain near surface unfrozen moisture contents at levels consistent with organic soil laden wetlands.

Model grid cells containing permafrost are those that include at least one subsurface layer remaining at or below 0 °C for two or more consecutive years. I simulate a Northern Hemisphere (NH) permafrost area of 16.10 x $10^6$ $km^2$ from 1970-1989, compared with estimates over the same period of 12.21-16.98 x $10^6$ $km^2$ for the total exposed NH land area underlain by permafrost (Zhang et al., 2000). Nearly all of this permafrost (15.63 x $10^6$ $km^2$) extends to within the top 1 m of the ground - this I term ‘near surface permafrost’ (N.S. permafrost). The models relatively high permafrost extent is attributable to the fact that the model contains ~8 % more exposed NH land area than observations due to its coarse-resolution grid (and lack of resolved lakes, islands and coastlines) as well as to regional temperature biases (Meissner et al., 2003).

Model simulations were conducted from 1800 to 2500 using forcing data from the harmonized Representative Concentration Pathway datasets (Moss et al., 2010). I forced the model according to the RCP forcing pathways by specifying atmospheric $CO_2$ concentrations, the radiative forcing from all non-$CO_2$ greenhouse gases, fractions of the land surface devoted to crop and pasture land and the direct effect of sulphate aerosols as an alteration of the surface albedo. I conducted simulations under the low (RCP3-PD), medium-low (RCP4.5), medium-high (RCP6.0) and high (RCP8.5) forcing scenarios. The results (Figure 5.1) suggest a strong acceleration in the rate of permafrost degradation beginning in the $21^{st}$ Century, with the potential for over 60 % loss of N.S. permafrost by year 2100 in the case of the high forcing scenario. The rate of simulated permafrost loss of the $21^{st}$ Century is comparable to other studies of permafrost degradation with global climate models under comparable

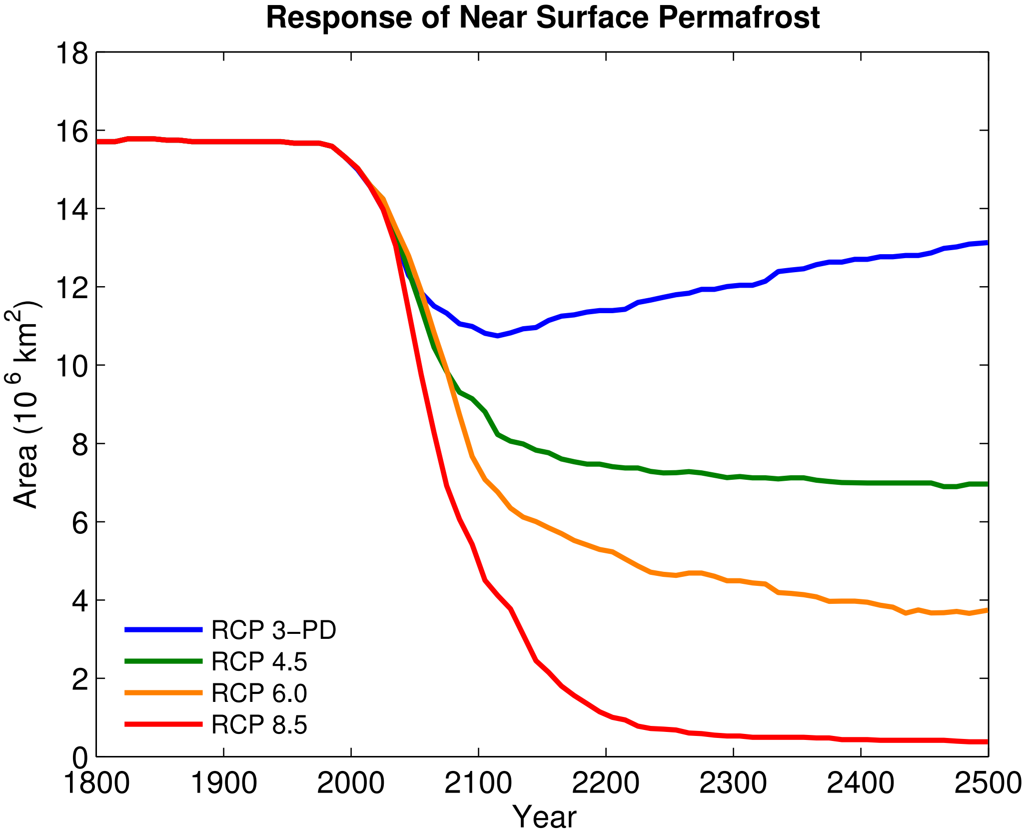


Figure 5.1: Time series of Northern Hemisphere near surface permafrost area simulated by the UVic ESCM under four Representative Concentration Pathway forcing scenarios.

forcing scenarios (Demchenko et al., 2001; Lawrence et al., 2008a; Saito et al., 2007).

## 5.3 Wetland Identification Scheme

I identify wetlands on the basis of terrain slope and soil moisture content (Kaplan, 2002) (Figure 5.2a). This use of model grid-scale climate indices and topographic thresholds is consistent with the approach used by other studies that identify wetlands within global climate models (Kaplan, 2002; Shindell et al., 2004; Weber et al., 2010). Grid cells containing wetlands are those having unfrozen soil moisture contents greater than 65 % of the saturated soil moisture content in the upper soil layer at least one day out of the year. Fractions of such grid cells containing wetlands are those with terrain slopes less than 0.20 % as identified from the ETOPO 2” Digital Elevation Map (U.S. Department of Commerce and Atmospheric Administration, 2006). Slope and moisture parameters were determined empirically by varying the parameter values, comparing the model output against observed global wetland maps, and selecting values that optimized the agreement between the model and observations. The simulated present-day wetland distribution generally agrees well with observed global wetlands datasets (Lehner and Doll, 2004; Matthews and Fung, 1987) (Figure 5.2b). Relative to these datasets, the model appears to slightly overestimate equatorial wetlands and underestimate wetlands around 20-40° N and 20-40° S, though many of the latter discrepancies are associated with quite small-scale features. It should also be noted, however, that the total global wetland area indicated by such datasets (and hence the model) is likely an underestimate (Finalyson and D'Cruz, 2005).

Wetlands are diverse ecosystems and are often identified, in part, by the presence of distinct wetland vegetation types (Wheeler, 1999), which are not represented as unique plant functional types in most climate models (including the UVic ESCM). Furthermore, many wetland ecosystems such as peatlands are characterized by complicated interactions between soil carbon, moisture and vegetation and feature a high degree of fine-scale surface heterogeneity, making the realistic simulation of these ecosystems challenging within the framework of a global-scale model (Frolking et al., 2009). While my approach uses a relatively simple scheme to identify wetlands, it is compatible with the level of complexity of the model's land surface scheme. Overall, the scheme succeeds in capturing the broad, global-scale distribution of wetlands. I focus on changes in wetlands associated with permafrost thaw in order to develop an understanding of how permafrost degradation will influence near surface hydrology on large scales.

A recent survey of the first-order controls of the abundance of Arctic lakes found that the presence of some form of permafrost increased the lake densities and area fractions by approximately 100-170 % relative to permafrost-free terrain (Smith et al., 2007). I suggest that these may generalize to wetlands; inspection of the relative abundance of high-latitude wetlands (Figure 5.3) and the distribution of permafrost free and permafrost bearing terrain (Figure 5.4) suggests that the occurrence of some form of permafrost strongly increases the abundance of wetlands, at least at pan-Arctic scales. The presence of permafrost is a key factor in determining the distribution of high-latitude wetlands in the model; in the simulated present day climate (averaged over the decade 2000-2010) 62 % of cells containing wetlands north of 45° N also con-

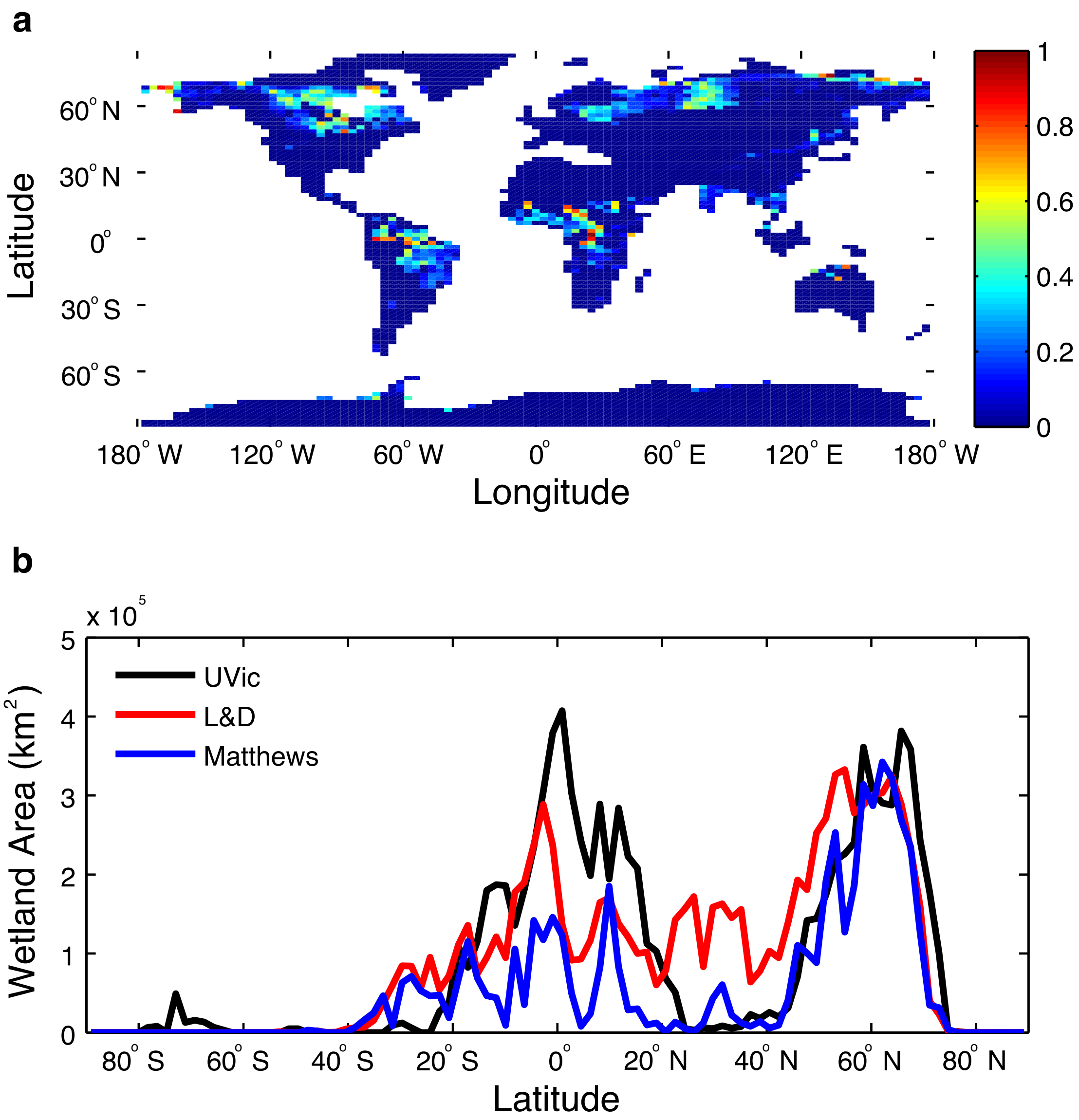


Figure 5.2: The UVic ESCM's present-day distribution of wetlands as compared with observations. a) Fractions of grid cells containing present-day (2000-2010) wetlands in the UVic model. b) Latitudinal distribution of wetland area. The model is able to reproduce the key features of the global wetland distribution when compared against the wetland databases of Lehner and Doll (2004) (L&D) and Matthews and Fung (1987) (M&F).

tain permafrost while 86 % of cells with wetlands north of 60 ° N contain permafrost. The mechanism underlying this correlation is clear: in the present-day climate, north of 45° N, the modelled permafrost table occurs no deeper than 1.7 m from the surface, and is generally substantially shallower than this. Below the permafrost table, frozen layers act as semi-impermeable barriers that greatly restrict the drainage of moisture from the upper soil layers. Thus, when these soil layers seasonally thaw, the liquid water tends to remain stored near the surface, resulting in seasonal wetlands.

## 5.4 Results

I simulate a total wetlands area north of 45° N to be 3.72 x $10^6$ km$^2$ in the present day climate and 3.76 x $10^6$ km$^2$, 3.18 x $10^6$ km$^2$, 2.79 x $10^6$ km$^2$, 1.11 x $10^6$ km$^2$ over the period 2490-2500 in the RCP 3-PD, RCP 4.5, RCP 6.0 and RCP 8.5 scenarios, respectively. I find that the rate of loss of wetlands is greatest when the rate of permafrost degradation is highest (Figure 5.5, left column). As warming occurs, there is an initial increase in wetland abundance at very high latitudes (> 70° N). This is a consequence of enhanced high-latitude precipitation, which results in increased soil moisture storage such that high latitude cells initially devoid of wetlands become able to satisfy the wetland soil moisture criterion during the thaw season in the now warmer climate. However, the increased abundance of wetlands at high latitudes is countered by a loss of wetlands associated with permafrost degradation such that there is a net loss in wetlands extent in the three higher emission scenarios. Wetland loss tends to occur first at low latitudes and shifts northward as high latitude warming progresses. I also show changes in the number of days out of the year that grid cells containing wetlands satisfy the moisture criterion (number of "days wet"). During the time period when wetland/N.S.permafrost loss is most pronounced at a particular latitude, there is an initial and strong increase in the average number of days wet for wetlands in regions of remaining N.S. permafrost (Figure 5.5, right column). This is subsequently followed by an abrupt decrease in number of days wet.

To better understand this observation, I selected a representative model grid cell containing both permafrost and seasonal wetlands in the present-day climate and tracked changes in the distribution of unfrozen moisture in the cells subsurface layers (Figure 5.6a-e). The number of days wet initially increases associated with an increase in length of the thaw season (Figure 5.6a vs. Figure 5.6b). As this occurs, the active layer and permafrost table deepen and the hydraulic conductivity (a function

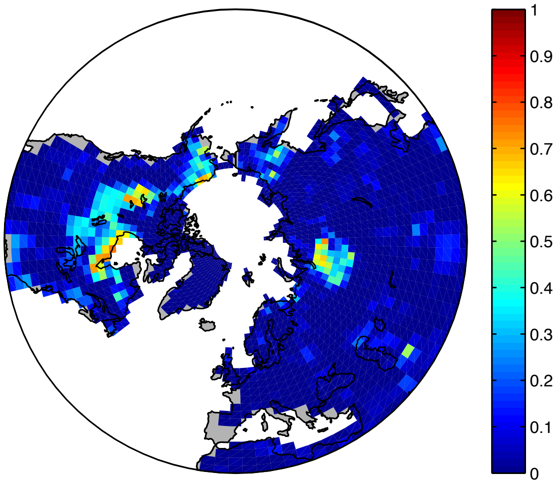


Figure 5.3: Polar projection plot of the wetlands distribution of Lehner and Doll (2004), as interpolated to the UVic ESCM model grid. I argue that a comparison of this plot with Figure 5.4 indicates that the presence of permafrost strongly increases the abundance of wetlands relative to permafrost-free terrain in the pan-Arctic region.

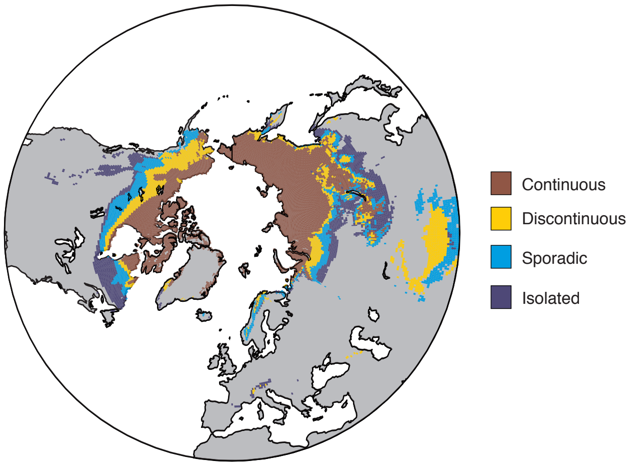


Figure 5.4: Zones of continuous, discontinuous, sporadic and isolated permafrost from the International Permafrost Association (IPA) digital permafrost map (Brown et al. (1998)).

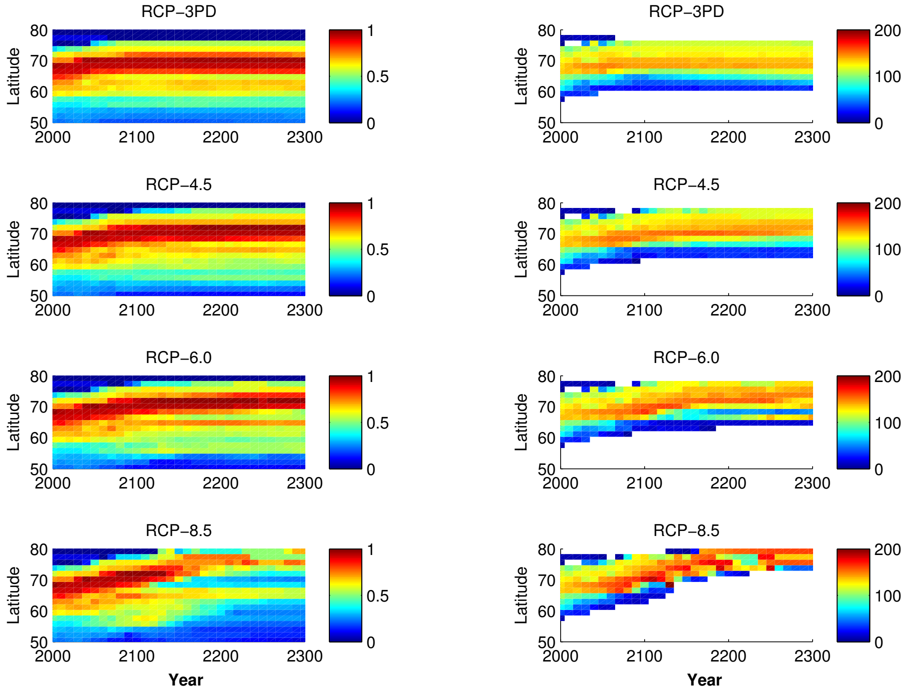


Figure 5.5: Hovmoller diagrams showing the evolution of northern wetlands under four RCP forcing scenarios. Wetland changes are generally most pronounced when rates of permafrost loss are highest. Left column: fraction of the total number of land grid cells at a given latitude that contain wetlands. Right column: average number of days out of the year that cells containing N.S. permafrost and wetlands satisfy the wetland moisture criterion. White indicates that there are no cells containing permafrost and wetlands at a particular latitude and a particular time.

of a layer's unfrozen moisture content) of the deeper layers increases. When the permafrost table deepens beyond about a meter, there is significant drainage of near surface moisture to deeper soil layers (Figure 5.6c) and the number of days that the upper soil layer satisfies the moisture criterion begins to sharply decline. The grid cell eventually transitions to a state where N.S. permafrost is absent, where the cell satisfies the moisture criterion only briefly at the outset of the thaw season (Figure 5.6d). An annually-averaged soil temperature of around 2.5 °C represents the point at which this significant deepening of the permafrost table typically occurs (Figure 5.7), and cells which initially contained permafrost and sustained seasonal wetlands transition to a state where cells satisfy the wetland moisture criterion at a greatly reduced number of days, if at all.

## 5.5 Discussion

The model's wetlands identification scheme is relatively simple since I do not explicitly simulate variations in water table depth or lateral water flow. In addition, the model lacks a representation of sub-gridscale variations in soil and vegetation characteristics and its representation of topography (including small-scale features) is static so that the formation of thermokarst terrain as permafrost thaws is not simulated. As well, some key boreal wetlands lie on the floodplains of major northern rivers and it is possible that their hydrology may be more strongly controlled by transient inundation events such as ice-damming, rather than permafrost, although such inundation events might also be expected to decrease in frequency in a warmer climate. All of these unresolved features and processes likely control the distribution and future evolution of high latitude wetlands to some degree. Despite these caveats, the model's present-day distribution of high latitude wetlands agrees with observations remarkably well, giving me confidence in its ability to capture the large-scale processes involved in their formation and persistence.

Accounting for 20-40 % of total emissions, the single biggest natural source of atmospheric methane arises from anaerobic decomposition of organic matter by methanogenic bacteria in naturally-occurring wetlands (Denman et al., 2007). The rate of methane production also has a strong dependence on temperature, with warmer conditions being associated with greater production. While northern wetlands are secondary to tropical wetlands as methane sources in the present climate (Denman et al., 2007), there has been much discussion as to a potential increase in methane emissions from

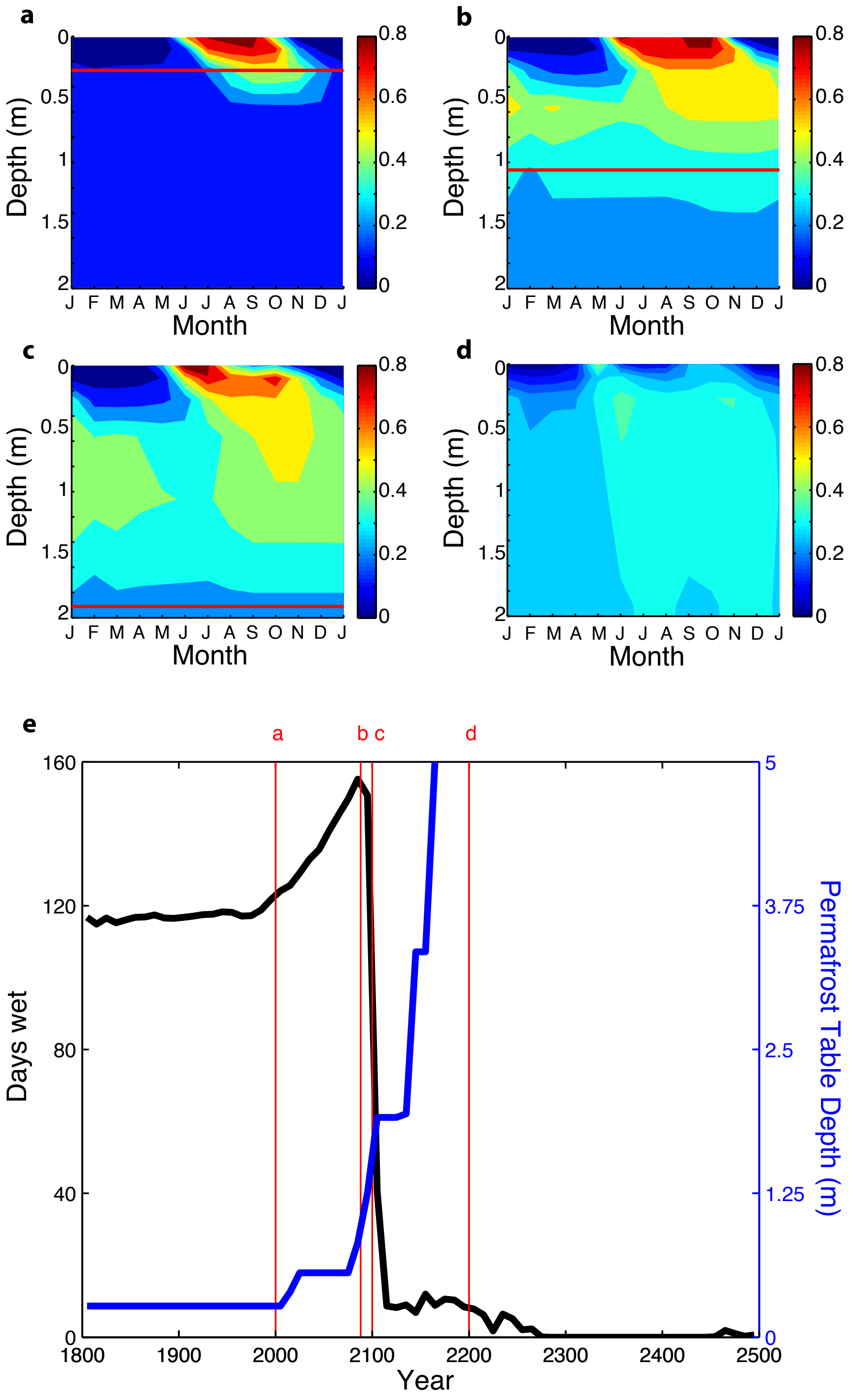


Figure 5.6: Evolution of near-surface soil moisture in a representative grid cell following the RCP8.5 scenario. In panels a-d, I plot the volumetric concentration of soil moisture ($m^3/m^3$) as the permafrost table (red line) deepens. In e), I plot the number of days of the year that the gridcell satisfies the moisture criterion ("days wet") and the permafrost table depth. Days wet initially increase due to a longer thaw season (a, b), but as warming continues and the hydraulic conductivity of deep layers increases, drainage of near surface moisture to deeper layers causes the days wet to sharply plummet (c, d).

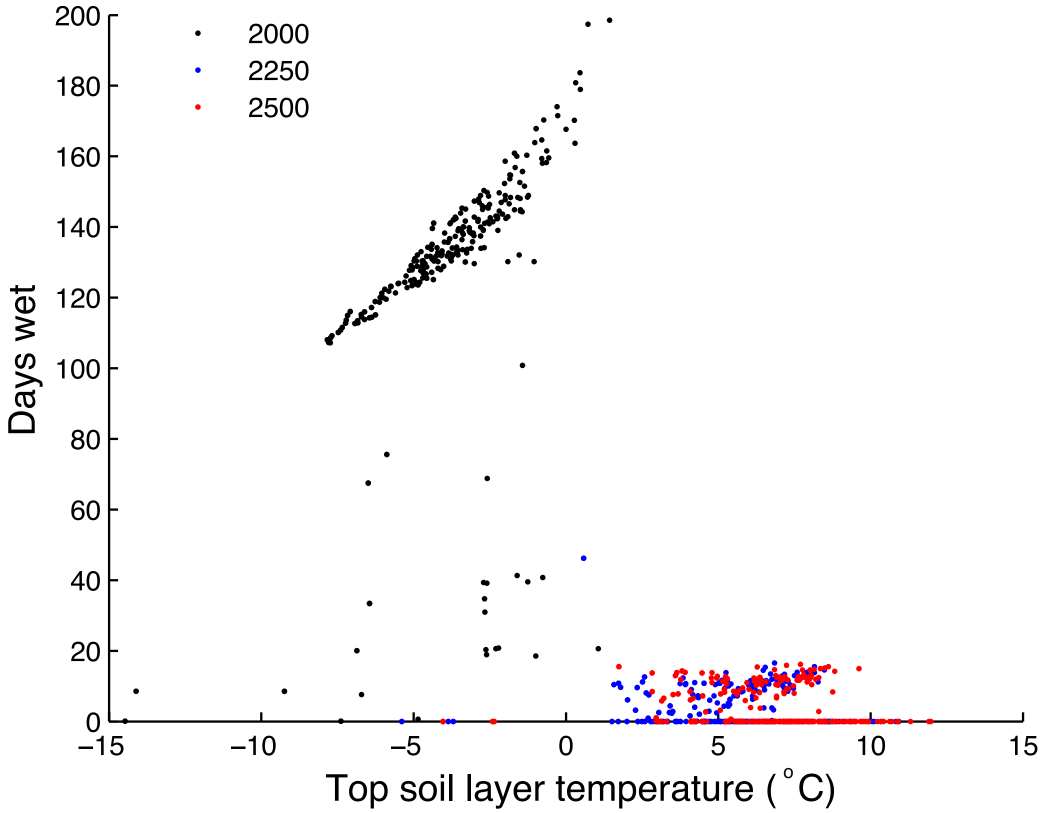


Figure 5.7: The evolution of wetlands in permafrost regions under the RCP8.5 forcing scenario. I plot the number of days of the year that grid cells satisfy the wetlands moisture criteria against the annually averaged top layer soil temperature. These data represent cells having a permafrost table depth of less than 0.5 m in year 2000 to illustrate how the deepening of the permafrost table exerts a control on wetlands. An annually average upper level soil temperature of about 2.5 °C represents the point at which there is significant deepening of the permafrost table and drainage to deeper soil layers.

northern wetlands under strong projected polar warming. Permafrost regions contain one of the largest vulnerable carbon pools with an estimated 1672 Pg of soil carbon in the northern circumpolar permafrost zone with 1024 Pg of the carbon within 3 m of the surface (Schuur et al., 2008; Tarnocai et al., 2009). As permafrost thaws, enhanced decomposition of soil carbon will alter carbon fluxes to the atmosphere. Wetlands play an important role in determining the magnitude of this high-latitude positive feedback to global warming. The presence of anoxic conditions in wetland soils leads to the release into the atmosphere of methane along with carbon dioxide during the decomposition process, with the former having a significantly higher global warming potential (Forster et al., 2007). My results suggest that the future global production of methane from high-latitude wetlands will have a complicated signal over the next century with decreasing overall wetland extent counteracting a greater number of wet days in remaining wetlands.

## 5.6 Summary

I this chapter, I used version U3.0M of the UVic Earth System Climate Model to demonstrate a link between high-latitude northern wetlands and the state of permafrost. I used a relatively simple wetland identification scheme that identifies wetland area on the basis of upper layer unfrozen soil moisture content and terrain slope. Despite, the relative simplicity of this scheme, the model is able to achieve a good reproduction of the global extent of wetlands. I showed that as near surface permafrost degrades, the areal extent of wetlands decreases, owing to an increase in the hydraulic conductivity of soil that allows for enhanced drainage of near surface soil moisture. I simulated a total wetlands area north of 45° N to be 3.72 x $10^6$ km$^2$ in the present day climate and 3.76 x $10^6$ km$^2$, 3.18 x $10^6$ km$^2$, 2.79 x $10^6$ km$^2$, 1.11 x $10^6$ km$^2$ over the period 2490-2500 in the RCP 3-PD, RCP 4.5, RCP 6.0 and RCP 8.5 scenarios, respectively.

The number of days that seasonal wetlands are present in regions with remaining permafrost initially increases through a lengthening of the thaw season, but dramatically decreases as the permafrost table deepens. I identify an annually averaged soil temperature threshold of 2.5 °C above which this transition to wetland reduction occurs. I suggest that the release of methane from high-latitude regions may have a complicated signal over the next century due to changes in wetland characteristics.

# Chapter 6

# Conclusion

## 6.1 Summary of Key Findings

In this dissertation, I have reported on the present-day and future characteristics of the Arctic climate focussing specifically on regions of perennially frozen ground, or permafrost. The research tool for these simulations was the UVic Earth System Climate Model - an intermediate complexity climate model whose land surface scheme was modified to enable it to simulate regions of seasonally frozen ground and permafrost. A number of major modifications were made to the model as outlined in Chapter 2 and Appendix B, of which the most significant were the replacement of the original single subsurface soil layer with a multiple layer soil and bedrock scheme as well as a representation of the thermal and hydrological impacts of soil freezing and thawing. I believe that the research presented in this thesis represents a significant step forward both in the development of the UVic ESCM as well as in the use of climate models to better understand the controls on, and the characteristics of permafrost.

In Chapter 3, I compared the simulated Arctic climate of the ESCM to the ERA-40 reanalysis. The model displays some regional temperature anomalies that were attributed to the degree of sophistication of the atmospheric module as well as to known anomalies in the simulated extent of sea ice. While the model does a reasonable job of capturing the annual mean Arctic precipitation, its seasonality is poorly captured, possibly due to the inability of the model to represent convective precipitation. Despite these biases, the fully coupled model (U3.0M) capably captures the extent and thickness of lying snow. A second version of the model, driven by data derived from the ERA-40 reanalysis (U3.0E), significantly under-represents the thickness of lying snow. This observation suggests that the ablation of snow may be overestimated in the new version of the ESCM, and that the fully coupled version of the model satisfactorily simulates the distribution of snow on account of excessive Arctic winter precipitation. A more thorough investigation of the characteristics of the model's snow scheme may be justified if the version of the model described in this thesis is integrated into a future version of the operational climate model.

A number of metrics were used to assess the distribution of permafrost and seasonally frozen ground in the ESCM. U3.0M simulates a present-day permafrost area of 15.59 x $10^6$ km$^2$ which must be interpreted as an overestimate on the order of $\sim$ 10 % on account of the fact that the ESCM overestimates Northern Hemisphere land mass due to its model grid configuration. This simulated area agrees well with the areal extent of zones of continuous and discontinuous permafrost (14.98 x $10^6$ km$^2$) as well as estimates of the total area of ground containing permafrost (12.21 - 16.98 x $10^6$ km$^2$) (Zhang et al. (2000)) although there are some regions near the southern permafrost boundary where permafrost is over- and under- abundant compared with observations and these generally coincide with identified atmospheric warm and cold biases. The simulated extent of seasonally frozen ground is 47.92 x $10^6$ km$^2$, a slight underestimate of the estimated extent of 48.12 x $10^6$ km$^2$ (Zhang et al. (2003)) likely due to the impact of atmospheric warm biases in the southern limits of the seasonally frozen ground region as well as the model's limited vertical resolution which means that the freezing of very thin soil layers is not captured. Ground temperatures in permafrost in U3.0M agree well with observations and the model does a good job of representing the transition zone between thin and thick permafrost. Modelled active layer thicknesses are consistently shallow compared with observations; results from a sensitivity analysis conducted indicates that this may be largely attributable to the model's limited vertical resolution. U3.0E does a better job of simulating regions of seasonally frozen ground than U3.0M, but the simulated permafrost area is overestimated and ground temperatures show a clear cold bias relative to observations; both of these latter observations may be explained by the fact that snow thicknesses are underestimated in this model configuration.

The results from Chapter 3 suggest that the fully coupled version of the climate model is a good tool for understanding the characteristics and temporal response of permafrost at pan-Arctic scales, with the stipulation that the permafrost characteristics simulated by the model are representative of mean values over large geographic areas. The high degree of heterogeneity in surface cover and subsurface characteristics means that the UVic ESCM is unsuitable for simulating permafrost characteristics at any one particular site, unless detailed site-specific data is supplied.

Chapter 4 examined the transient response of the Arctic climate from year 1800, through the historical period and into the future using forcing fields from the four Representative Concentration Pathways. The UVic ESCM's global mean surface air temperature response following RCP scenarios agrees well with the IPCC ensemble mean response following comparable SRES scenarios, although the ESCM's land/water warming contrast is lower than the ensemble mean and the model's low polar amplification means that Arctic warming is less pronounced relative to global mean warming that the IPCC ensemble mean. The model's low polar amplification potentially means that the transient response of permafrost and seasonally frozen ground is slower than it would otherwise be. The ESCM also simulates an overall increase in global precipitation as well as an increase in annual-mean high-latitude precipitation. The seasonality of the precipitation response, however, does not agree well with that shown by the IPCC models, with the ESCM showing a decrease in precipitation in the summer months when the IPCC models indicate an increase in precipitation; this is potentially associated with the same issue with the model that prevents it from adequately capturing the seasonal cycle in Arctic precipitation in the present day climate. These climate changes generally result in a greening of the Arctic in the ESCM with northward expansion of shrubs into grasslands and expansion of needleleaf trees into shrublands.

Like results from other models, the ESCM shows only modest decreases in the extent of permafrost over the historical period, followed by a rapid acceleration in the rate of loss beginning in the early $21^{st}$ Century, coinciding with an equivalent rapid rise in Arctic and global surface air temperature change. Permafrost degradation proceeds similarly along the four RCPs early in the mid-$21^{st}$ Century, but begins to diverge mid-Century. By the end of the century, the total area of permafrost (the area containing permafrost at any depth) decreased by between 28.51 % (RCP3-PD) and 53.82 % (RCP8.5) relative to its extent during the last two decades of the $20^{th}$ Century. The rate of permafrost loss in the UVic ESCM lies in the mid-range of rates of permafrost degradation from other models following comparable forcing scenarios. Near surface permafrost shows a significantly faster rate of degradation such that its area has decreased by between 28.87 % (RCP3-PD) and 65.91 % (RCP8.5) by 2100. Loss of permafrost begins in the southern limits of permafrost; thin permafrost degrades strictly by a deepening of the active layer whereas in thicker permafrost this processes is followed by the development and thickening of an perennially unfrozen talik layer. Radiative forcing in RCP3-PD peaks in 2050 and begins to decline thereafter and permafrost begins to stabilize early in the $22^{nd}$ Century in this scenario. Continued loss of permafrost occurs in the three higher RCP scenarios for the duration of the simulation. By 2500, as much as 81.05 % of total permafrost and 97.54 % of near surface permafrost has been lost, following the RCP8.5 scenario. These results point to the importance of considering long term changes in the Arctic climate and not simply restricting simulations to the $21^{st}$ Century. As well, results point to the importance of clearly specifying the depth of permafrost under consideration as near surface and deep permafrost layers respond and markedly different rates to warming.

In Chapters 3 and 4, I also conducted a set of experiments to test the sensitivity of the equilibrium permafrost distribution and its transient response following the RCP forcings. The most significant findings from these experiments are as follows. Firstly, the simulated extent of areas of permafrost and seasonally frozen ground was comparable in the standard configuration of the model to configurations with higher vertical resolution. The most significant difference between the different resolution configurations was an improved representation of active layer thickness as resolution is increased, resulting in a reduction in the model's shallow active layer bias. I suggest that the extra computational cost of increased vertical resolution is not essential unless the model is used in an application where accurate representation of active layer thickness is essential.

Secondly, of the trials conducted, the model was shown to be most sensitive to changes in the depth to bedrock. When the depth to bedrock is shallow, the active layer is deeper and permafrost warms more rapidly, both as a consequence of a reduction in total column soil moisture storage and accompanying latent heat change required to thaw ground ice. Consequently, the equilibrium distribution of near surface permafrost is considerably reduced in the shallow bedrock configurations, and the near surface permafrost that is present decays at an enhanced rate relative to deeper bedrock configurations. Somewhat counterintuitively, the shallow bedrock configuration shows a greater thickness of permafrost owing to a reduction in the thermal offset (warming) of near surface layers and a reduced depth at which the ground thermal conductivity switches from soil to bedrock; both factors result in an increase in the depth of the permafrost base in the shallow bedrock configurations.

Finally, the inclusion of excess ground ice results in a slightly slower rate of permafrost degradation, most noticeably in the case of the RCP3-PD and 4.5 scenarios, but allows for an estimate of the extent of ground subsidence accompanying permafrost loss. The southern limits of permafrost are most vulnerable to ground subsidence with subsidence ranging between 1 to 2.5 m by the end of the $21^{st}$ Century and with subsidence as much as 4 m by year 2500, again pointing to the importance of long term simulations when assessing hazards in regions of permafrost.

In Chapter 5, I considered the response of high-latitude wetlands to the RCP scenarios. I used a relatively simple wetland identification scheme that identifies wetland area on the basis of upper layer unfrozen soil moisture content and terrain slope. The model is able to achieve a good reproduction of the global extent of wetlands in the present day climate. The model suggests that high-latitude wetlands may be very sensitive to the state of permafrost. I showed that as near surface permafrost degrades, the areal extent of wetlands decreases, owing to an increase in the hydraulic conductivity of soil that allows for enhanced drainage of near surface soil moisture. I simulated a total wetlands area north of 45° N to be 3.72 x $10^6$ $km^2$ in the present day climate and 3.76 x $10^6$ $km^2$, 3.18 x $10^6$ $km^2$, 2.79 x $10^6$ $km^2$, 1.11 x $10^6$ $km^2$ over the period 2490-2500 in the RCP 3-PD, RCP 4.5, RCP 6.0 and RCP 8.5 scenarios, respectively. The loss of wetlands occurs against the background of an overall increase in Arctic precipitation. The model's wetlands identifications scheme is relatively simple and does not account for the full range of processes responsible for the sustenance of wetlands. Given the potential for major changes to these sensitive ecosystems, I suggest that this situation merits further study using models with a more sophisticated treatment of ground hydrology.

As a side note, I have also recently been involved in submitting wetland simulations to the Wetland and Wetland $CH_4$ Intercomparison of Models Project (WETCHIMP). The ESCM is one of ten models involved in this project which aims to coordinate a systematic study of models that represent wetlands and methane emissions over large spatial scales. While the ESCM does not currently include a representation of the methane cycle, it is one of only two models in the project that prognostically simulates wetland extent rather than relying on datasets that specify inundation area and, as such provides a unique modelling perspective to this intercomparison. The intercomparison is presently in the analysis and discussion phase and it is anticipated that results will be submitted for publication later this year.

## 6.2 Directions for Future Research

While I have investigated a number of environmental impacts of Arctic warming including vegetation change, permafrost loss and changes in wetlands, I have not investigated changes in the high-latitude carbon cycle and this would be the next logical step now that the UVic ESCM is capable of representing permafrost. One of the most important and poorly quantified impacts of permafrost degradation is the positive permafrost carbon feedback resulting from heightened microbial respiration of carbon in permafrost soils. At present, microbial decomposition of carbon by microbes is generally restricted to the small amounts of unfrozen moisture that occur in soils below 0 °C and the imbalance between photosynthesis and microbial decomposition mean that permafrost regions contain one of the largest vulnerable carbon pools. It is conservatively estimated that 1672 Pg of soil carbon is present in the northern circumpolar permafrost zone, with 1024 Pg C within 3 meters of the surface (Schuur et al., 2008). Loss of carbon from the permafrost pool will lead to a positive feedback on warming, influencing both the Arctic and global climate. The magnitude and timing of the feedback are key uncertainties that urgently need to be quantified. Initial research into the permafrost carbon feedback conducted by uncoupled ecosystem models (Koven et al., 2011; Schaefer et al., 2011) indicates that an additional 55 - 138 petagrams of carbon may be released to the atmosphere from permafrost soils by 2100, though this is likely an underestimate as such uncoupled models cannot account for the additional warming induced by the release of such carbon.

As demonstrated in Chapter 2, version 3.0 of the UVic ESCM underestimates the amount of carbon stored at high latitudes. High latitude carbon densities in version 3.0M are higher, indicating that inhibition of respiration by soil freezing is partly responsible for this observation. Yet, even in this version of the model, carbon densities are considerably lower than observed. A relatively simple approach to addressing the permafrost carbon feedback would be to simply specify carbon stores in permafrost. This would be straightforward to implement since each model soil layer contains its own carbon content and therefore permafrost carbon could simply be initialized and partitioned into carbon stores in soil layers beneath the active layer using the cumulative carbon distribution function as described in Chapter 2. Like excess ice, the model would not simulate the build up of this carbon but could account for its influence as permafrost thaw began to occur. The model would also be well suited to study the extent to which enhanced primary productivity and increased carbon uptake through the northwards expansion of trees and shrubs might offset the release of carbon from the soils. As well, a more detailed treatment of respiration of soil carbon within permafrost might also include some treatment of the effects of biogenic heat production accompanying soil respiration. A study by Khvorostyanov et al. (2008) suggests that the degree of heat released from decomposition is high enough to significantly impact the rate of permafrost thaw and may be sufficiently high in some case to sustain permafrost thaw even when the external forcing responsible for thaw is terminated.

Further work might be undertaken to improve the prognostic simulation of permafrost carbon in the model. At least two processes absent from the ESCM may account for the low high latitude carbon density. Firstly, Koven et al. (2009) demonstrated that the inclusion of a simple parameterization of cryoturbation (mixing processes that occur within cryotic soils) substantially increased soil carbon storage and improved the vertical distribution of carbon in their model. Koven et al. (2009)'s implementation of cryoturbation is based on a simple vertical diffusion scheme to distribute carbon within the soil column and would be relatively straightforward to incorporate in the ESCM. Secondly, the UVic model does not include a specific representation of peatland plant functional types. Northern peatlands play an important role in the high-latitude carbon cycle due to their high carbon sequestration rates and role as significant sources of methane (Wania, 2007) and inclusion of peatlands within the ESCM would likely also improve the simulated high-latitude soil carbon pool.

Ultimately, a simple representation of the methane cycle might be included in the model to fully account for the permafrost carbon feedback. Methane models available for inclusion in climate models cover a wide spectrum of complexity (Wania, 2007) ranging from simple empirical models that diagnose methane emissions based on only a few environmental controls to highly detailed process-based models. I believe the latter class of model is unsuitable for inclusion in the UVic ESCM as they have a level of computational complexity considerably higher than other elements of the ESCM's carbon cycle components and are incompatible with the level of sophistication of the model's subsurface hydrology scheme. A methane model based on that of Gedney et al. (2004), wherein methane emissions are diagnosed in terms of soil temperature and carbon density might be applied to account for methane emissions in fractions of model gridcells occupied by wetlands. Methane could then be treated in the atmospheric model of the ESCM as a well-mixed greenhouse gas that would decay into $CO_2$ based on a specified atmospheric residence time with its radiative effects handled as a reduction in outgoing long-wave radiative in a manner akin to that of $CO_2$ in the model. Implementation of even a simple representation of the methane cycle would allow quantification of the impact of large projected changes in northern wetlands on this aspect of the carbon cycle.

# Bibliography

ACIA, 2004: *Impacts of a Warming Arctic: Arctic Climate Impact Assessment*. Cambridge University Press, Cambridge, UK.

Ács, F., 2003: A comparative analysis of transpiration and bare soil evaporation. *Boundary-Layer Meteorology*, **109 (2)**, 139–162, doi:10.1023/A:1025473221779.

Alexeev, V., D. Nicolsky, V. Romanovsky, and D. Lawrence, 2007: An evaluation of deep soil configurations in the CLM3 for improved representation of permafrost. *Geophysical Research Letters*, **34 (L09502)**, doi:10.1029/2007GL029536.

Allison, I., et al., 2009: *The Copenhagen Diagnosis: Updating the World on the Latest Climate Science*. The University of New South Wales Climate Change Research Centre, Sydney, Australia.

Anisomov, O., N. I. Shiklomanov, and F. Nelson, 1997: Global warming and active-layer thickness: results from transient general circulation models. *Global and Planetary Change 15*, **15 (3-5)**, 61–77, doi:10.1016/S0921-8181(97)00009-X.

Anthony, K. M. W., P. Anthony, G. Grosse, and J. Chanton, 2010: Geologic methane seems along boundaries of Arctic permafrost thaw and melting glaciers. *Nature Geoscience*, doi:10.1038/ngeo1480.

Armstrong, R., M. Brodzik, K. Knowles, and M. Savoie, 2007: Global Monthly EASE-Grid Snow Water Equivalent Climatology. National Snow and Ice Data Center, Boulder, Colorado USA, Digital media.

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

Banks, D., 2008: *An Introduction to Thermogeology: Ground Source Heating and Cooling*. Wiley-Blackwell.

Bengtsson, L., K. Hodges, S. Koumoutsaris, M. Zahn, and N. Keenlyside, 2011: The changing atmospheric water cycle in Polar Regions in a warmer climate. *Tellus Series A - Dynamic Meteorology and Oceanography*, **63A (5)**, 907–920, doi: 10.1111/j.1600-0870.2011.00534.x.

Boelhouwers, J. and K. Hall, 2002: Periglacial and permafrost research in the Southern Hemisphere. *South African Journal of Science*, **98**, 46.

Bonan, G., 2008: *Ecological Climatology*. 2d ed., Cambridge University Press.

Brady, N. and R. Weil, 2001: *The Nature and Property of Soils*. Thirteenth ed., Prentice Hall.

Brook, E., D. Archer, E. Dlugokencky, S. Frolking, and L. D., 2008: Potential for abrupt changes in atmospheric methane. *Abrupt Climate Change*, A report by the U.S. Climate Change Science Program and the Subcommittee on Global Change Research. US Geological Survey, Reston, VA, 163–201.

Brown, J., O. Ferrians, J. Heginbottom, and E. Melnikov, 1998: Circum-Arctic map of permafrost and ground ice conditions. Digital media, National Snow and Ice Data Center / World Data Center for Glaciology, Boulder, CO.

Brown, J., K. 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.

Christensen, J., et al., 2007: Regional climate projections. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, M. Tignor, and H. Miller, Eds., Cambridge University Press, Cambridge, UK and New York, NY, USA, chap. 11, 847–940.

Clapp, R. and G. Hornberger, 1978: Empirical equations for some soil hydraulic properties. *Water Resources Research*, **14 (4)**, 601–604, doi: doi:10.1029/WR014i004p00601.

Clarke, L., J. Edmonds, H. Jacoby, H. Pitcher, J. Reilly, and R. Richels, 2007: *Scenarios of Greenhouse Gas Emissions and Atmospheric Concentrations*. Sub-report

2.1A of Synthesis and Assessment Product 2.1 by the U.S. Climate Change Science Program and the Subcommittee on Global Change Research. Department of Energy, Office of Biological & Environmental Research, Washington, D.C., USA.

Clifford, D., 2010: Global estimates of snow water equivalent from passive microwave instruments: history, challenges and future developments. *International Journal of Remote Sensing*, **31 (14)**, 3707–3726, doi:10.1080/01431161.2010.483482.

Cook, B., G. Bonan, S. Levis, and H. Epstein, 2007: The thermoinsulation effect of snow cover within a climate model. *Climate Dynamics*, **31 (1)**, 107–124, doi:10.1007/s00382-007-0341-y.

Cosby, B., G. Hornberger, R. Clapp, and T. Ginn, 1984: A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils. *Water resources research*, **20 (6)**, 682–690, doi:10.1029/WR020i006p00682.

Cox, P., R. Betts, C. Bunton, R. Essery, P. Rowntree, and J. Smith, 1999: The impact of new land surface physics on the GCM simulation of climate and climate sensitvity. *Climate Dynamics*, **15 (3)**, 183–203, doi:10.1007/s003820050276.

Dankers, R., E. Burke, and J. Price, 2011: Simulation of permafrost and seasonal thaw depth in the JULES land surface scheme. *The Cryosphere*, **5**, 1263–1309, doi:10.5194/tcd-5-1263-2011.

Davis, T., 2001: *Permafrost: A Guide to Frozen Ground in Transition*. University of Alaska Press, Fairbanks, Alaska.

Demchenko, P., A. Eliseev, I. Mokhov, A. Obukhov, V. Nechaev, and A. Velichko, 2001: Sensitivity of permafrost cover in the Northern Hemisphere to climate change. *Clivar Exchanges*, **6 (3)**, 9–11.

Denman, K., et al., 2007: Couplings between changes in the climate system and biogeochemistry. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, M. Tignor, and H. Miller, Eds., Cambridge University Press, Cambridge, UK and New York, NY, USA, chap. 7, 499–587.

Dharssi, I., P. Vidale, A. Verhoef, and B. Macpherson, 2009: New soil physical properties implemented in the Unified Model at PS18. Technical Report 528, Hadley Centre, Met Office.

Dunderdale, M., J. Muller, and P. Cox, 1999: Sensitivity of Hadley Centre climate model to different earth observation and cartographically derived land surface datasets. *The Contribution of POLDER and New Generation Spaceborne Sensors to Global Change Studies*, Meribel, France, 1–6.

Epstein, H., Q. Yu, J. Kaplan, and H. Lishke, 2007: Simulating future changes in arctic and subarctic vegetation. *Computing in Science and Engineering*, **9 (4)**, 12–23, doi:10.1109/MCSE.2007.84.

Farouki, O., 1981: Thermal properties of soils. Monograph 81-l, Cold Regions Research and Engineering Laboratory.

Finalyson, C. and R. D'Cruz, 2005: Inland water systems. *Ecosystems and Human Well-Being*, R. Hassan, R. Scholes, and N. Ash, Eds., Island Press, Washington, D.C., USA, Vol. 1: Current State and Trends, chap. 20, 551–583.

Forster, P., et al., 2007: Changes in atmospheric constituents and in radiative forcing. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, M. Tignor, and H. Miller, Eds., Cambridge University Press, Cambridge, UK and New York, NY, USA, chap. 2, 129–234.

French, H., 2007: *The Periglacial Environment.* 3d ed., Wiley, West Sussex, England.

Frolking, S., N. Roulet, and D. Lawrence, 2009: Issues related to incorporating northern peatlands into global climate models. *Carbon Cycling in Northern Peatlands*, A. Baird, L. Belyea, X. Comas, A. Reeve, and L. Slater, Eds., American Geophysical Union, Washington, D.C., USA, No. 184 in Geophysical Monograph.

Fujino, J., R. Nair, M. Kainuma, T. Masui, and Y. Matsuoka, 2006: Multi-gas mitigation analysis on stabilization scenarios using AIM global model. *Mutligas Mitigation and Climate Policy*, **The Energy Journal Special Issue (3)**, 343–354.

Gedney, N., P. Cox, and C. Huntingford, 2004: Climate feedback from wetland methane emissions. *Geophysical Research Letters*, **31 (L20503)**, doi: 10.1029/2004GL020919.

Goodrich, L., 1978: Some results of a numerical study of ground thermal regimes. *Proceedings of the 3rd International Conference on Permafrost*, National Research Council of Canada, Ottawa, Canada, 29–34.

Gravesen, R. and M. Wange, 2009: Polar amplification in a coupled climate model with locked albedo. *Climate Dynamics*, **33 (5)**, 629–643, doi:10.1007/s00382-009-0535-6.

Greco, S., R. Moss, D. Viner, and R. Jenne, 1994: Climate scenarios and socioeconomic projections for IPCC WG II assessment. Tech. rep., Consortium for International Earth Science Information Network.

Halsey, L., D. Vitt, and S. Zoltai, 1995: Disequilibrium response of permafrost in boreal continental western Canada to climate change. *Climatic Change*, **30 (1)**, 57–73, doi:10.1007/BF01093225.

Hamza, V., R. Cardoso, and C. Ponte Neto, 2008: Spherical harmonic analysis of Earth's conductive heat flow. *International Journal of Earth Sciences*, **97 (2)**, 205–226, doi:10.1007/s00531-007-0254-3.

Harris, S., H. French, J. Heginbottom, G. Johnston, B. Ladanyi, D. Sego, and R. van Everdinger, 1988: Glossary of permafrost and related ground-ice terms. Technical Memorandum 142, National Research Council of Canada.

Heginbottom, J., 2002: Permafrost mapping: a review. *Progress in Physical Geography*, **26 (4)**, 623–642, doi:10.1191/0309133302pp355ra.

Heginbottom, J., J. Brown, E. Melnikov, and O. Ferrians Jr., 1993: Circumarctic map of permafrost and ground ice conditions. *Proceedings of the 6th International Conference on Pemafrost*, South China University of Technology Press, Beijing, China, Lanzhou Institute of Glaciology and Geocryology, Chinese Academy of Sciences, Chinese Society of Glaciology and Geocryology, Vol. 2, 1132–1136.

Heginbottom, J. and M. Dubreuil, 1993: A new permafrost and ground ice map for the National Atlas of Canada. *Proceedings of the 6th International Conference on*

*Pemafrost*, South China University of Technology Press, Beijing, China, Lanzhou Institute of Glaciology and Geocryology, Chinese Academy of Sciences, Chinese Society of Glaciology and Geocryology, Vol. 1, 255–260.

Hinkel, K. and F. Nelson, 2003: Spatial and temporal patterns of active layer thickness at Circumpolar Active Layer Monitoring (CALM) sites in northern Alaska, 1995-2000. *Journal of Geophysical Research*, **108 (8168)**, doi:10.1029/2001JD000927.

Hinzman, L., N. Bettez, W. Bolton, F. Chapin, M. Dyourgerov, and C. Fastier, 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.

Hinzman, L., D. Kane, K. Yoshikawa, A. Carr, W. Bolton, and M. Fraver, 2003: Hydrological variations among watersheds with varying degrees of permafrost. *Proceedings of the 8th International Conference on Permafrost*, M. Phillips, S. Springman, and L. Aresnson, Eds., A.A. Balkema Publishers, Lisse, The Netherlands, 407–411.

Holland, M., C. Bitz, and A. Weaver, 2001: The influence of sea ice physics on simulations of climate change. *Journal of Geophysical Research*, **106 (C9)**, doi: 10.1029/2000JC000651.

Hornberger, G., J. Raffensperger, P. Wiberg, and K. Eshleman, 1998: *Elements of Physical Hydrology.* John Hopkins University Press, Baltimore, MA, USA.

Hurtt, G., et al., 2009: Harmonization of global land-use scenarios for the period 1500-2100 for IPCC-AR5. *iLEAPS Newsletter*, **7**, 6–8.

IPCC, 2007: *Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge University Press.

IPCC, 2007: Summary for policymakers. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, M. Tignor, and H. Miller, Eds., Cambridge University Press, Cambridge, UK and New York, NY, USA, 1–18.

Johansen, O., 1975: Thermal conductivity of soils. Ph.D. thesis, University of Trondheim, Trondheim, Norway.

Kaplan, J., 2002: Wetlands at the last glacial maximum: Distribution and methane emission. *Geophysical Research Letters*, **29 (1079)**, doi:10.1029/2001GL013366.

Kaplan, J. and M. New, 2006: Arctic climate change with a 2 °C global warming: timing, climate patterns and vegetation change. *Climatic Change*, **79 (3-4)**, 213–241, doi:10.1007/s10584-006-9113-7.

Kattsov, V. and E. Kallen, 2005: Future climate change: Modeling and scenarios for the Arctic. *Arctic Climate Impact Assessment*, P. Grabhorn, J. Weybright, and C. Grabhorn, Eds., Cambridge University Press, Cambridge, UK and New York, NY, USA, chap. 4, 100–150.

Khvorostyanov, D. V., G. Krinner, P. Ciais, M. Heimann, and S. Zimov, 2008: Vulnerability of permafrost carbon to global warming. Part I: model description and role of heat generated by organic matter decomposition. *Tellus B*, **60 (2)**, 250–264.

Koven, C., P. Friedlingstein, P. Ciais, D. Khvorostyanov, G. Krinner, and C. Tarnocai, 2009: On the formation of high-latitude soil carbon stocks: effects of cryoturbation and insulation by organic matter in a land surface model. *Geophysical Research Letters*, **36 (L21501)**, doi:10.1029/2009GL040150.

Koven, C., B. Ringeval, P. Friedlingstein, P. Ciais, P. Cadule, D. Khvorostyanov, G. Krinner, and C. Tarnocai, 2011: Permafrost carbon-climate feedbacks accelerate global warming. *Proceedings of the National Academy of Sciences*, **108 (36)**, 14 769–14 774, doi:10.1073/pnas.1103910108.

Kwok, R., G. Cunningham, M. Wensnahan, I. Rigor, H. Zwally, and D. Yi, 2009: Thinning and volume loss of the Arctic Ocean sea ice cover: 2003-2008. *Journal of Geophysical Research*, **114 (C07005)**, doi:10.1029/2009JC005312.

Lamarque, J.-F., et al., 2010: Historical (1850-2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: methodology and application. *Atmospheric Chemistry and Physics*, **10**, 7017–7039, doi:10.5194/acp-10-7017-2010.

Lawrence, D. and A. Slater, 2005: A projection of severe near-surface permafrost degradation during the 21st century. *Geophysical Research Letters*, **32 (L24401)**, doi:doi:10.1029/2005GL025080.

Lawrence, D. and A. Slater, 2008: Incorporating organic soil into a global climate model. *Climate Dynamics*, **30 (2-3)**, 145–160, doi:10.1007/s00382-007-0278-1.

Lawrence, D., A. Slater, V. Romanovsky, and D. Nicolsky, 2008a: Sensitivity of a model projection of near-surface permafrost degradation to soil column depth and representation of soil organic matter. *Journal of Geophysical Research*, **113 (F02011)**, doi:10.1029/2007JF000883.

Lawrence, D., A. Slater, and S. Swenson, 2011: Simulation of present-day and future permafrost and seasonally frozen ground in CCSM4. *Journal of Climate*, doi: 10.1175/JCLI-D-11-00334.1.

Lawrence, D., A. Slater, R. Tomas, M. Holland, and C. Deser, 2008b: Accelerated Arctic land warming and permafrost degradation during rapid sea ice loss. *Geophysical Research Letters*, **35 (L11506)**, doi:10.1029/2008GL033985.

Lehner, B. and P. Doll, 2004: Development and validation of a global dataset of lakes, reservoirs and wetlands. *Journal of Hydrology*, **296 (1-4)**, 1–22, doi: 10.1016/j.jhydrol.2004.03.028.

Lemke, P., et al., 2007: Observations: Changes in snow, ice and frozen ground. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, M. Tignor, and H. Miller, Eds., Cambridge University Press, Cambridge, UK and New York, NY, USA, chap. 4, 337–383.

Linell, K., 1973: Long term effects of vegetative cover on permafrost stability in an area of discontinuous permafrost. *North American Contribution, 2nd International Conference on Permafrost*, National Academy of Sciences, Washington, D.C., USA, 688–693.

Loveland, T. and A. Belward, 1997: The IGBP-DIS global 1 km land cover dataset, DIScover: first results. *International Journal of Remote Sensing*, **18 (15)**, 3289–3295, doi:10.1080/014311697217099.

Lunardini, V., 1981: *Heat Transfer in Cold Climates*. Van Nostrand Reinhold Company, New York, NY, USA.

Matthews, E. and I. Fung, 1987: Methane emissions from natural wetlands: Global distribution, area and environmental characteristics of sources. *Global Biogeochemical Cycles*, **1 (1)**, 61–86, doi:10.1029/GB001i001p00061.

Matthews, H., 2004: Land cover change, vegetation dynamics and the global carbon cycle: Experiments with the UVic Earth System Climate Model. Ph.D. thesis, University of Victoria, Victoria, BC, Canada.

McGuire, A., F. Chapin, J. Walsh, and C. Wirth, 2006: Integrated regional changes in arctic climate feedbacks: Implications for the global climate system. *Annual Review of Environmental Resources*, **31**, 61–91, doi: 10.1146/annurev.energy.31.020105.100253.

McGuire, A., J. Melillo, L. Joyce, D. Kicklighter, A. Grace, B. Moore, and C. Vorosmarty, 1992: Interaction between carbon and nitrogen dynamics in estimating net primary productivity for potential vegetation in North America. *Global Biogeochemical Cycles*, **6 (2)**, 101–124, doi:10.1029/92GB00219.

Meehl, G., et al., 2007: Global climate projections. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Forth Assessment Report of the Intergovernmental Panel on Climate Change*, S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, M. Tignor, and H. Miller, Eds., Cambridge University Press, Cambridge, UK and New York, NY, USA, chap. 10, 749–846.

Meinshausen, M., et al., 2011: The RCP greenhouse gas concentrations and their extension from 1765 to 2300. *Climatic Change*, **109 (1-2)**, 213–241, doi: 10.1007/s10584-011-0156-z.

Meissner, K., A. Weaver, H. Matthews, and P. Cox, 2003: The role of land surface dynamics in glacial inception: a study with the UVic Earth System Climate Model. *Climate Dynamics*, **21 (7-8)**, 515–537, doi:10.1007/s00382-003-0352-2.

Mitra, S., R. Wassman, and P. Vlek, 2003: Global inventory of wetlands and their role in the carbon cycle. ZEF - Discussion Papers on Development Policy 64, Center for Development Research.

Moss, R., et al., 2010: The next generation of scenarios for climate change research and assessment. *Nature*, **463 (7282)**, 747–756, doi:10.1038/nature08823.

Nakicenovic, N., J. Alcamo, G. Davis, B. de Vries, J. Fenhann, S. Gaffin, K. Gregory, and A. Grubler, 2000: *Special Report on Emissions Scenarios : a Special Report of Working Group III of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, UK and New York, NY, USA.

Nelson, F., O. Anisimov, and N. I. Shiklomanov, 2001: Subsidence risk from thawing permfrost. *Nature*, **410 (6831)**, 889–890, doi:10.1038/35073746.

Nelson, F., O. Anisimov, and N. I. Shiklomanov, 2002: Climate change and hazard zonation in the circum-Arctic permafrost regions. *Natural Hazards*, **26 (3)**, 203–225, doi:10.1023/A:1015612918401.

Nelson, F. and N. Shiklomanov, 1999: Variability of active-layer thickness at multiple spatial scales, North-central Alaska, USA. *Arctic, Antarctic and Alpine Research*, **31 (2)**, 179–186.

Nixon, M. F., 1998. Updated 2009.: Active layer monitoring, Arctic and sub-Arctic Canada. Boulder, Colorado USA: National Snow and Ice Data Center/World Data Center for Glaciology. Digital Media. (http://nsidc.org/data/ggd353.html).

Ramaswamy, V., et al., 2001: Radiative forcing of climate change. *Climate Change 2001: The Scientific Basis*, J. Houghton, Y. Ding, D. Griggs, M. Noguer, P. van der Linden, X. Dai, K. Maskell, and C. Johnson, Eds., Cambridge University Press, Cambridge, UK and New York, NY, USA, chap. 6, 349–416.

Reader, M. and G. J. Boer, 1998: The modification of greenhouse gas warming by the direct effect of sulphate aerosols. *Climate Dynamics*, **14 (7-8)**, 593–607, doi:10.1007/s003820050243.

Riahi, K., A. Grubler, and N. Nakicenovic, 2007: Scenarios of long-term socio-economic and environmental development under climate stabilization. *Technological Forecasting and Social Change*, **74 (7)**, 887–935, doi:10.1016/j.techfore.2006.05.026.

Richter, D. and D. Markewitz, 1995: How deep is soil? *BioScience*, **45 (9)**, 600–609.

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

Romanovsky, V., S. Smith, and H. Christiansen, 2010: 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, doi:10.1002/ppp.689.

Rowland, J., et al., 2010: Arctic landscapes in transition: Responses to thawing permafrost. *EOS*, **91 (26)**, 229–231, doi:10.1029/2010EO260001.

Saito, K., M. Kimoto, T. Zhang, K. Takata, and S. Emori, 2007: Evaluating a high-resolution climate model: Simulated hydrothermal regimes in frozen ground regions and their change under the global warming scenario. *Journal of Geophysical Research*, **112 (F02S11)**, doi:10.1029/2006JF000577.

Schaefer, K., T. Zhang, L. Bruhwiler, and A. P. Barrett, 2011: Amount and timing of permafrost carbon release in response to climate warming. *Tellus B*, **63 (2)**, 165–180, doi:10.1111/j.1600-0889.2011.00527.x.

Scholes, R. and E. Brown de Colstoun, 2011: ISLSCP II Global Gridded Soil Characteristics. Available online (http://www.daac.ornl.gov) from the ORNL Distributed Active Archive Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA.

Schuur, E., et al., 2008: Vulnerability of permafrost carbon to climate change: implications for the global carbon cycle. *BioScience*, **58 (8)**, 701–714, doi: 10.1641/B580807.

Serreze, M. and R. Barry, 2011: Processes and impacts of Arctic amplification: a research synthesis. *Global and Planetary Change*, **77 (1-2)**, 85–96, doi: 10.1016/j.gloplacha.2011.03.004.

Serreze, M. and C. Hurst, 2000: Representation of mean Arctic precipitation from NCEP-NCAR and ERA reanalyses. *Journal of Climate*, **13 (1)**, 182–201, doi: 10.1175/1520-0442(2000)013¡0182:ROMAPF¿2.0.CO;2.

Shindell, D., B. Walter, and G. Faluvegi, 2004: Impacts of climate change on methane emissions from wetlands. *Geophysical Research Letters*, **31 (L21202)**, doi:10.1029/2004GL021009.

Slater, A., et al., 2001: The representation of snow in land surface schemes: results from PILPS 2(d). *Journal of Hydrometeorology*, **2 (1)**, 7–25, doi:10.1175/1525-7541(2001)002¡0007:TROSIL¿2.0.CO;2.

Smith, L., Y. Sheng, and G. MacDonald, 2007: A first pan-Arctic assessment of the influence of glaciation, permafrost, topography and peatlands on Northern Hemisphere lake distribution. *Permafrost and Periglacial Processes*, **18 (2)**, 201–208, doi:10.1002/ppp.581.

Smith, L., Y. Sheng, G. MacDonald, and L. Hinzman, 2005: Disappearing Arctic lakes. *Science*, **308 (5727)**, 1429, doi:10.1126/science.1108142.

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

Smith, S. and M. Burgess, 2002: A digital database of permafrost thickness in Canada. Geological Survey of Canada, Open File 4173 - Available online, doi: doi:10.4095/213043.

Smith, S. and T. Wigley, 2006: Multi-gas forcing stabilization with minicam. *The Energy Journal*, **Special Issue 3**, 373–392.

Tarnocai, C., J. Canadell, E. Shuur, P. Kuhry, G. Mazhitova, and S. Zimov, 2009: Soil organic carbon pools in the northern circumpolar permafrost region. *Global Biogeochemical Cycles*, **23 (GB2023)**, doi:10.1029/2008GB003327.

Uppala, S., et al., 2005: The ERA-40 re-analysis. *Quarterly Journal of The Royal Meteorological Society*, **131 (612)**, 2961–3012, doi:10.1256/qj.04.176.

U.S. Department of Commerce, N. O. and N. G. D. C. Atmospheric Administration, 2006: 2-minute gridded global relief data (ETOPO2v2). http://www.ngdc.noaa.gov/mgg/fliers/06mgg01.html.

van den Hurk, B., P. Viterbo, A. Beljaars, and A. Betts, 2000: Offline validation of the ERA40 surface scheme. Technical Memorandum 295, European Centre for Medium-Range Weather Forecasts, Research Department.

van Vuuren, D. P., M. den Elzen, P. Lucas, B. Eickhout, B. Strengers, B. van Ruijven, S. Wonink, and R. van Houdt, 2007: Stabilizing greenhouse gas concentrations at low levels: an assessment of reduction strategies and costs. *Climatic Change*, **81 (2)**, 119–159, doi:10.1007/s10584-006-9172-9.

Viterbo, P. and A. Beljaars, 1995: An improved land surface parameterization scheme in the ECMWF model and its validation. *Journal of Climate*, **8**, 2716–2748, doi: 10.1175/1520-0442(1995)008¡2716:AILSPS¿2.0.CO;2.

Walsh, J., 2005: Cryosphere and hydrology. *Arctic Climate Impact Assessment - Scientific Report*, P. Grabhorn, J. Weybright, and C. Grabhorn, Eds., Cambridge University Press, Cambridge, UK and New York, NY, USA, chap. 6, 183–242.

Wania, R., 2007: Modelling northern peatland land surface processes, vegetation dynamics and methane emissions. Ph.D. thesis, Department of Earth Sciences, University of Bristol.

Ward, R. and M. Robinson, 2000: *Principles of Hydrology*. 4th ed., McGraw-Hill, London, UK.

Weaver, A., et al., 2001: The UVic Earth System Climate Model: Model description, climatology and applications to past, present and future climates. *Atmosphere-Ocean*, **39 (4)**, 361–428, doi:10.1080/07055900.2001.9649686.

Weber, S., A. Drury, W. Toonen, and M. van Weele, 2010: Wetlands methane emissions during the Last Glacial Maximum estimated from PMIP2 simulations: climate, vegetation and geographic controls. *Journal of Geophysical Research*, **115 (D06111)**, doi:10.1029/2009JD012110.

Weller, G., 2005: Summary and synthesis of the ACIA. *Arctic Climate Impact Assessment - Scientific Report*, P. Grabhorn, J. Weybright, and C. Grabhorn, Eds., Cambridge University Press, Cambridge, UK and New York, NY, USA, chap. 18, 989–1020.

Wheeler, B., 1999: Water and plants in freshwater wetlands. *Eco-hydrology: Plants and Water in Terrestrial and Aquatic Environments*, A. Baird and R. Wilby, Eds., Routledge, chap. 5, 127–180.

Williams, P. and M. Smith, 1989: *The Frozen Earth: Fundamentals of Geocryology*. Cambridge University Press, Cambridge, UK.

Wilson, M. and A. Henderson-Sellers, 1985: A global archive of land cover and soils data for use in general circulation climate models. *Journal of Climatology*, **5 (2)**, 119–143, doi:10.1002/joc.3370050202.

Wise, M., et al., 2009: Implications of limiting $CO_2$ concentrations for land use and energy. *Science*, **324 (5931)**, 1183–1186, doi:10.1126/science.1168475.

Woo, M.-K. and T. Winter, 1993: The role of permafrost and seasonal frost in the hydrology of northern wetlands in North America. *Journal of Hydrology*, **141 (1-4)**, 5–31, doi:10.1016/0022-1694(93)90043-9.

Zhang, T., 2005: Influence of the seasonal snow cover on the ground thermal regime: An overview. *Reviews of Geophysics*, **43 (RG4002)**, doi: doi:10.1029/2004RG000157.

Zhang, T., R. Barry, K. Knowles, J. Heginbottom, and J. Brown, 1999: Statistics and characteristics of permafrost and ground-ice distribution in the Northern Hemisphere. *Polar Geography*, **23 (2)**, 132–154, doi:10.1080/10889379909377670.

Zhang, T., R. Barry, K. Knowles, F. Ling, and R. Armstrong, 2003: Distribution of seasonally and perennially frozen ground in the Northern Hemisphere. *Proceedings of the 8th International Conference on Permafrost*, M. Phillips, S. Springman, and L. Arenson, Eds., A.A. Balkema Publishers, Zurich, Vol. 2, 1289–1294.

Zhang, T., J. Heginbottom, R. Barry, and J. Brown, 2000: Further statistics on the distribution of permafrost and ground-ice in the Northern Hemisphere. *Polar Geography*, **24 (2)**, 126–131, doi:10.1080/10889370009377692.

Zinke, P., A. Stangenberger, W. Post, W. Emanuel, and J. Olson, 1986: Worldwide organic carbon and nitrogen data. Tech. Rep. ONRL/CDIC-18, Carbon Dioxide Information Centre, Oak Ridge, Tenessee. doi:10.3334/CDIAC/lue.ndp018.

# Appendix A

# Model Variables

| **Parameter** | **Description** | **Value** |
|---|---|---|
| $\Delta z$ | Soil layer thickness | 1.0 m |
| b | Clapp-Hornberger Exponent | 6.63 |
| $K_s$ | Saturated Hydraulic Conductivity | $5.0 \cdot 10^{-4}$ kgm$^{-2}$s$^{-1}$ |
| $\lambda_{soil}$ | Soil thermal conductivity | 0.75 W m$^{-1}$ K$^{-1}$ |
| C | Soil volumetric heat capacity | $3.3 \cdot 10^5$ J m$^{-3}$ K$^{-1}$ |
| R | Gas constant | 287.05 J kg$^{-1}$ K$^{-1}$ |
| P | Surface air pressure | $10^{\ 5}$ Pa |
| $\mathrm{L}_c$ | Latent heat of condensation | $2.501 \cdot 10^6$ J kg$^{-1}$ |
| $\mathrm{L}_f$ | Latent heat of fusion | $3.34 \cdot 10^6$ J kg$^{-1}$ |
| $\mathrm{c}_p$ | Specific heat capacity of air (constant P) | 1005 J kg$^{-1}$ K$^{-1}$ |
| $\theta_c$ | Volumetric soil moisture content at critical point | 0.242 m$^3$ m$^{-3}$ |
| $\theta_s$ | Volumetric soil moisture content at saturation | 0.458 m$^3$ m$^{-3}$ |
| $\theta_w$ | Volumetric soil moisture content at wilting point | 0.136 m$^3$ m$^{-3}$ |
| $\sigma$ | Stefan-Boltzmann constant | $5.67 \cdot 10^{-8}$ W m$^{-2}$ K$^{-1}$ |
| $\kappa_s$ | Specific soil respiration rate | 5.9 x $10^{-9}$ s$^{-1}$ |
| $q_{10}$ | Soil respiration temperature constant | 2.0 |

Table A.1: Constants in version 3.0

| **Parameter** | **Description** | **Units** |
|---|---|---|
| $T_a$ | Surface air temperature | K |
| $T^*$ | Skin temperature | K |
| $T_g$ | Ground (soil) temperature | K |
| $R_n$ | Net radiation | $W\ m^{-2}$ |
| $SW_{\downarrow,abs}$ | Absorbed downwards shortwave radiation | $W\ m^{-2}$ |
| $LW_{\downarrow}$ | Downwards longwave radiation | $W\ m^{-2}$ |
| SH | Surface-to-atmosphere sensible heat flux | $W\ m^{-2}$ |
| LH | Surface-to-atmosphere latent heat flux | $W\ m^{-2}$ |
| GH | Surface-to-ground sensible heat | $W\ m^{-2}$ |
| $r_a$ | Aerodynamic resistance | $s\ m^{-1}$ |
| $r_s$ | Surface resistance | $s\ m^{-1}$ |
| $r_{s,soil}$ | Surface resistance, bare soil tile | $s\ m^{-1}$ |
| $r_{s,PFT}$ | Surface resistance, PFT tile | $s\ m^{-1}$ |
| $\rho_a$ | Surface air density | $kg\ m^{-3}$ |
| $q_{sat}$ | Saturated specific humidity | $kg_{H_2O}\ kg_{air}^{-1}$ |
| $q_a$ | Specific humidity | $kg_{H_2O}\ kg_{air}^{-1}$ |
| $\theta$ | Volumetric soil moisture concentration | $m_{water}^3\ m_{soil}^{-3}$ |
| E | Evapotranspiration rate | $kg\ m^{-2}\ s^{-1}$ |
| $M_u$ | Unfrozen soil moisture areal concentration | $kg\ m^{-2}$ |
| RF | Rainfall rate | $kg\ m^{-2}\ s^{-1}$ |
| SM | Snowmelt rate | $kg\ m^{-2}\ s^{-1}$ |
| K | Hydraulic conductivity | $kg\ m^{-2}\ s^{-1}$ |
| W | Water flux (drainage) | $kg\ m^{-2}\ s^{-1}$ |
| $S_{snow}$ | Sublimation from snow | $kg\ m^{-2}\ s^{-1}$ |
| $\beta$ | Soil moisture stress factor | |
| $\mathcal{C}$ | Vegetation carbon content | $kg\ m^{-2}$ |
| f | Vegetation gridcell areal fraction | |
| N | Net Primary Productivity | $kg\ m^{-2}\ s^{-1}$ |
| $\Lambda$ | Fraction of NPP used to augment f | |
| $\Gamma_{local}$ | Local litterfall rate | $kg\ m^{-2}\ s^{-1}$ |
| $\Gamma_{largescale}$ | Large scale litterfall rate | $kg\ m^{-2}\ s^{-1}$ |
| $\Gamma$ | Total gridcell litterfall rate | $kg\ m^{-2}\ s^{-1}$ |

Continued on Next Page…

Table A.2 – Continued

| | | |
|---|---|---|
| $\alpha$ | PFT competition factor | |
| $\mathcal{C}_s$ | Soil carbon | kg m$^{-2}$ |
| $R_{soil}$ | Soil respiration rate | kg m$^{-2}$ s$^{-1}$ |
| $f(\theta_l)$ | Respiration moisture function | |
| f(T) | Respiration temperature function | |

Table A.2: Variables in UVic Version 3.0

| **Parameter** | **Description** | **Value** |
|---|---|---|
| $b_o$ | Clapp-Hornberger exponent, organic soil | 2.7 |
| $\Phi_{s,o}$ | Matric potential, organic soil | 0.0103 m |
| $K_{s,o}$ | Saturated hydraulic conductivity, organic soil | $2.0 \cdot 10^{-2}$ kg m$^{-2}$s$^{-1}$ |
| $\theta_{s,o}$ | Porosity, organic soil | 0.900 m$^3$ m$^{-3}$ |
| $C_{dry,o}$ | Dry soil heat capacity, organic soil | $0.25 \cdot 10^6$ J m$^{-3}$ K$^{-1}$ |
| $\lambda_{dry,o}$ | Dry soil thermal conductivity, organic soil | 0.05 W m$^{-2}$ |
| $\rho_p$ | Density of peat | 130 kg m$^{-3}$ |
| $c_w$ | Specific heat capacity, water | 4180 J kg$^{-1}$ K$^{-1}$ |
| $c_i$ | Specific heat capacity, ice | 2100 J kg$^{-1}$ K$^{-1}$ |
| $\rho_w$ | Density of pure water | 1000 kg m$^{-3}$ |
| $\rho_i$ | Density of ice | 917 kg m$^{-3}$ |
| $\rho_{snow}$ | Density of snow | 330 kg m$^{-3}$ |
| $\lambda_{snow}$ | Thermal conductivity of snow | 0.250 W m$^{-1}$ K$^{-1}$ |
| $\epsilon$ | Phase change constant | 114.3 m K$^{-1}$ |
| g | Acceleration due to gravity | 9.81 m s$^{-2}$ |
| $T_m$ | Melting point of pure, unconfined water | 273.15 K |

Table A.3: Additional constants in version 3.0M

| **Parameter** | **Description** | **Units** |
|---|---|---|
| b | Clapp-Hornberger exponent | |
| $b_m$ | Clapp-Hornberger exponent, mineral soil | |
| $\phi_s$ | Matric potential | m |
| $\phi_{s,m}$ | Matric potential, mineral soil | m |
| $K_s$ | Saturated hydraulic conductivity | kg $m^{-2}$ $s^{-1}$ |
| $K_{s,m}$ | Saturated hydraulic conductivity, mineral soil | kg $m^{-2}$ $s^{-1}$ |
| $C_{dry}$ | Dry soil volumetric heat capacity | J $K^{-1}$ $m^{-3}$ |
| $C_{dry,m}$ | Dry soil volumetric heat capacity, mineral soil | J $K^{-1}$ $m^{-3}$ |
| $\theta_s$ | Porosity | $m^3_{water}$ $m^{-3}_{soil}$ |
| $\theta_{s,m}$ | Porosity, mineral soil | $m^3_{water}$ $m^{-3}_{soil}$ |
| $\lambda_{dry}$ | Dry soil thermal conductivity | W $m^{-1}$ $K^{-1}$ |
| $\lambda_{dry,m}$ | Dry soil thermal conductivity, mineral soil | W $m^{-1}$ $K^{-1}$ |
| $\lambda_{cs}$ | Thermal conductivity, compacted soil | W $m^{-1}$ $K^{-1}$ |
| $\%_{Clay}$ | Percentage of clay-sized particles in mineral soil | % |
| $\%_{Sand}$ | Percentage of sand-sized particles in mineral soil | % |
| $\%_{Silt}$ | Percentage of silt-sized particles in mineral soil | % |
| CCP | Cumulative Carbon Profile Function | % |
| $\rho_c$ | Carbon density | kg $m^{-3}$ |
| $\sigma_c$ | Carbon areal density | kg $m^{-2}$ |
| z | Depth in soil profile | m |
| $z_{bottom}$ | Depth of bottom boundary of soil layers | m |
| $z_{top}$ | Depth of top boundary of soil layers | m |
| $\Delta z_n$ | Thickness of soil layer n | m |
| $f_o$ | Fraction of soil that is organic | |
| $f_m$ | Fraction of soil that is mineral | |
| $M_{l,n}$ | Liquid soil moisture content in a layer | kg $m^{-2}$ |
| $M_{f,n}$ | Frozen soil moisture content in a layer | kg $m^{-2}$ |
| $W_n$ | Water flux through base of layer n | kg $m^{-2}$ $s^{-1}$ |
| $\Psi$ | Soil water suction | m |
| $E_n$ | Fraction of evapotranspiration from layer n | kg $m^{-2}$ $s^{-1}$ |
| $S_{soil}$ | Sublimation from bare soil tiles | kg $m^{-2}$ $s^{-1}$ |
| $\rho_{root}$ | Root density | kg $m^{-3}$ |

Continued on Next Page…

Table A.4 – Continued

| | | |
|---|---|---|
| $\rho_{root,0}$ | Root density at z = 0 | kg $m^{-3}$ |
| $d_{r,PFT}$ | Rooting depth | m |
| D | Total soil depth | m |
| $B_{n,PFT}$ | Soil moisture stress function within a layer and PFT | |
| $\lambda_{soil}$ | Soil thermal conductivity | W $m^{-1}$ $K^{-1}$ |
| $\lambda_{sat}$ | Soil thermal conductivity at saturation | W $m^{-1}$ $K^{-1}$ |
| $\lambda_{cs}$ | Thermal conductivity of compacted soil | W $m^{-1}$ $K^{-1}$ |
| $\lambda_b$ | Soil thermal conductivity at layer boundaries | J $K^{-1}$ m $^{-3}$ |
| $\theta_l^s$ | Liquid volumetric soil moisture concentration at saturation | $m^3_{water}$ $m^{-3}_{soil}$ |
| $\theta_f^s$ | Frozen volumetric soil moisture concentration at saturation | $m^3_{water}$ $m^{-3}_{soil}$ |
| $C_s$ | Soil heat capacity | J $K^{-1}$ m $^{-3}$ |
| $T_{g,n}$ | Temperature of soil layer n | K |
| $G_n$ | Heat transport by diffusion through base of layer n | W $m^{-2}$ |
| $J_n$ | Heat transport by advection of water into layer n | W $m^{-2}$ |
| $\Delta G$ | Gibbs Free Energy | J kg $^{-1}$ |
| $T_{max}$ | T above which all soil moisture is unfrozen | K |
| $\theta_l^{max}$ | Maximum amount of unfrozen moisture at T | $m^3_{water}$ $m^{-3}_{soil}$ |
| $C_{s,eff}$ | Effective soil heat capacity including phase changes | J $K^{-1}$ m $^{-3}$ |
| $f_{frozen}$ | Fraction of a gridcell with surface ground ice | |
| $f_{liquid}$ | Fraction of a gridcell with surface ground water | |
| $r_{s,soil,evap}$ | Surface resistance for soil evaporation | s $m^{-1}$ |
| $r_{s,soil,sub}$ | Surface resistance for soil sublimation | s $m^{-1}$ |
| $\lambda_{eff}$ | Effective thermal conductivity with snow present | W $m^{-1}$ $K^{-1}$ |

Table A.4: Additional Variables in UVic Version 3.0M

## A.1 UVic Version 3.0M

## A.2 Moisture Units

This and subsequent sections will make use of several different, interchangeable units for the moisture contents within a soil layer and they are here defined:

- **M** ($\frac{kg}{m^2}$): Areal density of total soil moisture per unit horizontal area.
- $\boldsymbol{\theta}$ ($\frac{m^3}{m^3}$) : Volumetric concentration of moisture - the volume of water per unit volume of soil (that is, the fraction of the volume of the soil that is water) if all the soil moisture were in liquid form.

Soil moisture is constrained such that $\theta$ never exceeds the saturation volumetric soil moisture concentration, $\theta_s$, which is a soil layer-dependent constant. $\theta$ is related to the volumetric concentrations of the liquid ($\theta_l$ - volume of liquid water : volume of soil) and frozen ($\theta_f$ - volume of ice : volume of soil) soil moisture.

$$\theta = \theta_l + \frac{\rho_i}{\rho_w}\theta_f \tag{A.1}$$

where $\rho_w$ and $\rho_i$ are the respective densities of liquid water and ice. With $\Delta$ z as the soil layer thickness, M may be related to these quantities:

$$M = M_u + M_f = \Delta z \rho_w \left(\theta_l + \frac{\rho_i}{\rho_w}\theta_f\right) \tag{A.2}$$

- $\mathbf{S}_l$**,** $\mathbf{S}_f$ (dimensionless): Soil moisture contents expressed as a fraction of liquid water content at saturation.

$$S_l = \frac{\theta_l}{\theta_s} \qquad S_f = \frac{\rho_i}{\rho_w}\frac{\theta_f}{\theta_s} \tag{A.3}$$

Two additional parameters are important in determining moisture availability to plants - the volumetric moisture contents at the wilting ($\theta_w$) and critical (or field capacity) points ($\theta_c$). The wilting point represents a level of soil suction above which plants are unable to extract moisture from the soil, taken here to be 1.53 MPa (= 156.4 m equivalent of water) [1]. while the 0.033 MPa (= 3.364 m equivalent of water). The field capacity corresponds to a soil suction value below which the suction is too weak to retain moisture against the force of gravity, here this point corresponds to a suction of 0.033 MPa (= 3.364 m equivalent of water). This point also corresponds to

---

[1] Wikipedia gives 1.5 MPa

the moisture content of the soil just after gravitational drainage has ceased; when the soil moisture contents exceed this level it is assumed that the soil offers no resistance to plants' ability to extract moisture. Both $\theta_w$ and $\theta_c$ are calculated by inverting equation 2.31:

$$\theta_w = \theta_s \left( \frac{3.364m}{\Psi_s} \right)^{-\frac{1}{b}} \tag{A.4}$$

$$\theta_c = \theta_s \left( \frac{156.4m}{\Psi_s} \right)^{-\frac{1}{b}} \tag{A.5}$$

# Appendix B

# Impact of Revisions to Land Surface Scheme

## B.1 Offline Model Configuration

The purpose of this appendix is to assess the impacts of the various modifications to the land surface scheme outlined in Chapter 2, using the offline configuration of the model, UVic 3.0E. Four offline versions of the model are compared in this section. They are as follows:

1. **3.0Ea** - Offline configuration of the land surface scheme (LSS) from version 3.0 of the model without any changes implemented by the author.
2. **3.0Eb** - Offline configuration of the LSS from version 3.0M with the soil hydrological and thermal parameters of version 3.0 (Table 2.1, left column.) and no soil freeze-thaw processes. Here as well, soil thermal properties have no soil moisture dependence, as in UVic 3.0.
3. **3.0Ec** - Offline configuration of the LSS from version 3.0M with no soil freeze-thaw processes. Here, soil thermal conductivity and heat capacity depend on the soil moisture content.
4. **3.0E** - Offline configuration of the LSS from version 3.0M with all modifications described in Chapter 2.

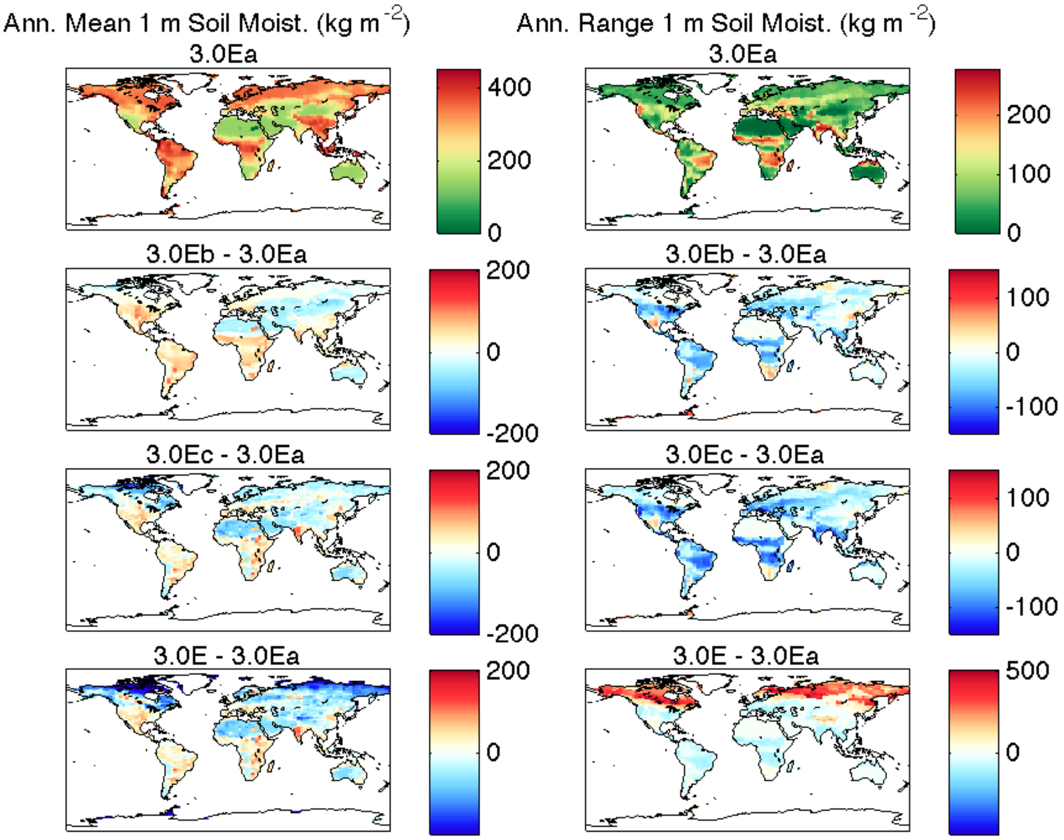


Figure B.1: Changes in near surface soil moisture storage. The left hand column shows the annual mean soil moisture content in the top 1 m of soil (kg m$^{-2}$) while the right hand column shows the annual range in near surface soil moisture content (kg m$^{-2}$). The top row show these quantities as simulated in v. 3.0Ea (single layer) of the ESCM. The bottom three rows shows the differences between v. 3.0Eb, 3.0Ec and 3.0E and v. 3.0Ea.

## B.2 Changes in Soil Moisture

I firstly examine changes in near surface soil moisture content at a global scale. Since v. 3.0Ea has only one soil layer with a thickness of 1 m, to obtain results from the multilayer versions of the model, the soil moisture content was summed over the top four layers (to a depth of 1.06 m). The multilayer moisture contents were then scaled down by a factor of 1.06 so that the total moisture contents in the top meter of soil were compared across model versions.

### B.2.1 Impact of Inclusion of Multiple Soil Layers

Figure B.1 compares the annual mean soil moisture and range in moisture content in the four configurations. The first analysis investigates the impact on moisture content as the LSS is switched from having a single 1 m soil layer to multiple layers with a total depth of 10 m (v. 3.0Eb); these impacts are shown in the second row of panels in the figure. V. 3.0Ea of the model does not simulate of the effects of moisture infiltration through soil layers to reach the level at which drainage can occur. Thus, matric potential plays no role in restricting the movement of moisture through the soil, as is the case in the multilayer scheme. Any moisture that enters the soil column in v. 3.0Ea is immediately able to begin draining through the bottom of the single soil layer. Additionally, in v. 3.0Eb, deeper soil layers (at depths below 1 m) may interact with near surface layers. As well, evapotranspiration rates may be different as a result of a change in the moisture availability to plants, with evapotranspiration rates dependent on the vertical distribution of soil moisture as well as the rooting depth of the vegetation in each tile.

The largest change in mean, near surface soil moisture storage occurs in the tropics where the highest precipitation rates occur and moisture content typically increases in v. 3.0Eb. A plot of the change in the mean rate of precipitation - evaporation (P-E, Figure B.2) shows a strong decrease in P-E in the tropics, indicating that the evaporation rate within the tropics actually increases (as the precipitation field remains fixed in these simulations). Thus, an increase in near surface soil moisture storage occurs despite an overall increase in evapotranspiration. Therefore, the inclusion of the effects of soil moisture movement through the soil column in v. 3.0Eb results in an increase in the near surface soil moisture content and a reduction in drainage from the base of the soil column. Another strong increase in soil moisture concentration is seen throughout the relatively dry central and southern United States. Throughout some parts of this region, there is a reduction in evaporation (Figure B.2) which would result in an increase in near surface soil moisture content; elsewhere, the increase is simply a result in the redistribution of soil moisture within the soil column with increased soil moisture storage near the surface.

A slight reduction in soil moisture concentration in v. 3.0Eb occurs in desert regions, such as the Saraha and Central Australia. Here, P-E rates remain unchanged. In these regions, precipitation is consistently extremely low throughout the year. Here, the multilayer model simulates a strong gradient in soil moisture content with depth with very low values near the surface and higher values at depth. This is attributable to the fact that the influence of the force of gravity in Richard's equation governing water flow would tend to result in higher moisture contents at depth in situations where there is virtually no significant water influx from the surface. As well, in the PFT-free deserts, evaporation occurs strictly from the surface layer. Both factors tend to reduce the near surface soil moisture content relative to soil moisture content at depth. As a result, when comparing the multilayer model against v. 3.0Ea (which simulates no soil moisture gradient), the near surface soil moisture concentration decreases in such regions.

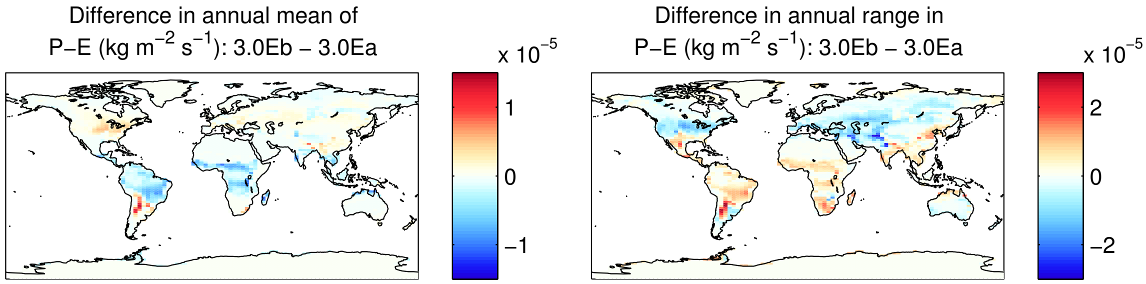


Figure B.2: Changes in precipitation minus evapotranspiration (P-E) rates between version 3.0Eb and version 3.0Ea (kg m$^{-2}$s$^{-1}$). Left panel: Changes in annual mean P-E; right panel: changes in annual range in P-E

The annual range in near surface soil moisture content is reduced throughout most of the world in v. 3.0Eb (second row, right column, Figure B.1). Most of the changes in the annual range in soil moisture content, may be explained based on changes in the range in P-E between the two versions (Figure B.2). An increase in the P-E range results from an increase in E during dry (low P) or a decrease in E under wet (high P) conditions. Conversely, a decrease in the annual P-E range results from an increase in E during wet conditions, or a decrease in E during dry conditions.

### B.2.2 Impact of New Soil Parameters

Version 3.0Ec of the model includes spatially varying mineral soil characteristics and the inclusion of organic soils. Figure B.3 shows the change in key soil moisture parameters averaged over the top 1 m of soil in this version. To understand the individual impact of changing these soil moisture parameters, a separate set of four model runs was conducted in which one soil parameter at a time was changed from its v. 3.0Eb value; these results are displayed in Figure B.4.

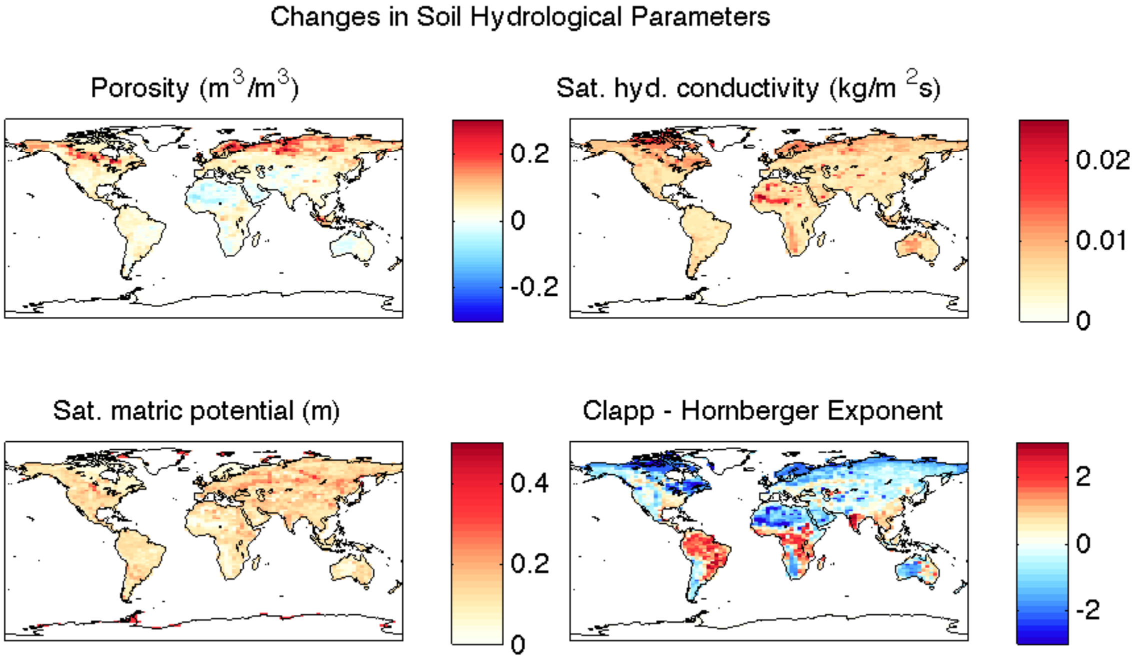


Figure B.3: Changes in soil hydrology parameters in v. 3.0Ec to v. 3.0Eb. For comparison purposes, the variables have the following constant values in version 3.0Eb. Porosity ($\theta_s$): 0.458 m$^3$/m$^3$, saturated hydraulic conductivity (K$_s$): 0.0005 kg/m$^2$s, saturated matric potential ($\phi_s$): 0.0495 m, Clapp-Hornberger exponent (b): 6.6

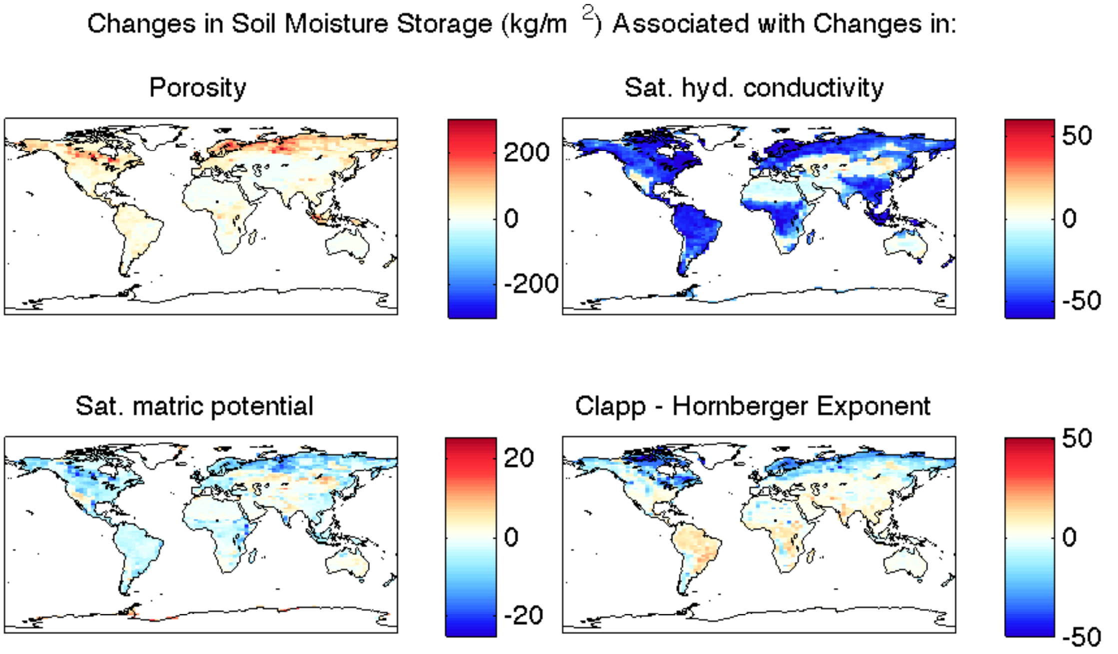


Figure B.4: Changes in annual mean soil moisture storage in the top 1 m of soil associated with moisture parameter changes in Figure B.3.

Porosity measures the volume of soil moisture at saturation relative to the total volume of all soil constituents (including soil moisture); it is therefore, a measure of the soil's moisture storage capacity. At a global scale, porosity remains fairly similar between the two model versions, with the notable exception of regions of where the organic soil content is high, owing to the high porosity of such soils. As the top left panel of Figure B.4 shows, there is a very strong increase in soil moisture storage in these organic rich regions when only the porosity is altered.

The saturated hydraulic conductivity governs the rate of soil moisture movement between soil layers as well as the rate of drainage through the base of the soil column. The soil hydraulic conductivity in v. 3.0Ea (and v .3.0Eb) was extremely low compared with commonly cited mean soil values and is increased by between one to two orders of magnitude in version 3.0Ec, meaning that soil moisture should now move more rapidly through the soil column. As such, the effects of soil water percolation that were discussed in the previous section as a reason for increased near surface soil moisture storage in v. 3.0Eb are diminished as the rate of percolation is substantially increased. Indeed, the top right panel of Figure B.4 indicates that, with the exception of desert regions, near surface soil moisture storage strongly decreases as a result of the increased hydraulic conductivity. In regions of very dry soil, what little soil moisture is present is retained by the very strong matric potential that occurs at low moisture concentrations, such that there is little-to-no change here when the hydraulic conductivity increases. The third row of Figure B.1 indicates that regions of higher soil moisture storage in v. 3.0Eb (as compared with v. 3.0Ea) such as the tropics and the central US now show typically show closer agreement with v. 3.0Ea. Investigation of all the soil moisture parameter changes suggests that it is the hydraulic conductivity changes that are most strongly responsible for this change.

Matric potential at saturation term influences the strength of water fluxes between layers having differing soil moisture contents. This quantity increases in v. 3.0Ec relative to 3.0Eb by up to an order of magnitude, but the spatial pattern is complex as is the impact on soil moisture storage (bottom left panel, Figure B.4). An increase in matric potential should tend to reduce variations in soil moisture content across the soil profile in a grid cell. The overall pattern of changes on the near surface soil moisture storage thus depends on not only on the strength of the matric potential change, but on the nature of the soil moisture profile in a grid cell. Where soil moisture typically tends to be higher near the surface, than at depth, surface soil moisture storage decreases; where it is higher at depth, surface moisture storage increases.

Finally, the Clapp-Hornberger (C-H) exponent determines the dependence of the matric potential and hydraulic conductivity on soil moisture content (Figure B.5). Where the coefficient is low (soils with high organic contents or high sand content), the matric potential varies relatively slowly with soil moisture, but the point at which soil hydraulic conductivity begins to rapidly increase occurs at a relatively low soil moisture concentration. The reverse is true for soils with a high Clapp-Hornberger exponent (clay rich soils). The strongest impacts of the C-H exponent changes (bottom right panel, Figure B.4) occur in the tropics where there is typically an increase in near surface soil moisture storage associated with an increase in the C-H coefficient. Here, there are two competing effects. Firstly, the matric potential varies more strongly with soil moisture. As soil moisture concentrations are generally higher near the surface in such regions, this effect should lead to reduced near-surface soil moisture storage. Conversely, the moisture level above which the hydraulic conductivity begins to strongly increase shifts to higher values, explaining the increased soil moisture storage, suggesting that the hydraulic conductivity effect is the dominant impact of the change in the C-H exponent. The strongest reduction in soil moisture storage in seen in the high latitudes of the Northern Hemisphere in regions of organic / sand rich soils. Here, the Clapp-Hornberger coefficient decreases, resulting in the opposite effect on hydraulic conductivity, leading to reduced soil moisture storage in the upper layers of the soil.

The overall pattern of changes in soil moisture storage when all four soil parameters are changed together is complicated (Figure B.6) and is not simply the direct sum of the individual soil moisture changes outlined above. Comparing the overall pattern to the individual soil moisture parameter changes suggests that there is no one single parameter change that dominates over all others at a global level. It is significant however, that the magnitude of a change in soil moisture storage associated with a change in a single parameter is not necessarily indicative of its overall importance when all soil moisture parameters are change. For example, by far the strongest change shown in Figure B.4 is associated with the increase in the saturated soil moisture concentration where soil carbon content is highest. This strong increase strongly dominates over the combined decreases in soil moisture storage in such regions from changes in the hydraulic conductivity, matric potential and the C-H exponent. If soil moisture changes added linearly, one would expect to see a very strong increase in soil moisture storage ($\sim$100 kg/m$^2$ or greater) in regions dominated

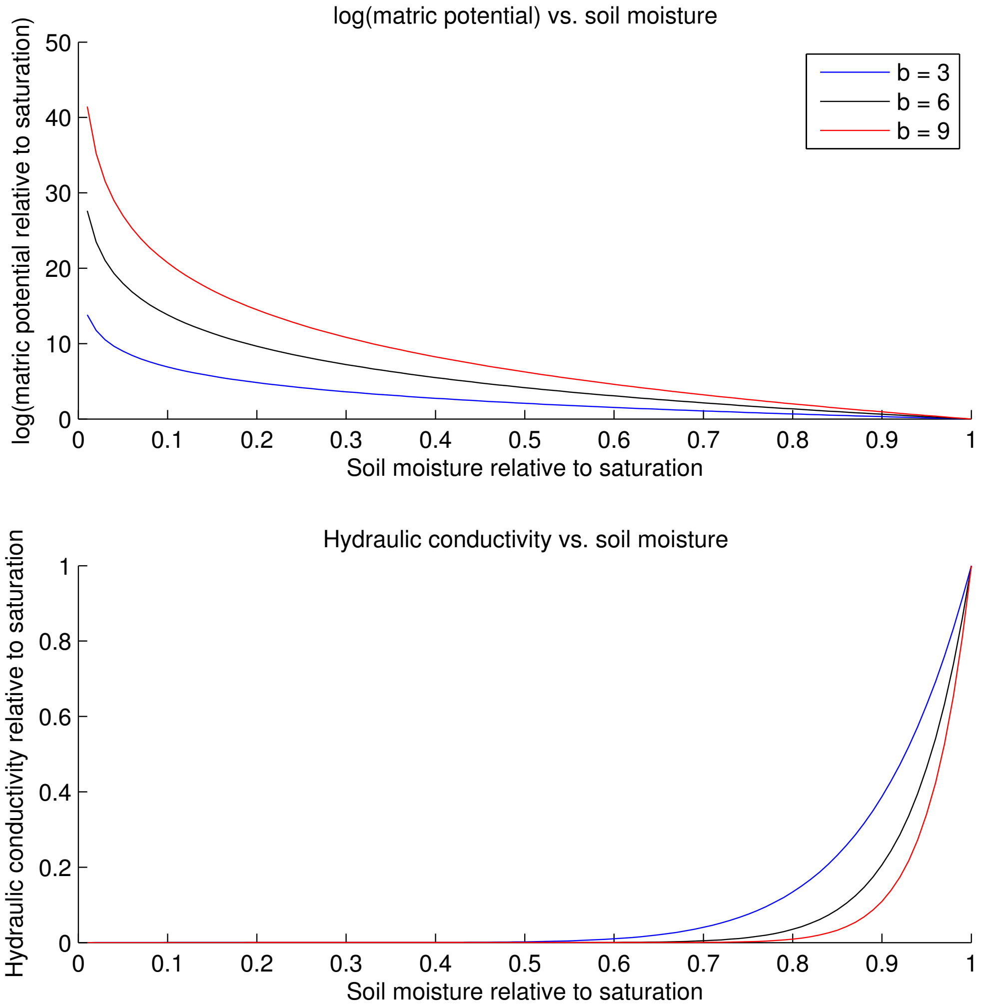


Figure B.5: Matric potential and hydraulic conductivity as a function of soil moisture content. Top panel: logarithm of matric potential as a fraction of its saturation value as soil moisture is varied from zero to saturation. Bottom panel: soil hydraulic conductivity as a fraction of its saturation value as soil moisture is varied from zero to saturation. In both cases, three separate curves are plotted for representative values of the Clapp-Hornberger coefficient from version 3.0Ec's soil dataset. Note that version 3.0Eb has a Clapp-Hornberger coefficient of 6.6.

by organic soils, but such changes are not revealed in Figure B.6. Lastly, it is noted that the spatial pattern in the annual range in near surface soil moisture content is very similar between this version of the model and v. 3.0Eb (Figure B.1), although the annual range in near surface soil moisture storage is smaller in v. 3.0Ec.

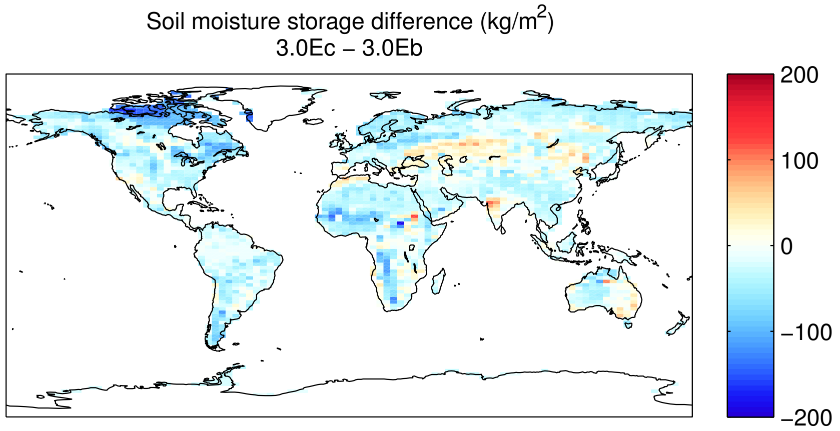


Figure B.6: Annual mean near surface soil moisture difference (kg m$^{-2}$) between v. 3.0Ec and 3.0Eb

### B.2.3 Impact of Soil Freeze-Thaw Processes

Version 3.0E of the model allows for soil phase changes. As expected, there are no changes in the annual mean soil moisture storage nor its range in regions that are not subject to soil freezing and thawing (Figure B.1). Where soil moisture freezing does occur, the annual mean liquid soil moisture content is typically strongly reduced. Here, however, the annual range in liquid soil moisture strongly increases as water can remain pooled near the surface during thaw seasons as a result of the presence of liquid soil layers below. As such, in soil layers within the active layer, liquid soil moisture can vary from near zero to near saturation over the course of the year, explaining the very high range in liquid soil moisture in such regions.

While the liquid soil moisture content decreases in permafrost regions, there is a strong increase in *total* (frozen + liquid) near surface soil moisture content (Figure B.7). The annual range in total soil moisture content strongly decreases here due to the fact that drainage is sharply reduced by the presence of frozen soil layers, increasing soil moisture storage. Conversely, there is a strong increase in the total soil moisture range in regions that experience seasonal frost. Here, drainage is impeded

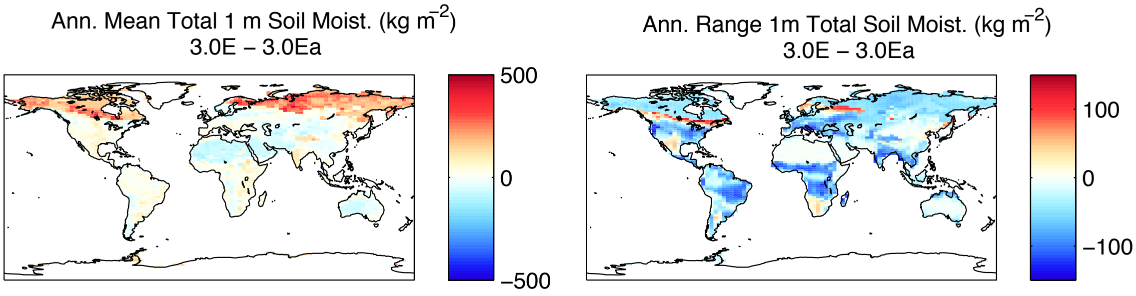


Figure B.7: Changes in total soil moisture storage (kg m$^{-2}$) in v. 3.0E relative to v. 3.0Ea

only in the winter, allowing moisture to build up in the winter months, then drain once the frozen layers thaw. As a final observation, it is noted that of all the model processes that have been investigated, the inclusion of soil freeze-thaw effects seem to have the strongest impacts on soil moisture budgets as evidenced by the plots in bottom row of Figure B.1 and Figure B.7.

## B.3 Changes in Soil Temperature

I next consider changes in soil temperature/ For comparison purposes, in the multilayer versions of the model, a mean soil temperature was calculated by averaging the soil temperature over the top four soil layers (to a depth of 1.06 m), weighting the temperatures by the relative thickness of each soil layer. The range in this depth-average temperature was then used to determine the range in the near surface soil temperature. Figure B.8a shows the annual mean and range in near surface (top 1 m) soil temperature in v. 3.0Ea as well as changes in theses quantities in the multilayer versions of the model relative to v. 3.0Ea. Changes in soil temperature may result if there is a change in the surface skin temperature, since it is the skin temperature that governs the soil heat flux. As such, to help understand the observed changes in soil temperature a plot of the mean and range in the skin temperature across the versions was also made (Figure B.8b). Changes in the skin temperature that do occur across the model versions are linked to changes in the soil temperature, evapotranspiration and sublimation and albedo (associated with changes in the distribution of lying snow). A detailed analysis of changes to the skin temperature would therefore be quite complicated and is beyond the scope of this work. However, it is noted that changes in skin temperature are typically much smaller than the equivalent soil tem-

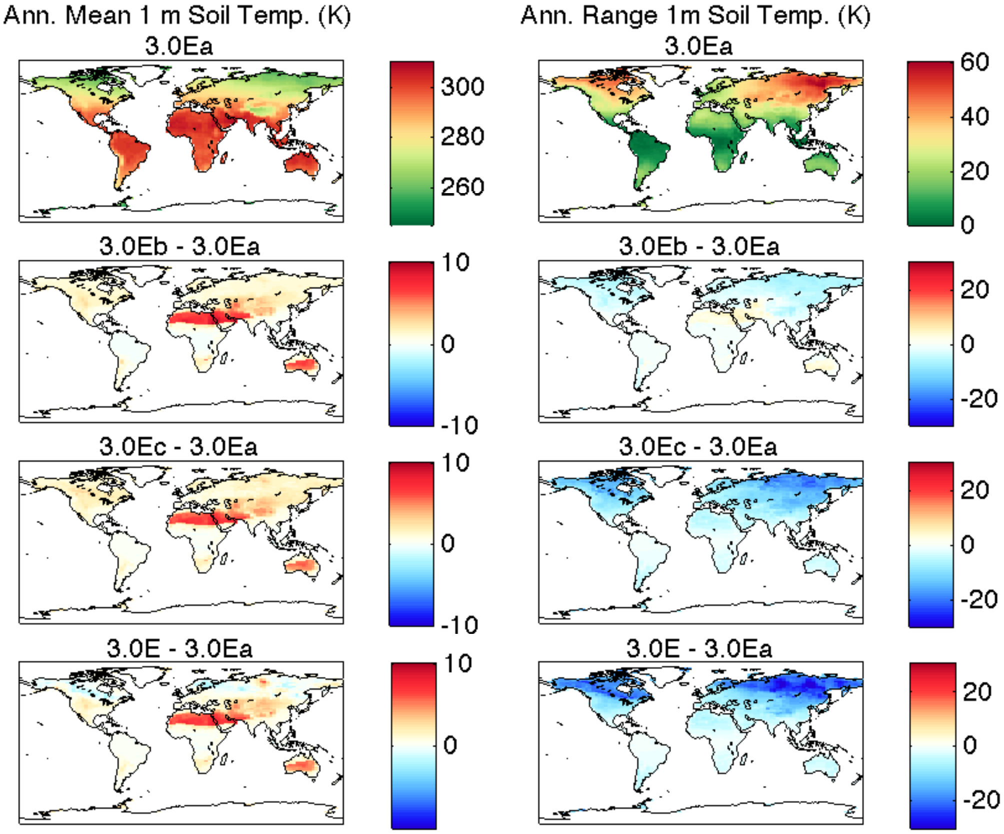


Figure B.8a: Annual mean soil temperature (K) and soil temperature range (K). The top row shows the annual mean temperature and annual temperature range in version 3.0Ea of the model. Subsequent rows show changes in these quantities relative to version 3.0Ea

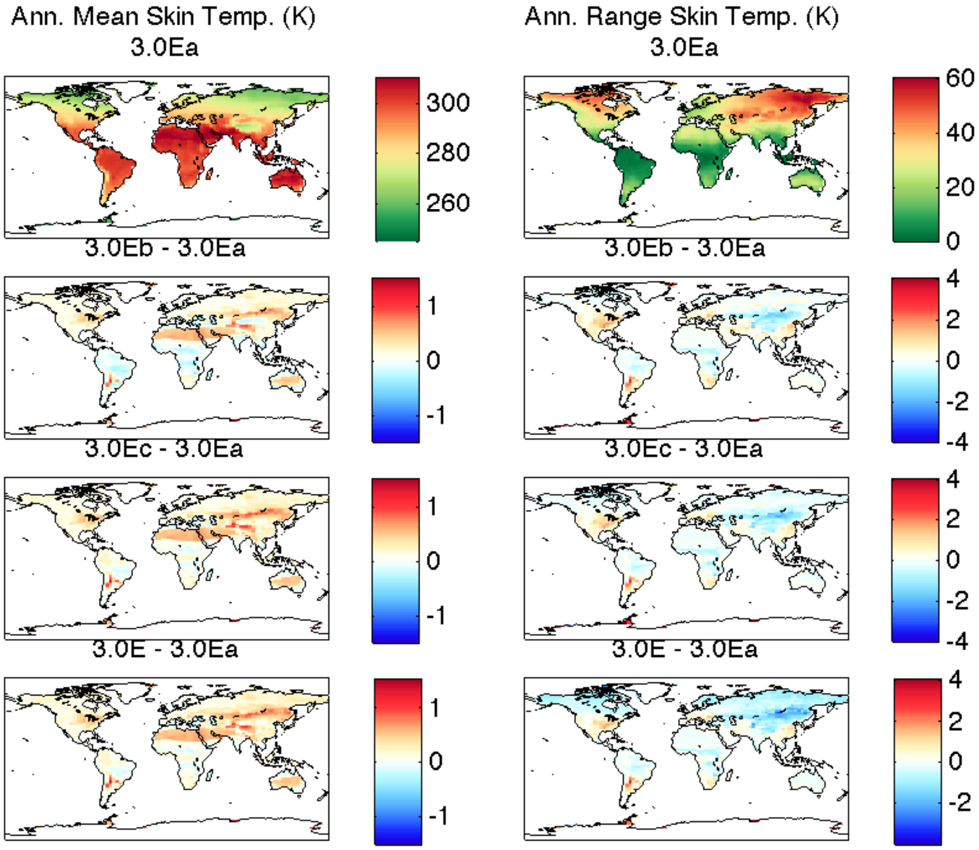


Figure B.8b: Annual mean skin temperature(K) and skin temperature range (K). Changes in skin temperature across model versions are, in general, substantially smaller than corresponding changes in soil temperature.

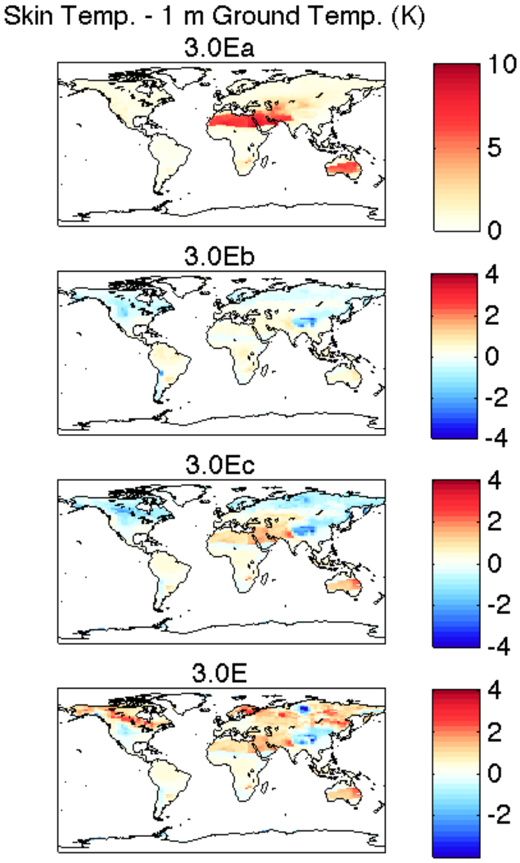


Figure B.9: Difference between annual mean surface temperature ($T^*$) and annual mean near surface ground temperature in top 1 m of soil ($T_g$) (K).

perature changes (Figure B.8a vs. B.8b). Therefore, other factors must be invoked to explain the most significant soil temperature changes.

### B.3.1 Impact of Inclusion of Multiple Soil Layers

The strongest impact of inclusion of multiple soil layers (3.0Ea vs. 3.0Eb) in the model is a sharp increase in mean soil temperature of between 2 - 10 K in desert regions. In v. 3.0Ea and 3.0Eb, soil temperatures are influenced solely by the ground sensible heat flux and the soil thermal conductivity is constant and not depending on soil moisture content. One would therefore expect that, once the land surface scheme had equilibrated, the annual mean soil temperature should be nearly identical to the annual mean surface (skin) temperature. As Figure B.9 reveals, while this is the case over large regions of the surface, there is a strong gradient between the skin temperature, T$^*$ and the ground temperature, T$_g$ in areas of the world where a substantial portion of the gridcell is dominated by bare soil; in the multilayer versions of the model, this gradient is typically about an order of magnitude smaller. The explanation for this anomaly has to do with errors introduced in solving the surface energy budget. The surface energy budget is a function of temperature (b(T)) with T$^*$ as the temperature that balances the equation.

$$b(T^*) = SW_{\downarrow,abs} + LW_{\downarrow} - \sigma \boldsymbol{T^{*4}} - \frac{\rho_a c_p(\boldsymbol{T^*} - T_a)}{\boldsymbol{r_a}} - \frac{\rho_a(\boldsymbol{q_{sat}(T*)} - q_a)(L_c)}{\boldsymbol{r_a} + \boldsymbol{r_c}} - \lambda_{soil}\frac{\boldsymbol{T^*} - T_g}{\frac{\Delta z}{2}} = 0 \tag{B.1}$$

The model uses a one step Newton-Raphson method to determine T$^*$, by approximating the energy budget as a linear function of temperature:

$$b(T^*) = b(T_o) + \frac{db}{dT}(T^* - T_o) \tag{B.2}$$

where T$_o$ is the initial guess for T$^*$ and the derivative of b is also evaluated at T$_o$. Since, by definition of T$^*$, b(T$^*$) = 0, it follows that:

$$T^* = T_o - \frac{b(T_o)}{\frac{db}{dT}} \tag{B.3}$$

The air temperature (T$_a$) is used as the initial guess value (T$_o$). The accuracy of the solution then depends on how well the linearized function approximates the true function. Since the sensible heat and ground heat terms depend linearly on temperature, and the shortwave and incident longwave radiation terms are constant as far as the LSS energy budget code is concerned, the two terms that determine the solution accuracy are the outgoing longwave radiation term and the latent heat flux term which are approximated as follows:

$$\sigma(T)^4 \approx \sigma(T_a)^4 + 4\sigma(T_a)^3(T - T_a) \tag{B.4}$$

$$\frac{\rho_{air}(\boldsymbol{q_{sat}}(\boldsymbol{T}) - q_1)(L_c)}{\boldsymbol{r_a} + \boldsymbol{r_s}} \approx \frac{\rho_{air}(\boldsymbol{q_{sat}}(\boldsymbol{T_a}) + \frac{dq_{sat}}{dT}(T - T_a) - q_1)(L_c)}{\boldsymbol{r_a} + \boldsymbol{r_s}} \tag{B.5}$$

As the longwave radiation term depends on T to the fourth power and q$_{sat}$ has an approximately exponential dependence on temperature, both the energy flux terms will increase with T faster than the linear functions used to approximate them. Consequently, the value of T$^*$ that is calculated in Equation B.3 will be higher than the correct value of T$^*$ that balances the energy budget and consequently, the outgoing longwave radiation and latent heat terms will be higher than they should be when they are diagnosed based on T$^*$. As the ground heat flux term is diagnosed to balance the energy budget and assure that the model conserves energy, this implies that the ground heat flux will consistently be lower than its correct value.

The errors in T$^*$ and in the ground heat term will be greatest where the difference between the air temperature and the calculated value of T$^*$ is greatest. This occurs in regions where there are extensive amounts of bare soil. Here, the grid cell surface roughness is substantially lower than in heavily vegetated regions. As the surface roughness governs the degree of sensible and latent heat transfer (which determine the strength of the coupling between surface temperature and air temperature), both of these heat fluxes are comparatively low in bare soil regions. The latent heat term is also strongly affected by the reduced moisture availability in deserts. Thus, in deserts, T$^*$ is often substantially higher than the air temperature and the ground heat flux is strongly reduced below its correct value, explaining the large difference between T$^*$ and T$_g$ in version 3.0Ea of the model, as shown in top panel of Figure B.9.

This strong difference between skin temperature and soil temperature is substantially reduced in the multilayer versions of the model as the second panel of Figure B.9 shows. In v. 3.0Eb of the model, the soil thermal conductivity again has no dependence on soil moisture content. The reduced temperature difference between

$T^*$ and $T_g$ is a consequence of the fact that the midpoint of the first soil layer and the surface are now separated by one-tenth of the distance as in v. 3.0Ea (ie. top soil layer thickness of 1.0 m vs. 0.1 m). As the soil temperature drops below the skin temperature due to the error in the energy budget code, the gradient between the soil and skin temperature for an equivalent soil-skin temperature difference in v. 3.0Eb is thus 10 times greater than in v. 3.0Ea. This soil-skin temperature gradient determines the degree to which the surface energy budget is out of balance when it is diagnosed at the initial guess value of $T_a$ and thus, the soil heat flux term more strongly acts to reduce $T^*$ below its erroneously high value, improving the agreement between soil and skin temperatures.

In versions 3.0Ec and 3.0E of the model, the soil thermal conductivity depends on soil moisture content. However, in most desert regions, the simulated soil moisture content remains consistently low with little annual variation (Figure B.1), so the assertion that soil thermal conductivity is relatively constant throughout the year is reasonable here. In these versions, the temperature difference between $T^*$ and $T_g$ is desert regions is higher than in v. 3.0Eb. This is a consequence of the revised soil thermal conductivity parameters. In v. 3.0Ea and 3.0Eb, the soil thermal conductivity is equal to 0.75 W/m·K, substantially higher than observed dry soil values. In v. 3.0Ec and 3.0E, dry soil thermal conductivity values range between 0.187 and 0.270 W/m·K and the low soil moisture contents in desert regions means that the actual soil thermal conductivity values are close to the dry soil value. As the soil thermal conductivity appears in the numerator of the soil heat flux term, a reduction in soil thermal conductivity would have the opposite impact on the soil heat flux error as decreasing the top soil layer thickness. Hence, as a result of using more realistic dry soil thermal conductivity values, the soil heat flux error is increased in desert regions. Nonetheless, the temperature difference between $T^*$ and $T_g$ in desert regions in v. 3.0Ec and 3.0E remains nearly an order of magnitude lower than in v. 3.0Ea.

Outside of desert regions, one would not expect the annual mean ground temperature ($\overline{T_g}$) to be equal to the annual mean skin temperature ($\overline{T^*}$), owing to variations in the ground thermal conductivity throughout the year. In v. 3.0Eb, the only such annual variation is due to the presence of lying snow which reduces the ground thermal conductivity in winter months, insulating the ground and inhibiting it from cooling during the winter months (leading to $\overline{T_g} > \overline{T^*}$) and from warming during the thaw season (leading to $\overline{T_g} < \overline{T^*}$). The alteration of ground thermal conductivity owing to the presence of lying snow was not represented in v. 3.0Ea. The second panel of

Figure B.9 suggests that winter insulation effect dominates with warmer ground temperatures most pronounced in the vicinity of the Tibetan Plateau and the Southern Rockies.

The second row, second column of Figure B.8a shows the changes in the annual range of soil temperature in the multilayer version of the model. In desert regions, the range in soil temperatures increases, a consequence of the fact that the soil temperatures now more closely align to the more variable surface temperatures. Elsewhere in the world, the range of soil temperatures is comparable or lower than in v. 3.0Ea. The lower soil temperature ranges are observed in the regions of the world which show the strongest range in soil temperature (top right panel, Figure B.8a). Here, the thermal inertial of the deeper soil column in v. 3.0Eb, serve to dampen the range in near surface soil temperature.

### B.3.2 Impact of New Soil Parameters

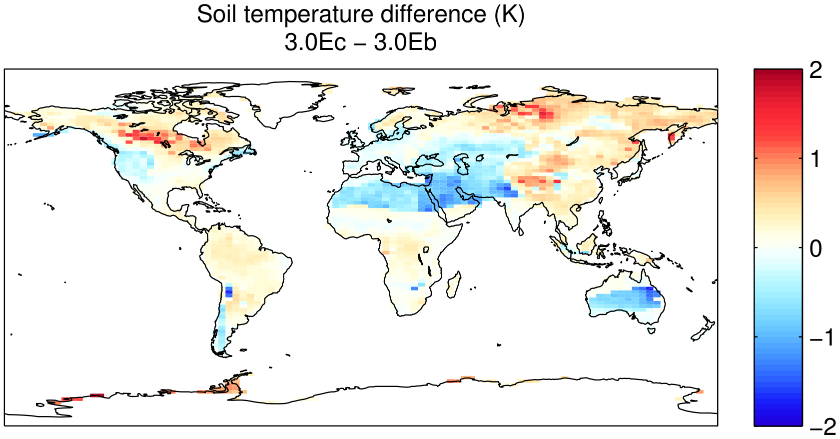


Figure B.10: Difference in annual mean soil temperature between version 3.0Ec and 3.0Eb (K).

Figure B.10 shows the difference in annual mean soil temperature between v. 3.0Ec (with new soil parameters) and 3.0Eb, clearly showing the aforementioned temperature change in desert regions associated with the new dry soil thermal conductivity values in this version of the model. Changes in mean soil temperature outside desert regions require more detailed analysis. In v. 3.0Ec, soil moisture influences both the thermal conductivity of the soil and the soil heat capacity. Both the thermal conductivity and heat capacity are increasing functions of soil moisture content. An increase in the soil thermal conductivity at certain times of the year would tend to cause soil layers to more rapidly respond to changes in the skin temperature. At the same, time an increased soil heat capacity owing to the presence of additional moisture in the soil increases the thermal inertia of soil layers, leading to an opposing effect in terms of the soil's ability to respond to surface temperature changes. The overall effect of the inclusion of the impact of soil moisture on thermal conductivity and heat capacity then depends on which of these two changes dominates and how the seasonality in soil thermal parameters corresponds to the seasonality in the skin temperature. To study changes in thermal characteristics of the soil over the course of the year, an 'effective' thermal diffusivity (k$_{eff}$, m$^2$/s) for the top meter of the soil was calculated. This is defined as the effective thermal conductivity for heat transport across the first four layers of soil (that span the top 1 m) divided by the average heat capacity of these layers:

$$k_{eff} = \frac{\frac{\sum_{n=1}^{4} \Delta z_n}{\sum_{n=1}^{4} \frac{\Delta z_n}{k_n}}}{\frac{\sum_{n=1}^{4} C_n \Delta z_n}{\sum_{n=1}^{4} \Delta z_n}} \tag{B.6}$$

The effective thermal conductivity term (numerator in equation B.6) is calculated by treating the layers as though they were resistors in series, and summing the resistances (inverses of the thermal conductivities) to find an overall thermal resistance across the top four layers. It is noted that layers rich in soil carbon tend to have particularly low thermal conductivities and heat capacities owing to their high porosity. In organic rich soils, given the form of the cumulative carbon function introduced in the previous chapter, it is the topmost layers that tend to have the lowest thermal conductivities. Although such layers are thinner than the deeper soil layers, they can have a disproportionate influence on the effective soil thermal conductivity and on the effective thermal diffusivity. The effective heat capacity for the top meter of soil is simply the depth-weighted average heat capacity of the layers.

An examination of the effective thermal diffusivity across the top four soil layers can explain most of the changes in the annual mean soil temperature between v. 3.0Eb and v.3.0Ec (Figure B.10). These changes can be understood in terms of a differences between the annual mean skin temperature ($\overline{T^*}$) and mean ground temperature ($\overline{T_g}$) as a result of seasonal variations in k$_{eff}$. The strength of these changes depends on three factors. Firstly, the degree to which k$_{eff}$ varies throughout the year (left panel,

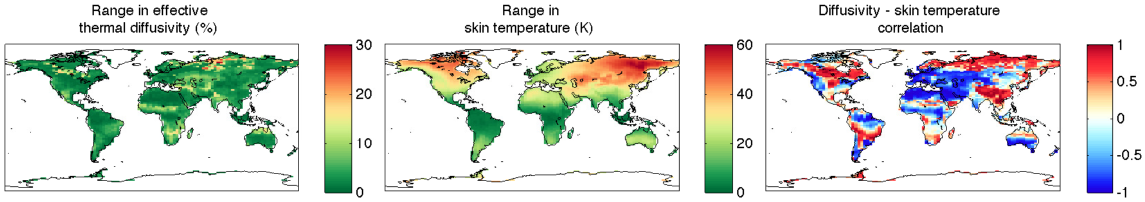


Figure B.11: Annual range in k$_{eff}$ (left panel) as a percentage of annual mean value, annual range in T$^*$ (center panel) and correlation between k$_{eff}$ and T$^*$ in v. 3.0Ec of the model

Figure B.11) determines the strength of seasonal variations in the soil's ability to equilibrate with the surface. The strongest variations in k$_{eff}$ are shown to occur in the tropics where there are strong seasonal soil moisture variations (top right panel, Figure B.1) as well as in regions of high organic soil content, where even a moderate increase in the soil moisture concentration near the surface can strongly increase k$_{eff}$. Variations in the thermal diffusivity will have little overall effect on soil temperature if the skin temperature is very stable throughout the year; hence, the degree to which skin temperatures vary (centre panel, Figure B.11) is a second determinant in the difference between between $\overline{T^*}$ and $\overline{T_g}$. Lastly, annual variation in k$_{eff}$ and of T$^*$ must be strongly correlated (or anti correlated) for there to to be a strong difference between annual mean soil temperature and annual mean skin temperature. The correlation (Cor) between k$_{eff}$ and T$^*$ in each grid cell (right panel, Figure B.11) is calculated based on the covariance (Cov) matrix for these variables:

$$Cor(k_{eff}, T^*) = \frac{Cov(k_{eff}, T^*)}{\sqrt{Cov(k_{eff}, k_{eff}) \cdot Cov(T^*, T^*)}} \tag{B.7}$$

To understand this last factor, consider the idealized situation in which the annual variation both in k$_{eff}$ and T$^*$ is sinusoidal in nature. Figure B.12 shows such a situation, displaying variation about the mean value in k$_{eff}$ and T$^*$ with both displayed in arbitrary units. In the first panel, both signals are in phase and are perfectly correlated (Cor(k$_{eff}$, T$^*$) = 1): the thermal diffusivity is highest when the skin temperature is highest, and the diffusivity is lowest when the skin temperature is lowest. Thus, the soil should more readily equilibrate to the skin temperature during the warmest months, leading to mean soil temperatures that are warmer than mean skin temperatures (and a warmer soils in v. 3.0Ec vs. 3.0Eb). In the second panel, k$_{eff}$

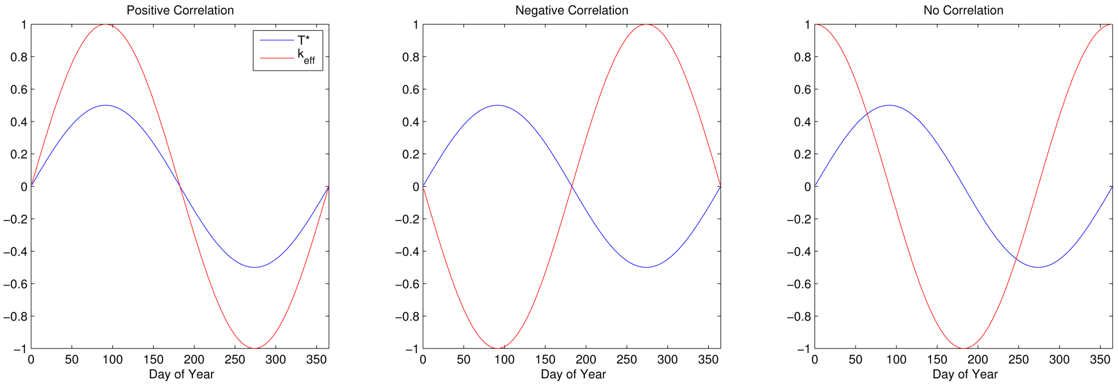


Figure B.12: Idealized correlation between effective thermal conductivity and skin temperature. In these plots, T* and $k_{eff}$ are plotted as variations relative to their mean value in normalized units. The plots are simply intended to show idealized patterns of variability in $k_{eff}$ and T*. The first panel shows a situation in which the sinusoidal variations are exactly in phase (Cor = 1), the second where the variations are $180^o$ out of phase (Cor = -1) and the third where they are $90^o$ out of phase (Cor = 0)

and T* are 180° out of phase and perfectly anti-correlated (Cor($k_{eff}$, T*) = -1. By the same logic, with this type of correlation, the soil temperature should most rapidly equilibrate in the coldest months leading to an annual mean soil temperature that is less than the annual mean skin temperature (cooler soils in v. 3.0Ec vs. 3.0Eb). Finally, the third panel shows the case where there is no overall correlation between $k_{eff}$ and T*. Here, the annual mean soil temperature should be exactly equal to the annual mean skin temperature as maximum and minimum $k_{eff}$ values correspond the the annual mean T* values.

The strongest variations in skin temperature tend to occur at high latitudes in North America and in mid-to-high latitudes in Central-Eastern Eurasia. Examining these regions in Figure B.10, one can see warming of the soil (in 3.0Ec vs. 3.0Eb) where there is a strong positive correlation between $k_{eff}$ and T*, and cooling where there is a strong negative correlation coefficient. Regions where there is a high organic soil content (and associated variability in $k_{eff}$) appear as 'hotspots' in these regions. The above analysis is a simplified description of the changes in v. 3.0Ec of the model (for instance, the form of $k_{eff}$ described above, assumes that heat transfer through the soil occurs exclusively from diffusion, with no heat transfer by the percolation of water), but the qualitative agreement between the trends in Figure B.11 and the soil

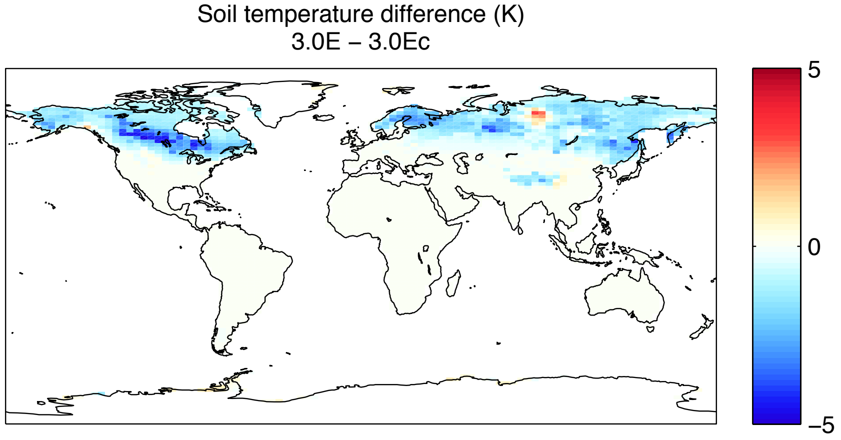


Figure B.13: Difference in annual mean soil temperature between version 3.0E and 3.0Ec (K).

temperature changes in Figure B.10 suggest that this analysis is reasonable.

The annual range in soil temperatures is reduced in this version of the model (second column, third row, Figure B.8a) as a consequence of the lower soil thermal conductivity values throughout and the higher heat capacity of the soil due to the inclusion of the heat capacity of water.

### B.3.3 Impact of Soil Freeze-Thaw Processes

Figure B.13 shows the difference in annual mean soil temperature between v. 3.0E (which allows for soil freeze-thaw processes) and v. 3.0Ec of the model. As anticipated, soil temperature differences are restricted to regions of permafrost and seasonal frost and result from three key changes:

1. **Changes in soil moisture storage**. The presence of ice in the soil strongly restricts the hydraulic conductivity. In regions of permafrost, this results in a year round increase in total soil moisture content and a reduction in the range of soil moisture throughout the year (Figure B.7) in 3.0E vs. 3.0Ec. In regions of seasonal frost, the annual range in soil moisture moisture increases as a consequence of the restriction of soil moisture movement during the winter months, and the annual mean total moisture content slightly increases.

2. **Changes in thermal conductivity/heat capacity of moisture**. The thermal conductivity of ice is roughly four times that of water while its volumetric

heat capacity is about half that of water. Given that the total soil moisture content in regions of permafrost is fairly constant throughout the year, the presence of ice in the winter months acts to strongly increase $k_{eff}$ through both the thermal conductivity effect (increases the overall thermal conductivity of soil in winter) and the heat capacity effect (decreases the heat capacity of soil in winter months). Although the moisture content in regions of seasonal frost is not stable throughout the year, the thermal conductivity is highest in winter months too owing to the increased moisture storage in the soil in winter months as well as the presence of ground ice in this season.

3. **Soil phase changes**. During periods when phase changes occur, the effective heat capacity of the ground is strongly increased as a result of the extra energy that the soil must lose (during freezing periods) and gain (during thawing period) associated with phase changes of water. The overall effect of this term on the annual mean soil temperature is difficult to assess a priori, as it depends on the degree of symmetry of soil freezing and thawing throughout the year and how well correlated soil freezing and thawing is with changes in skin temperature.

Figure B.13 indicates that overall, annual mean soil temperatures typically cool in v. 3.0E relative to v. 3.0Ec. As in the previous section, these results are analyzed in terms of variations in $k_{eff}$, $T^*$ and the correlation between these quantities. The annual range in $k_{eff}$ (left panel, Figure B.14) sharply increases in regions that experience soil-freeze thaw processes as a consequence of the very strong increase in $k_{eff}$ in the winter months relative to the summer months. The annual range in skin temperature (center panel, Figure B.14) remains very similar to v. 3.0Ec. Most interesting however, is the change in the correlation between $k_{eff}$ and $T^*$ (right panel, Figure B.14). Here, throughout the high latitudes of the Northern Hemisphere, $k_{eff}$ and $T^*$ are very strongly anti-correlated, reflecting the fact that $k_{eff}$ is highest in the winter when the skin temperature is the lowest. This is a sharp change from v. 3.0Ec wher,e throughout much of this region, $k_{eff}$ and $T^*$ were strongly positively correlated. Thus, the soil cooling between v. 3.0E and 3.0Ec is explained by the heightened thermal diffusivity in the winter months which allows soil temperatures to more rapidly equilibrate to the skin temperature during this season, causing the annual mean soil temperature to drop. As a consequence of the large amount of energy associated with soil water phase changes, the annual range in soil temperatures

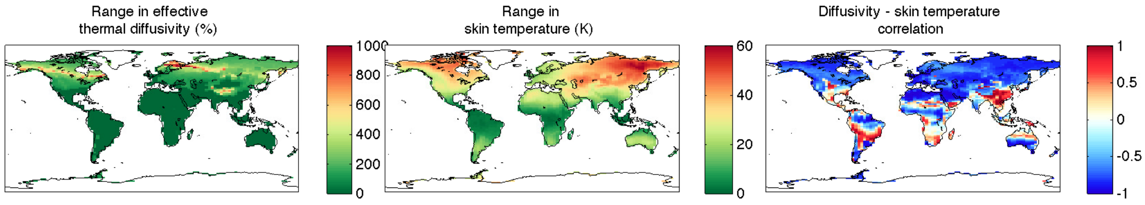


Figure B.14: Annual variation in $k_{eff}$ as a percentage of annual mean value (left panel), annual range in $T^*$ (center panel) and correlation between $k_{eff}$ and $T^*$ in version 3.0E of the model

in regions that experience such phase changes is reduced in 3.0E relative to 3.0Ec (second column, fourth row, Figure B.8a).

### B.3.4 Hovmoller Diagrams

As a means of summarizing the findings concerning changes in near surface soil moisture storage and temperature, a set of Hovmoller diagrams is now presented. The diagrams show how moisture and temperature vary with depth and with time of year and represent subsurface conditions at a representative high latitude cell in permafrost.

The high latitude grid cell is centered at 62.1°N, 127.8 °W in a region of permafrost (as simulated by v. 3.0E). The region experiences a strong precipitation maximum in the mid-summer although the soil remains fairly moist year-round (left column, Figure B.15, red curve, Figure B.16 ). The soil moisture levels reach a maximum in the early summer, beginning to rise in May-June as snow melt occurs and continuing to rise through to July as the region experiences its annual precipitation maximum. While the single layer v. 3.0Ea cannot capture the vertical soil moisture gradient, in this region of the world the gradient is fairly small in the top meter of the soil. As such, the Hovmoller diagrams for moisture in v. 3.0Ea and v. 3.0Eb resemble each other closely. A second consequence of this relatively weak vertical soil moisture gradient is that the evapotranspiration curves of v. 3.0Ea and v. 3.0E are both quite similar and track the precipitation rate fairly closely, with the exception of the winter months where snow cover reduces the evapotranspiration rate to zero.

The inclusion of new soil properties in v. 3.0Ec a results in an increase in the vertical soil moisture gradient through the impact of the increased hydraulic conductivity with the result that near surface soil moisture storage is reduced. Despite

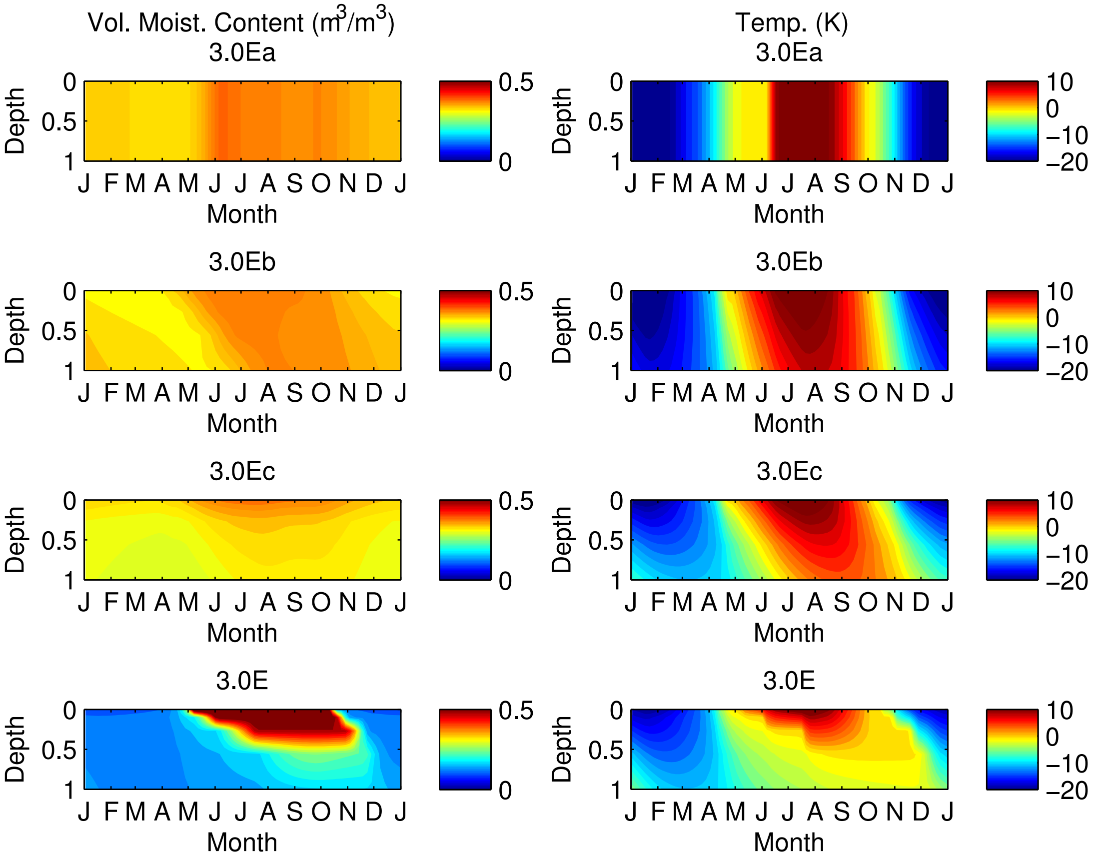


Figure B.15: Hovmoller Diagrams showing variations in soil moisture content and temperature in the top 1 m of soil for a high latitude grid cell. The soil moisture content is plotted in terms of volumetric soil moisture concentration (m$^3$ $H_2O$/m$^3$ soil)

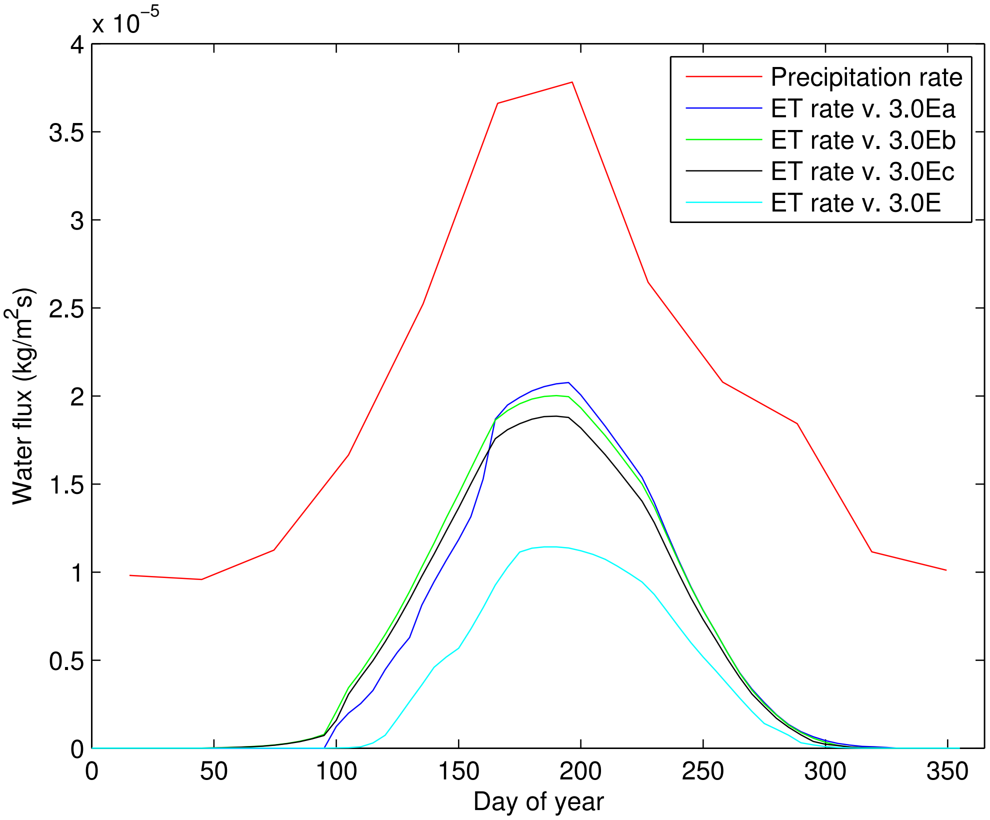


Figure B.16: Rates of precipitation and evapotranspiration in three model versions: high latitude grid cell. The precipitation field remains the same for all model versions.

this reduction in near surface soil moisture availability, the annual evapotranspiration timeseries in the gridcell remains fairly similar, decreasing slightly.

By far the largest impact in terms of soil moisture storage across the various versions comes in version 3.0E which allows for soil freezing and thawing to occur. Here, the liquid soil moisture content plummets sharply in the winter months in all soil layers. Soil moisture content rises in the summer months in a ‘stepped’ pattern. As the soil warms, thawing tends to occur strongly in one layer at a time. This is a consequence of the large amount of latent heat needed to thaw the ground which causes the thawing layer’s temperature to rise towards zero very slowly, producing a strong temperature gradient between the layer above (which has already thawed) and the currently thawing layer. Once this layer thaws, its temperature rapidly rises, causing a strong temperature gradient between this layer and the layer below, allowing that layer to thaw as well.

A second consequence of the presence of permafrost is the fact that the hydraulic conductivity of soil layers is strongly reduced when frozen moisture is present. This allows for pooling of a substantial amount of moisture near the surface such that the top soil layers remain close to saturation during the summer months - conditions representative of seasonal wetlands. These near-saturation conditions are absent in the moisture profiles of version 3.0Ec; this observation points to an important consequence of permafrost thaw in the model. Without the reduction in soil permeability accompanying frozen soil in the ground, such seasonal wetlands cannot be sustained in most pan-Arctic grid cells. This observation is addressed in Chapter 5 of the thesis. Lastly, it is noted that the evapotranspiration rate is strongly reduced when frozen soil is simulated. As of yet, the model does not simulate plant physiological adaptations to moisture stress and PFT root profiles remain identical globally. Thus, even though soil moisture availability in the upper soil layers strongly increases in the summer months, plant roots still have a certain density of roots in the deep layers which have very low moisture availability. Consequently the overall plant moisture availability is lowered as is the evapotranspiration rate. In actuality, plant rooting depths are generally shallower in regions of frozen soil and future versions of the model may include a representation of variable PFT rooting depth in response to moisture stress.

Next, the soil temperature profile is considered (right column, Figure B.15). One interesting feature of the annual temperature range in v. 3.0Ea is the month-long period from May to June where the soil temperature remains at the freezing point.

This is a consequence of the snow melt scheme in version 3.0Ea in which snow melt occurs strictly from heat transfer from the soil layer such that the soil temperature is constrained to never rise above the freezing point of water so long as lying snow is present. In subsequent versions of the model, snow melt occurs both from the ground as well as when the skin temperature exceeds the freezing point. As a consequence, snow melt begins to occur earlier in this version of the model (well before the soil temperature crosses 0 ° C). Other than this, the temperature time series of v. 3.0Ea and 3.0Eb actually agree fairly closely, although the latter shows a tilted isotherms as a consequence of the fact that deep soil temperatures do not instantaneously respond to changes in near surface temperature and there is a resulting phase shift between the top soil temperature and the temperature of deeper soil layers with the phase shift increasing with depth. The tilt of the isotherms is indicative of the degree of the phase lag and the soil thermal diffusivity. A weaker tilt indicates a smaller phase shift and higher diffusivity, while a stronger tilt is indicative of a greater phase shift and a lower diffusivity. This grid cell has a relatively high thermal diffusivity in v. 3.0Eb, such that there is a relatively low phase lag between soil layers. The inclusion of new soil thermal properties (v. 3.0Ec) increases this phase lag and reduces the soil temperature range at depth.

V. 3.0E which shows the strongest and most interesting changes in the soil thermal profile. The seasonal asymmetry in thermal diffusivity results in a substantially higher diffusivity in the winter months with the result that deep soil layers more rapidly respond to temperature changes in the surface layers, as opposed to the summer months where the lower thermal diffusivity results in the soil layers having weaker thermal coupling. The soil temperature profile shows a somewhat jagged profile during periods of the year when rapid phase changes occur. Interestingly, this jaggedness occurs only when phase changes are included. The jaggedness of the soil profile is a consequence of the fact that the soil column is discretized into a series of layers each of which tends to undergo phase changes one at a time, as previously described . The choice of 8 soil layers was chosen on the basis of the substantial computational cost when many more soil layers were added. While this jaggedness is clearly unphysical, it raises the interesting question of how substantially it impacts the broad pattern of soil temperature changes. To this end, a 'high soil resolution' version of the model was prepared (and described in more detail in Chapter 3) with four times as many soil layers and with each soil layer in v. 3.0E split into four sublayers of equal thickness. The model was spun up to equilibrium (taking nearly 5 months of

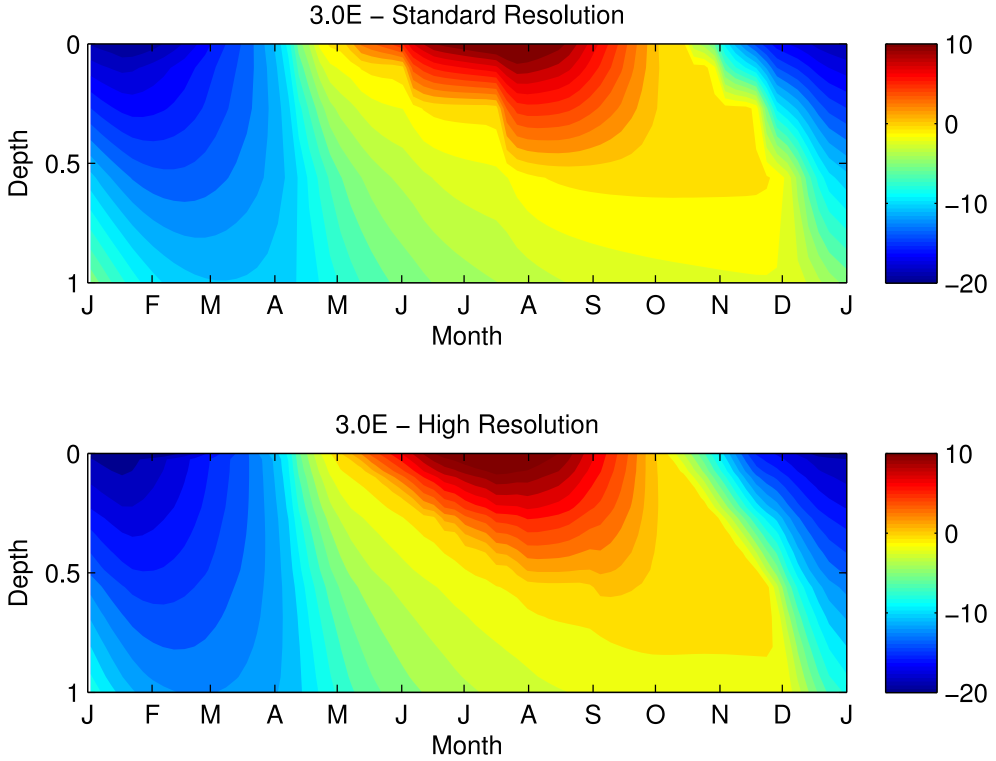


Figure B.17: Hovmoller diagram showing the annual soil temperature cycle in a grid cell that experiences soil freeze-thaw processes. Top panel: diagram for a low resolution version of the model; bottom panel: diagram for a high resolution version of the model.

spin up time, vs. 2 weeks in the 10-layer version) and the soil temperature profile for this grid cell was plotted agains the 8 layer version of the model. As shown in Figure B.17, the jaggedness of the soil temperature plot is substantially reduced in this high resolution version of the model, yet the overall range of soil temperatures at a particular depth remains quite similar with the exception that the warming pulse in the summer months extends to slightly deeper depths.

Other results from the high resolution version of the model are addressed in later sections of the model, but it is my opinion that given other biases in the model and uncertainties in input parameters, the use of a substantially higher resolution soil scheme is not justified in the study of permafrost in a coupled model given the extremely high extra computational cost that is required.

## B.4 Summary

In this chapter, UVic 3.0E, an offline version of the UVic climate model, was introduced in which the land surface scheme is driven by atmospheric variables derived from the ECMWF ERA-40 reanalysis product. This configuration of the model was used to study the impact of modifications made to the land surface scheme for the purpose of permafrost simulation. The offline configuration of the model was used for this purpose to facilitate analysis without requiring an investigation into differences due to feedbacks between the land surface and other climate system components. Key impacts observed as changes were applied to the land surface scheme were as follows.

**Switch from single soil layer configuration to multiple soil layers (v. 3.0Ea to v. 3.0Eb)**:

- Near surface soil moisture storage increases in most regions due to simulation of the effects of infiltration of moisture through the soil column (v. 3.0Eb) as opposed to immediate drainage of moisture from the base of the single soil layer.

- There is a strong increase in soil temperatures in desert regions in the multilayer version of the model as a consequence of a limitation in the numerical scheme used to determine the skin temperature in the model's surface energy budget code. There is also a slight warming of soil temperatures in regions subject to lying snow due to the thermal insulation effect of snow cover that was not represented in v. 3.0Ea. The annual range in near surface soil temperatures also tends to decrease due to the additional thermal inertia of deeper soil layers.

**Inclusion of revised soil parameters and dependence of soil thermal parameters on soil moisture content. (v. 3.0Eb to v. 3.0Ec)**:

- Changes were made to the soil porosity, saturated hydraulic conductivity, saturated matric potential and Clapp-Hornberger coefficient. The overall pattern of changes to soil moisture depends in a complicated manner on gridcell-specific changes to each of these four parameters, changes to which do not add linearly in terms of the overall effect on soil moisture. The overall effect is a reduction in near surface soil moisture content nearly everywhere.

- In v. 3.0Ec, both the soil thermal conductivity and heat capacity vary with soil moisture content and thus vary through time. The overall effect of variability in thermal conductivity and heat capacity on mean soil temperature may be understood in terms of how the effective thermal diffusivity in a grid cell varies with time throughout the year. In regions where there is a strong positive (negative) correlation between variations in the skin temperature and the thermal conductivity, the soil temperature tends to warm (cool) in v. 3.0Ec relative to 3.0Eb with the strength of the warming (cooling) dependent on the degree to which the effective thermal conductivity and skin temperature vary throughout the year.

- The revised soil parameters generally act to reduce the soil thermal conductivity and the soil heat capacity. Consequently the annual range in soil temperature values is generally reduced in v. 3.0Ec relative to v. 3.0Eb.

**Inclusion of soil freeze-thaw processes (v. 3.0Ec to v. 3.0E)**

- There are no changes to soil moisture storage nor soil temperatures outside of regions subject to soil freezing and thawing, as one would expect given that there are no atmosphere / land surface feedbacks represented in these model configurations.

- Of all the changes outlined above, the inclusion of soil freeze/thaw processes has the largest impact on soil moisture content. The annual mean *liquid* soil moisture content strongly decreases in regions of permafrost and seasonally frozen ground, but the annual range in liquid moisture content strongly increases due to the pooling of liquid moisture near the surface during the thaw season. Also in these regions there is a strong increase in the total (frozen + liquid) moisture content and the annual range in total moisture content typically decreases due to a restriction in moisture drainage.

- Inclusion of soil freeze/thaw processes typically results in a cooling of near surface soil temperatures, largely as a consequence of the much higher thermal conductivity (and thermal diffusivity) in the winter months owing to the presence of ice in the soil. This allows the ground to more rapidly equilibrate to skin temperatures at this time of the year, resulting in cooling relative to v. 3.0Ec. As a consequence of the large amount of energy associated with soil water phase changes, the annual range in soil temperatures in regions of permafrost and seasonally frozen ground is also reduced in v. 3.0E..