Diagnosing Present and Future Permafrost from Climate Models

ANDREW G. SLATER
National Snow and Ice Data Center, Cooperative Institute for Research in Environmental Sciences,
University of Colorado Boulder, Boulder, Colorado

DAVID M. LAWRENCE
National Center for Atmospheric Research, Boulder, Colorado

(Manuscript received 13 June 2012, in final form 24 January 2013)

ABSTRACT

Permafrost is a characteristic aspect of the terrestrial Arctic and the fate of near-surface permafrost over the next century is likely to exert strong controls on Arctic hydrology and biogeochemistry. Using output from the fifth phase of the Coupled Model Intercomparison Project (CMIP5), the authors assess its ability to simulate present-day and future permafrost. Permafrost extent diagnosed directly from each climate model’s soil temperature is a function of the modeled surface climate as well as the ability of the land surface model to represent permafrost physics. For each CMIP5 model these two effects are separated by using indirect estimators of permafrost driven by climatic indices and compared to permafrost extent directly diagnosed via soil temperatures. Several robust conclusions can be drawn from this analysis. Significant air temperature and snow depth biases exist in some model’s climates, which degrade both directly and indirectly diagnosed permafrost conditions. The range of directly calculated present-day (1986–2005) permafrost area is extremely large ($\sim 4-25 \times 10^6$ km$^2$). Several land models contain structural weaknesses that limit their skill in simulating cold region subsurface processes. The sensitivity of future permafrost extent to temperature change over the present-day observed permafrost region averages ($1.67 \pm 0.7 \times 10^6$ km$^2/\text{C}$) but is a function of the spatial and temporal distribution of climate change. Because of sizable differences in future climates for the representative concentration pathway (RCP) emission scenarios, a wide variety of future permafrost states is predicted by 2100. Conservatively, the models suggest that for RCP4.5, permafrost will retreat from the present-day discontinuous zone. Under RCP8.5, sustainable permafrost will be most probable only in the Canadian Archipelago, Russian Arctic coast, and east Siberian uplands.

1. Introduction

Permafrost forms an integral part of the terrestrial Arctic system as well as the “third pole” of the Tibetan Plateau. A permanently frozen soil matrix and an active layer that experiences annual freeze-thaw cycles can have a particularly large impact on hydrologic fluxes (Rouse et al. 1997; Yoshikawa and Hinzman 2003; Smith et al. 2005) as well as upon flora (Lloyd et al. 2003; Christensen et al. 2004), fauna (Clark et al. 1997; Wrona et al. 2006), geomorphology (Rowland et al. 2010), and biogeochemical cycling (Schuur et al. 2008, 2009). Socioeconomic impacts of permafrost degradation may also prove costly (Larsen et al. 2008). Permafrost degradation has recently been observed in many locations including Sweden (Akerman and Johansson 2008), Canada (Fortier et al. 2007; Thibault and Payette 2009), Alaska (Jorgenson et al. 2006), and Tibet (Cheng and Wu 2007). Ground temperatures are warming across the majority of permafrost regions (Romanovsky et al. 2010) and occasionally large rapid changes are seen (Oberman 2008). Although observations of ground temperature are limited in space, there is a strong indication that permafrost is warming and thawing across much of the permafrost domain in response to recent climate change, which raises the question—how much and how fast will permafrost thaw over the course of the twenty-first century if the climate continues to warm?

Currently, the best tools for peering into Earth’s potential future climate are coupled global climate models (CGCMs) or “Earth system models.” These models
typically constitute fully coupled atmosphere, ocean, land, sea ice, and often biogeochemical components. The fifth phase of the Coupled Model Intercomparison Project (CMIP5; http://cmip-pcmdi.llnl.gov/cmip5/) has brought together CGCMs from international groups to perform a series of coordinated experiments that project future climate based on scenarios of anthropogenic influences. Based on results from climate models, several prior efforts have performed assessments of future permafrost in limited regions (Sazonova and Romanovsky 2003; Zhang et al. 2008) or for single climate models (Stendel and Christensen 2002; Lawrence and Slater 2005; Saito et al. 2007; Lawrence et al. 2008) or for periods up to 2050 (Anisimov et al. 1997; Dankers et al. 2011). While there is agreement across these studies that permafrost degradation will occur with a warming climate, the extent and rate of the process is highly debated.

Our aim in this study is to assess the present state and future trajectory of permafrost within the CMIP5 CGCMs. We also strive to isolate the influence of the simulated model climate from that of each CGCM’s embedded land surface model in terms of the diagnosed present-day conditions and future fate of permafrost. In the following sections we introduce the data and methods of diagnosing permafrost, both directly from soil temperatures and indirectly using simplified models. We then assess the CMIP5 model climates over present-day permafrost regions, compare different diagnoses of permafrost, suggest reasons for particular model results, and, lastly, show where permafrost degradation is likely to occur under climate warming scenarios.

2. Data

Monthly-mean data are used from reanalysis projects and the collection of CMIP5 climate models. Only the Northern Hemisphere is considered. All data are from grid boxes with greater than 30% land fraction. Glacier points are removed from the analysis and land fractions are accounted for in coastal regions. We use as many models as possible for each portion of our analysis, thus the number used can differ as a result of data availability.

a. Reanalysis data

As a test of the permafrost diagnostics in the present-day climate, four recent reanalysis products, the European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis Interim (ERA-Interim; Dee et al. 2011), the National Aeronautics and Space Administration (NASA) Modern-Era Retrospective Analysis for Research and Applications (MERRA; Rienecker et al. 2011), the National Oceanic and Atmospheric Administration (NOAA) Modern-Era Retrospective Analysis for Research and Applications (MERRA; Rienecker et al. 2011), the National Oceanic and Atmospheric Administration (NOAA) Climate Forecast System Reanalysis and Reforecast (CFSR; Saha et al. 2010), and the Japanese 25-yr Reanalysis (JRA-25; Onogi et al. 2007), were used. These reanalyses start at the beginning of the satellite period, 1979, and are updated to 2011 and beyond.

b. Climate model data

Historic and future simulations by coupled climate models were obtained from the CMIP5 archive. Complete model expansions are provided in Table 1 and
a summary of the models included in this study is given in Table 2. We use the CMIP5 simulation designated as historical, and the representative concentration pathway (RCP) cases RCP2.6, RCP4.5, RCP6.0, and RCP8.5. The historical simulations run from 1850 or 1860 through to the end of 2005 and are forced with observed changes in atmospheric constituents (CO2, O3, etc.) due to anthropogenic or volcanic forcing. The RCP simulations apply an expected increase in radiative forcing; for example, RCP8.5 assumes an extra 8.5 W m−2 forcing by 2100. RCP cases go from 2006 to 2099 and were stitched to the end of their respective historical run. Not all models provided required variables for all simulations, particularly for noncore CMIP5 cases (e.g., RCP2.6 and RCP6). Full details of the CMIP5 simulations are available from Taylor et al. (2012). Model variables used here are surface air temperature (TAS; equivalent to 2-m temperature), snow mass (SNW), snow depth (SND), and soil temperatures (TSL).

When diagnosing permafrost directly from modeled soil temperatures and making comparisons to simplified models (section 3b), we use only the first ensemble member reported by each group (i.e., simulation r1i1p1) and each model maintains its native grid. For present-day assessment of model climates (section 4a) and projections based on climatic change (see section 5), we 1) averaged climate variables across all ensemble members (up to 10) for each model and 2) interpolated data onto a common grid [that of the Community Climate System Model, version 4.0 (CCSM4.0), 0.9° × 1.25°, the highest-resolution model available] to allow for homogenous comparison.

c. Observations: The IPA map

The International Permafrost Association (IPA) map (Heginbottom et al. 1993; Brown et al. 1997) is perhaps the best available data of permafrost distribution at the global scale. Data used in its compilation date from 1960 to 1993; in some marginal permafrost regions, degradation may have since occurred.

For purposes of comparison, the IPA data were transformed from its 25-km equal-area Special Sensor Microwave Imager (SSM/I) Earth (EASE) grid onto the common CCSM4.0 grid by area weighting the occurrence of continuous, discontinuous, isolated, or sporadic permafrost. Models currently cannot represent subgrid-scale permafrost distributions, hence, in this study we only use the discontinuous and continuous zones in all analysis (i.e., where we expect 50% or more of the ground to be underlain by permafrost) and exclude glaciers. We refer to this 16.2 × 106 km2 region (blue areas in Fig. 1) as the “present-day permafrost area.” This model equivalent permafrost area differs from those reported by Zhang et al. (2000) but is more appropriate for model evaluation.

3. Permafrost diagnostic methods

Permafrost can be diagnosed directly from the CMIP5 modeled soil temperatures or indirectly using simplified permafrost models. Modeled soil temperatures are a function of both the simulated surface climate and the ability of the land model to represent permafrost. To separate these two factors and give insights into causes of particular CMIP5 model results we make use of simplified permafrost models that are driven primarily by climate variables.

a. Direct diagnosis of permafrost from modeled soil temperature

For those models that provide soil temperatures, “near-surface” permafrost can be diagnosed as such: if soil at a depth within 3.5 m of the surface (based on the lower boundary of a model’s soil layers) maintains a temperature of 0°C or less for the present and prior year, it is considered to contain permafrost. This definition (designated here as TSL) was chosen because active layers rarely reach this depth and it is within the range of most model soil depths. There is a large range of complexity in land surface models within CMIP5; for example, soil/ground column depths span 3–47 m, with 3–23 soil layers (Fig. 2). Horizontal resolutions are similarly variable, ranging from 2.8° × 2.8° to 0.9° × 1.25° (Table 2).

An alternative diagnosis of permafrost sustainability is the annual mean soil temperature (MST) at some depth, say 3.5 m, below the surface. If MST is below 0°C and assuming constant soil heat capacity, it is suggestive of permafrost at some deeper depth. Such a diagnostic can ameliorate issues created by coarse soil column discretization within land models. TSL and MST diagnose extremely similar permafrost areas except in the case of three models, which show a substantial difference: CanESM2, GISS-E2-R, and MPI-ESM-LR [discussed further in section 4b(1)]. Averaged across the present-day permafrost region, modeled values of MST differ greatly (see Table 2). However, MST is not an explicit definition of permafrost nor does it inform about the upper-soil state, which is important for hydrology, biology, and biogeochemical cycling.

b. Indirect diagnosis via simplified permafrost models

We apply two diagnostic models—the surface frost index (SFI) and the Kudryavtsev method (KUD)—both described below. These models require climatic inputs of annual maximum temperature, annual minimum
TABLE 2. Model characteristics. Here F/R refers to a force-restore scheme is used in which there are only two layers (i.e., daily and seasonal thermal stores). Composite-SG refers to a shifting vertical grid of snow into soil is used in the thermal calculation depending on snow depth.

<table>
<thead>
<tr>
<th>Model name</th>
<th>MST: Soil temperature (°C)*</th>
<th>Snow depth factor (m)*</th>
<th>Resolution (lat × lon)</th>
<th>Land model</th>
<th>Snow scheme structure**</th>
<th>No. soil layers</th>
<th>Depth of soil (m)</th>
<th>Model references</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCC-CSM1</td>
<td>−8.63</td>
<td>−0.18</td>
<td>0.408</td>
<td>2.8 × 2.8</td>
<td>BCC_AVIM1.0</td>
<td>Multilayer</td>
<td>10</td>
<td>3.43; Ji (1995); BCC (2012)</td>
</tr>
<tr>
<td>CanESM2</td>
<td>−7.40</td>
<td>−0.71</td>
<td>0.266</td>
<td>2.8 × 2.8</td>
<td>CLASS2.7</td>
<td>Bulk layer</td>
<td>3</td>
<td>4.10; Verseghy (2000)</td>
</tr>
<tr>
<td>CCSM4.0</td>
<td>−9.20</td>
<td>−1.49</td>
<td>0.442</td>
<td>0.9 × 1.25</td>
<td>CLM4</td>
<td>Multilayer</td>
<td>15</td>
<td>43.74; Oleson et al. (2010); Lawrence et al. (2011)</td>
</tr>
<tr>
<td>CNRM-CM5</td>
<td>−10.43</td>
<td>—</td>
<td>0.308</td>
<td>1.4 × 1.4</td>
<td>SURFEX</td>
<td>Bulk layer F/R</td>
<td>—</td>
<td>Boone et al. (2000); LeMoigne et al. (2009)</td>
</tr>
<tr>
<td>CSIRO-Mk3.6.1</td>
<td>−10.17</td>
<td>—</td>
<td>0.211</td>
<td>1.875 × 1.875</td>
<td>CSIRO</td>
<td>Multilayer</td>
<td>5</td>
<td>4.60; Gordon et al. (2002)</td>
</tr>
<tr>
<td>GFDL-ESM2M</td>
<td>−7.37</td>
<td>−7.57</td>
<td>0.319</td>
<td>2.0 × 2.5</td>
<td>LM3</td>
<td>Multilayer</td>
<td>23</td>
<td>10.00; Dunne et al. (2012)</td>
</tr>
<tr>
<td>GISS-E2-H</td>
<td>−7.96</td>
<td>—</td>
<td>0.390</td>
<td>2.0 × 2.5</td>
<td>Model II-LS</td>
<td>Composite</td>
<td>6</td>
<td>3.50; Rosenzweig and Abramopoulos (1997)</td>
</tr>
<tr>
<td>GISS-E2-R</td>
<td>−8.05</td>
<td>−5.04</td>
<td>0.352</td>
<td>2.0 × 2.5</td>
<td>Model II-LS</td>
<td>Composite</td>
<td>6</td>
<td>3.50; Rosenzweig and Abramopoulos (1997)</td>
</tr>
<tr>
<td>HadCM3</td>
<td>−11.41</td>
<td>—</td>
<td>0.248</td>
<td>2.24 × 3.75</td>
<td>MOSES2.2</td>
<td>Implicit</td>
<td>4</td>
<td>3.00; Essery et al. (2001)</td>
</tr>
<tr>
<td>HadGEM2-CC</td>
<td>−11.62</td>
<td>−8.83</td>
<td>0.249</td>
<td>1.24 × 1.875</td>
<td>MOSES2.2</td>
<td>Implicit</td>
<td>4</td>
<td>3.00; Essery et al. (2001)</td>
</tr>
<tr>
<td>HadGEM2-ES</td>
<td>−10.40</td>
<td>−7.53</td>
<td>0.251</td>
<td>1.24 × 1.875</td>
<td>MOSES2.2</td>
<td>Implicit</td>
<td>4</td>
<td>3.00; Essery et al. (2001)</td>
</tr>
<tr>
<td>INM-CM4</td>
<td>−8.40</td>
<td>−4.55</td>
<td>0.256</td>
<td>1.5 × 2.0</td>
<td>—</td>
<td>Bulk layer</td>
<td>20</td>
<td>15.00; Volodin and Lykosov (1998a,b)</td>
</tr>
<tr>
<td>MIROC5</td>
<td>−8.17</td>
<td>−6.74</td>
<td>0.286</td>
<td>1.4 × 1.4</td>
<td>MATSIRO</td>
<td>Multilayer</td>
<td>6</td>
<td>14.00; Takata et al. (2003)</td>
</tr>
<tr>
<td>MIROC-ESM</td>
<td>−6.68</td>
<td>−4.12</td>
<td>0.340</td>
<td>2.8 × 2.8</td>
<td>MATSIRO</td>
<td>Multilayer</td>
<td>6</td>
<td>14.00; Takata et al. (2003)</td>
</tr>
<tr>
<td>MPI-ESM-LR</td>
<td>−8.75</td>
<td>−7.67</td>
<td>0.296</td>
<td>1.875 × 1.875</td>
<td>ECHAM5</td>
<td>Composite-SG</td>
<td>5</td>
<td>9.58; Roeckner et al. (2003)</td>
</tr>
<tr>
<td>MRI-CGCM3</td>
<td>−9.60</td>
<td>−3.16</td>
<td>0.312</td>
<td>1.125 × 1.125</td>
<td>HAL</td>
<td>Multilayer</td>
<td>14</td>
<td>10.00; Yukimoto et al. (2012)</td>
</tr>
<tr>
<td>NorESM1-M</td>
<td>−10.80</td>
<td>−3.24</td>
<td>0.444</td>
<td>1.875 × 2.5</td>
<td>CLM4</td>
<td>Multilayer</td>
<td>15</td>
<td>43.74; Oleson et al. (2010); Lawrence et al. (2011)</td>
</tr>
<tr>
<td>Reanalyses</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CFSRR</td>
<td>−8.18</td>
<td>—</td>
<td>0.421</td>
<td>0.312 × 0.312</td>
<td>Noah</td>
<td>Composite</td>
<td>4</td>
<td>4.00; Ek et al. (2003); Saha et al. (2010)</td>
</tr>
<tr>
<td>ERA-Interim</td>
<td>−8.43</td>
<td>—</td>
<td>0.300</td>
<td>0.7 × 0.7</td>
<td>TESSEL</td>
<td>Bulk layer</td>
<td>4</td>
<td>2.89; Viterbo et al. (1999)</td>
</tr>
<tr>
<td>JRA-25</td>
<td>−8.26</td>
<td>—</td>
<td>0.281</td>
<td>1.125 × 1.125</td>
<td>SiB</td>
<td>Composite F/R</td>
<td>—</td>
<td>Sellers et al. (1986); Sato et al. (1989)</td>
</tr>
<tr>
<td>MERRA</td>
<td>−8.37</td>
<td>—</td>
<td>0.322</td>
<td>0.50 × 0.66</td>
<td>Catchment</td>
<td>Multilayer</td>
<td>6</td>
<td>10.00; Koster et al. (2000)</td>
</tr>
</tbody>
</table>

* Mean over present-day continuous and discontinuous permafrost area for years 1986–2005, if data were available.

** According to scheme of Slater et al. (2001).
temperature, and mean winter snow depth to generate an annual sinusoidal steady-state climate. The models give equilibrium-type solutions, hence these three climate inputs are averaged for the 20 years prior to the year for which diagnostics are presented (e.g., the year 2000 uses data averaged over 1981–2000). As these models provide equilibrium solutions, the disappearance of permafrost based on their metrics does not necessarily mean that deep permafrost will have completely degraded from that area at the time of diagnosis; it only means that permafrost is no longer sustainable under those climatic conditions. These methods have been used to make inferences about near-surface permafrost (Anisimov and Nelson 1996), but they are underpinned by steady-state assumptions.

The winter period for computing snow depth is defined as the time during which air temperature is below 0°C. Consequently, as temperatures warm, the length of the snow season also becomes shorter. While the annual temperature cycle is sinusoidal, snow depth is seasonally asymmetric because of the accumulation process. To obtain a meaningful average snow depth that reflects its impact as ground insulation, we weight the depth of snow for each month by the degree to which the temperature that month is below freezing. That is, for all months \( i \) with temperature \( T \) below freezing for the water year October to September, mean snow depth \( D \) is given as

\[
D = \frac{\sum_i D_i \times T_i}{\sum_i T_i} \quad \text{for } T < 0^\circ C. \tag{1}
\]

Each month snow water equivalent (SWE) and snow depth were used to compute snow density, which is needed for estimating snow thermal conductivity. The mean winter density was weighted in the same way as depth. All CMIP5 models provide SWE data, while some (e.g., CSIRO and HadGEM2) omit snow depth, in which case a snow density of 250 kg m\(^{-3}\) was used for conversion. The CFSRR snow depth is an analyzed field but uses an unrealistic constant density of 100 kg m\(^{-3}\) so we again applied 250 kg m\(^{-3}\).

1) SURFACE FROST INDEX

The surface frost index (SFI) as described by Nelson and Outcalt (1987) is a climate-based index:

\[
SFI = \frac{\sqrt{DDF^*}}{\sqrt{DDF^*} + \sqrt{DDT}}. \tag{2}
\]
where DDF* is the sum of freezing degree days modified for snow insulation, and DDT is the sum of thawing degree days computed from the sinusoidal climate. It uses no information about the surface state (soils, vegetation, etc.), but provides an idea of where the climate would be conducive to the existence of permafrost. The only departure we make from the original SFI method is the computation of mean winter snow depth as given above. Thermal conductivity and heat capacity of snow is computed as per the original paper as a function of density. The SFI ranges from 0.0 to 1.0, with values of 0.5–0.6 indicating sporadic permafrost, 0.60–0.67 indicating extensive (akin to discontinuous) permafrost, and values above 0.67 taken as continuous permafrost.

2) KUDRYAVTSEV METHOD

The Kudryavtsev method (KUD) (Kudryavtsev et al. 1977) is an analytic model that produces a steady-state (equilibrium) solution under the assumption of constant sinusoidal air temperature forcing with a mean winter snow cover and damping from organic matter or vegetation. Theoretical derivation of the model is given in Yershov (1998, ch. 1) with an additional excellent description available in Sazonova and Romanovsky (2003) and Romanovsky and Osterkamp (1997). This method also forms the basis for the State Hydrometeorological Institute (SHI) permafrost model (Anisimov et al. 1997).

The required spatially explicit mineral soil textures (% sand, silt, and clay) are taken from Zobler (1986) and computation of thermal properties follow Farouki (1981). Soil saturations are derived as a function of the Water and Global Change (WATCH) soil moisture estimates (Harding et al. 2011) and are kept constant. Excess soil ice is not accounted for, nor is the impact of ground subsidence.

The thickness of the organic mat (considered a combination of organic material in the soil, mosses, and ground cover vegetation) was assumed to be a maximum of 75 cm and scaled to the density of soil organic matter distribution given in the Global Soil Data Task (2000); hence over most of the Arctic permafrost region the prescribed organic mat is 10–25 cm thick. Thermal diffusivity of the organic mat under both frozen and thawed conditions is somewhat subjective as it is dependent upon its composition, but we follow the values used by Anisimov et al. (1997). The KUD model has two primary outputs, active layer thickness and the mean annual temperature at the top of the permafrost, which can be used to indicate if permafrost is sustainable. When forced with reanalysis data we produce permafrost temperatures in eastern Siberia similar to those in Sazonova and Romanovsky (2003).

4. Analysis of permafrost in CMIP5 models

a. Climate of the present-day permafrost region

The two most important climatic variables for permafrost are air temperature and snow depth. We use the reference period of 1986–2005 to perform a climate evaluation over the area where the IPA map suggests that continuous and discontinuous permafrost exists. Data were interpolated to the CCSM4.0 grid and all averages are area weighted.

Climate means are given in Table 2, but in Fig. 3 we use the mean of the four reanalysis products as a reference and show departures for each model. Out of 17 models with data for this period, 8 models have a cold
bias of 0.5°C or more, with 5 having a 2°C or colder bias. Only 3 models have a warm bias of 0.5°C or more (Fig. 3a). When viewed in terms of seasonality (Figs. 3b,c), the majority of the discrepancy between the models and the reanalyses occurs in the cold winter months and fits with the difficulty models have in modeling stable boundary layer conditions (ECMWF-GABLS 2012). The range of mean annual minimum temperature averaged over the permafrost domain is greater than 10°C across the CMIP5 models, while their mean annual maximum only spans 5°C. The reanalyses also display a greater range in winter. BCC-CSM1.1, CSIRO, the Hadley models, and NorESM1-M all have cold biases throughout the year while CanESM2, MIROC5, and MIROC-ESM are warmer throughout the year. For normalized winter snow depths (Fig. 3d), the reanalyses lack agreement despite two of them (ERA-Interim and CFSR) being analyzed (assimilated) fields; however, their mean provides a reference. Ten of the models have an absolute anomaly of 0.05 m or more and most models have too little snow compared to the reference. CCSM4.0 and NorESM1-M (which includes the atmosphere and land components from CCSM4.0) show a high snow bias, which degrades the skill of modeling permafrost, particularly in certain regions (see Lawrence et al. 2012). The same problem afflicts BCC-CSM1.1. A normalized snow depth anomaly of 0.1 m will impact ground temperatures given that much of the Arctic and Tibet have peak snow depths of less than 0.5 m (Brown and Brasnett 2010) and, importantly, the ability of snow to insulate the ground from cold winter temperatures is nonlinear at shallow depths (Zhang 2005; Lawrence and Slater 2010).

b. Simulation of present and future permafrost by CMIP5 models

For the 13 models that provided soil temperatures, we compare the range of permafrost areas diagnosed via both TSL and SFI for 1900–2099 (using r1i1p1; Fig. 4a) under RCP8.5. In 6 of the 13 instances, SFI gives greater starting (1900) and ending (2099) area; however, four instances do the opposite with TSL having greater area, there are two instances where TSL is within the bounds of SFI, and there is one instance where TSL and SFI have the same minimum but TSL had a larger maximum. As defined by TSL, the ability of many CMIP5 models to directly simulate permafrost is quite poor. For example, across the 13 CMIP5 models in Fig. 4a, the range of maximum permafrost areas since 1900 via TSL is 5–27 × 10⁶ km². In contrast, this range is only half as large via SFI at 11–21 × 10⁶ km². The fact that TSL produces both larger and smaller permafrost areas than respective SFI results suggests that the land models have more effect on the overall range of results than do climate discrepancies; for example, compare the results of the BCC-CSM1.1 and GFDL models in Fig. 4a. Compared to IPA map observations, all CMIP5 model climates overestimate permafrost area on the Tibetan Plateau via SFI and KUD for 1986–2005, and all but two models overestimate this area using the TSL diagnostic (not shown). Based on the IPA map, the Tibetan Plateau contains less than 7% of Northern Hemisphere continuous and discontinuous permafrost area [but more recent data suggest that discontinuous permafrost may be slightly more prevalent (Yang et al. 2010)]. Permafrost changes in the Arctic are therefore of greater consequence to the global climate.

Recall that the large variation in permafrost extent computed by TSL across the CMIP5 models is a function of both climate and land model differences while SFI differences are driven purely by climate differences. The disparity in permafrost extent between the two methods therefore suggests that with regard to assessing permafrost, the role of the land surface model can be as significant as that of the CGCM’s simulated surface climate. The relationship between the influences of climate compared to the land model can be seen in Fig. 4b where we show SFI area against TSL area for each year and for each model. Note that time is not an axis in Fig. 4b, but permafrost area decreases over time with warming (in general for each model/scenario, time can be thought of as running along each line from upper right to lower left). Further, we emphasize that this is a comparative analysis and the ideal trajectory within this two model space is unknown. Nonetheless several points can be inferred from this figure. First, the influence of climate biases is demonstrated by the large range of SFI permafrost area for the year 2000 (10.5–19.5 × 10⁶ km², shown by dots in Fig. 4b) compared to the observed area of about 16 × 10⁶ km². Second, where permafrost area is greater than 2 × 10⁶ km² for either method, there is a very linear relationship between SFI and TSL areas. Below 2 × 10⁶ km² nonlinear behavior occurs as permafrost persists only in extremely cold regions or may cross thresholds in the models as warming becomes greater. Third, despite the differences in climate trajectories between various RCP’s, the SFI to TSL relationship remains similar for individual models. The linear relation prompts us to examine the sensitivity of permafrost to temperature change over the permafrost region [section 4b(2)] while the consistent differences between permafrost area as diagnosed by TSL and SFI prompts inquiry into the influence of the land model physics relative to the surface climate simulation [section 4b(1)].

1) IMPACT OF THE LAND SURFACE MODEL

CMIP5 models with their trace well above the 1:1 line in Fig. 4b (e.g., GFDL-ESM2M, HadGEM2, MIROC)
FIG. 4. (a) The range, from maximum to minimum, of permafrost area under RCP8.5 via SFI and diagnosed directly from TSL for the period 1900–2099. The gray bar is indicative of the present-day area estimate. (b) Permafrost area as computed by SFI plotted against TSL area, for each year of 1900–2099 for all RCPs. Both SFI and TSL data have a 10-yr smoothing. Time progresses along each model line; the upper-right start of each line is the year 1900, the dot on the lines is 2000, and the lower-left end of each line is 2099.
can be considered to have land models that produce “colder” soils in frost-prone regions, and the reverse is true for models that sit below 1:1 in Fig. 4b (e.g., BCC-CSM1.1, CanESM2, GISS-E2-R, and MPI-ESM-LR). Note that the 1:1 line is only a reference point; it is neither truth nor observation. Models that start their trace to the left side of the diagram (i.e., have low SFI values) have warmer, less permafrost-prone, climates (e.g., CanESM2). The MIROC models are subject to having both a “colder” land model and a warm climate. The ideal position for a model on Fig. 4b is unknown, as SFI is simply another model, but one that acts as a climate normalizing agent.

In several instances, the same land model is applied in slightly different CMIP5 models—for example, CCSM4.0 and NorESM1-M both use CLM4, the HadGEM2 models both use the Met Office Surface Exchange Scheme (MOSES2.2), and the MIROC models share the same scheme (see Table 2). While the climate of these models can be different (see Fig. 3), the behavior of each pair in Fig. 4b is similar, thus indicating good discrimination of land model influence (e.g., the HadGEM2 models practically overlap each other, as do CCSM4.0 and NorESM1-M). The MIROC models both show an accelerated decline in TSL area once SFI estimates are 2.5 × 10^6 km^2 or less, suggesting that this more frost-prone land model finally succumbs to the warming climate, possibly crossing a threshold of sorts. MIROC5 runs at twice the resolution of MIROC-ESM, which may partly explain their different positions, but their behavior is similar.

According to the classification of Slater et al. (2001), the MOSES2.2 land model (HadGEM2; Essery et al. 2001) has an implicit snow scheme, which effectively negates or reduces the insulating capacity of snow. That is, the top soil layer in this model takes on the same temperature of the air–snow interface. Consequently, the soils tend to be too cold. A complimentary study to this one by Koven et al. (2013) demonstrates this problem quite well. Despite the high-resolution soil column in GFDL’s Land Model (LM3; 20 layers, with 15 layers in the top 3 m; Fig. 2) and the use of a multilayer snow model, it shows a structural deficiency of minimal snow insulation, causing it to have very cold soil temperatures (not shown) and to simulate excessive permafrost. Another example of model structure heavily impacting the embedded permafrost simulation is the CanESM2 model, which uses the three soil layer Canadian Land Surface Scheme (CLASS) land model (Verseghy 2000), the third layer spanning 0.35–4.10-m depth (Fig. 2). Because of its shallow upper bound and low latent heat sink (maximum total frozen soil moisture content is below 150 mm for much of the permafrost region due to the imposition of bedrock), this layer is subject to large annual temperature fluctuations; thus, it tends not to stay frozen throughout the year in more marginal permafrost zone climates. Permafrost area via annual mean soil temperature in CanESM2 gives a mean area of 20.8 × 10^6 km^2 for 1986–2005 (compared to 3.75 × 10^6 km^2 via TSL), but by 2099 under RCP8.5, MST permafrost area is down to 2.9 × 10^6 km^2. The MPI-EMS-LR land model does not include explicit soil freezing processes (Roeckner et al. 2003) and thus has large seasonal soil temperature amplitudes, which explains the large MST (22.7 × 10^6 km^2; 1986–2005) area compared to TSL (10.1 × 10^6 km^2; 1986–2005). Improved process representation can also be detected; for example, BCC-CSM1.1 uses BCC_AVIM1 as its land model, which includes neither organic soil matter nor a deep soil column. The BCC_AVIM1.0, which uses a 10-layer soil column based on the old CLM3 model (Oleson et al. 2004), thus maintains less permafrost than CLM4 (used by CCSM4.0; Lawrence et al. 2008, 2012), resulting in BCC-CSM1.1 having a lower ratio of TSL:SFI permafrost area than CCSM4.0 (Fig. 4b).

2) Permafrost Climate Sensitivity

For each CMIP5 model in Fig. 4, the respective change in air temperature over the present-day permafrost region (Fig. 1) was regressed against the change in permafrost area for instances where permafrost area was greater than 2 × 10^6 km^2. Permafrost area was computed via TSL, SFI, and KUD methods for all RCPs. The linearity of this relationship is shown in Fig. 5; similar linear relations have been found for sea ice (Mahlstein and Knutti 2012). Across all 13 CMIP5 models for each of TSL, SFI, and KUD, the mean coefficient of determination R^2 of the linear regressions is above 0.80, 0.85, 0.91, and 0.94 for RCP2.6, RCP4.5, RCP6.0, and RCP8.5 respectively, suggesting a good linear fit. The fit becomes better with higher RCPs as interannual- to decadal-scale variability has less influence relative to the warming signal, but there is considerable variability in the sensitivity of the models (see Table 3). For example, the TSL sensitivity of INMCM4 is a notable outlier at only 0.75 × 10^6 km^2°C^{-1}. For all three methods (TSL, SFI, and KUD), the mean sensitivity decreases by about 0.15 × 10^6 km^2°C^{-1} from RCP2.6 to RCP8.5, as very cold regions with more resilient permafrost slowly become subject to degradation under RCP8.5. Of interest is that models with widely different TSL estimates of permafrost can have similar sensitivities. For example, in RCP8.5 the sensitivities of GFDL-ESM2M and MPI-ESM-LR are 1.7 and 1.75 × 10^6 km^2°C^{-1} respectively despite GFDL-ESM2M having more permafrost in 2099 than MPI-ESM-LR has in 1900.

It is important to be aware that such regression analysis is purely empirical and the sensitivities do not apply on a regional basis. Indeed, the SFI and KUD cases show...
that different climate trajectories applied to a fixed model can produce a variety of sensitivities, therefore suggesting that spatial and temporal differences in climate change (e.g., seasonal differences, differences in polar amplification, snow change regionality) can play an important role in determining how fast permafrost degrades. The variance of sensitivity across the models is greater under TSL than SFI or KUD, which is perhaps not surprising given the added degrees of freedom that the prognostic land models generate in the TSL method. The higher sensitivity of KUD compared to SFI may partly stem from the fact that it overestimates permafrost area in marginal regions, thus permafrost retreats rapidly in places. The KUD method has been typically applied to colder, continuous permafrost regions.

The metric “percentage change in permafrost area” has been used to make comparisons between modeling studies (Saito et al. 2007; Zhang et al. 2008; Dankers et al. 2011), but results presented here suggest it is not very informative for comparing future states because the sensitivity of permafrost area to warming can be quite similar while the absolute areas and percentage changes are quite different.

5. Future projections based on climate change

Climate biases presented in Fig. 3, along with the spread of permafrost area shown in Fig. 4 and the sensitivity analysis in Fig. 5, suggest that an assessment of future permafrost might be more informative when performed with respect to climatic change rather than by raw diagnosis from CMIP5 models. Issues such as inherent climate biases can be ameliorated through this approach but uncertainty will remain as some models clearly perform better than others across a range of diagnostics.

Figure 6 gives an indication of the mean trend and projected change in temperature over the present-day permafrost region (Fig. 1). For each CMIP5 model and reanalysis, the average temperature over the region is computed and their respective 1986–2005 average is subtracted. Despite the absolute climate biases in individual models (Fig. 3), the rate of change over the last 25 years (0.46°C per decade based on the reanalysis data) is captured very well by the mean of CMIP5 models. The mean temperature change by 2099 across available models is 2.2°C, 3.8°C, 4.5°C, and 7.8°C for RCP2.6, RCP4.5, RCP6.0, and RCP8.5, but there is a large range across the models; for example, HadGEM2-Earth System (ES) warms 16.2°C (+11.3°C) while GISS-E2-R only warms by +1.7°C (+4.6°C) for RCP4.5 (RCP8.5).

<table>
<thead>
<tr>
<th>Model</th>
<th>TSL</th>
<th>SFI</th>
<th>KUD</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCC-CSM1</td>
<td>-1.58</td>
<td>-2.02</td>
<td>-2.13</td>
</tr>
<tr>
<td>CanESM2</td>
<td>-1.18</td>
<td>-1.72</td>
<td>-2.01</td>
</tr>
<tr>
<td>CCSM4.0</td>
<td>-1.34</td>
<td>-1.54</td>
<td>-2.02</td>
</tr>
<tr>
<td>CNRM-CM5</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>CSIRO-Mk3.6.1</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>GFDL-ESM2M</td>
<td>-1.69</td>
<td>-2.05</td>
<td>-2.23</td>
</tr>
<tr>
<td>GISS-E2-H</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>GISS-E2-R</td>
<td>-1.75</td>
<td>-1.88</td>
<td>-1.87</td>
</tr>
<tr>
<td>HadCM3</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>HadGEM2-CC</td>
<td>-1.47</td>
<td>-1.54</td>
<td>-1.64</td>
</tr>
<tr>
<td>HadGEM2-ES</td>
<td>-1.46</td>
<td>-1.52</td>
<td>-1.63</td>
</tr>
<tr>
<td>INM-CM4</td>
<td>-0.75</td>
<td>-1.97</td>
<td>-2.32</td>
</tr>
<tr>
<td>MIROC5</td>
<td>-1.24</td>
<td>-1.27</td>
<td>-1.67</td>
</tr>
<tr>
<td>MIROC-ESM</td>
<td>-1.08</td>
<td>-1.14</td>
<td>-1.46</td>
</tr>
<tr>
<td>MPI-ESM-LR</td>
<td>-1.75</td>
<td>-2.07</td>
<td>-2.09</td>
</tr>
<tr>
<td>MRI-CGCM3</td>
<td>-1.32</td>
<td>-1.97</td>
<td>-1.97</td>
</tr>
<tr>
<td>NorESM1-M</td>
<td>-1.27</td>
<td>-1.66</td>
<td>-1.99</td>
</tr>
</tbody>
</table>

**TABLE 3.** The sensitivity (10^6 km^2°C^-1) of permafrost area to changes in air temperature over the region of present-day observed permafrost, under RCP8.5 forcing.
To project permafrost on the basis of climate change we can apply the SFI or KUD methods, remembering though that they showed an ~25%–38% greater sensitivity than TSL. To create forcing datasets we take an average of the four reanalysis products for both air temperature and snow depth over 1986–2005, then for each CMIP5 model and RCP, at each grid box, the temperature change and a proportional change in snow from this time period is added to the reanalysis mean to form trajectories of future climate. Unfortunately this method does not account for the fact that biases can (and do) change over time. The number of models that provided required temperature and snow data for RCP4.5 and RCP8.5 was 15, while only 10 had data for RCP2.6 and 9 for RCP6.0.

For projecting future permafrost under these trajectories, the SFI area has been tuned to match observed area during the reanalysis forced period by using a minimum frost index value of 0.58, rather than the 0.60 suggested by Nelson and Outcalt (1987). Given the range of warming per scenario and the different sensitivities caused by alternate climate evolution, it is not surprising to see sizable variance for each RCP projection of permafrost area (Fig. 7). Mean values of near-surface permafrost area

![Fig. 6. Mean change in surface air temperature over present-day permafrost area. The number of models contributing to each RCP estimate is given in brackets. Shaded area is one standard deviation across models.](image)

![Fig. 7. Projected change in sustainable permafrost area based on climate change from the present day via the SFI. Shaded areas represent one standard deviation across the CMIP5 models and the dashed black line is the model equivalent present-day total area of continuous and discontinuous permafrost.](image)
(recall that the SFI method tells us the permafrost suitability area under climate equilibrium, not the actual permafrost area at 2099) by 2099 are 10.0, 7.5, 5.9, and $2.1 \times 10^6 \text{km}^2$ for RCP2.6, RCP4.5, RCP6.0, and RCP8.5 respectively. While the areas suitable for permafrost in RCP2.6 and RCP4.5 appear to have stabilized by 2099, RCP6.0 and RCP8.5 are still on a declining trajectory.

Areas of sustainable near-surface permafrost by the year 2099 for all RCPs are given in Fig. 8, which shows the number of climate trajectories (one trajectory per CGCM) that retain permafrost for a given location.

Under RCP2.6 the majority of present-day continuous permafrost is likely to remain with some transforming to discontinuous. The increased warming of RCP4.5 and RCP6.0 see permafrost retreat farther north with Alaska showing vulnerability to degradation. RCP8.5 sees a wide span in expectations with a couple of cooler models keeping permafrost over large portions of Eurasia and Canada while others remove virtually all permafrost (which is not surprising given 11°C warming). However, the Canadian Archipelago, eastern Siberian uplands, high Arctic coastal Russia (e.g., the Taymir
6. Summary and conclusions

We have applied several methods of diagnosing permafrost over a range of climate models and future scenarios of climate change. From this analysis a number of salient points have become apparent and these are summarized here.

Permafrost relevant climate bias in temperatures or snow accumulation over the present-day permafrost region can be significant when compared to reanalyses for the 1986–2005 period. Impacts of biases are particularly evident from the large range in permafrost area diagnosed using only the climate indices of each model (e.g., the SFI method). Models display a greater range of temperature biases during the winter period, and shallow Arctic snow depths mean that small differences (e.g., 10 cm) seen in average snow depth will impact insulating abilities.

The range of permafrost area directly diagnosed from modeled soil temperatures (TSL) is extremely large (~4–25 $\times 10^6$ km$^2$ averaged over 1986–2005) and the range in mean soil temperature (MST; Table 2) over the present-day permafrost region for 1986–2005 is over 7°C across the models. Hence, it can be concluded that in general permafrost is not well represented by the ensemble of CMIP5 models, though not all models perform poorly. These large ranges are due to two factors: 1) differences in simulated surface climate, and 2) varying abilities of the underlying land surface models, with the latter likely having a larger impact (Fig. 4a; section 4b). When TSL is compared to climate-based permafrost indices such as the surface frost index (SFI) we see that some land models are particularly prone to producing cold soils with large permafrost areas while others encourage thawing in the upper levels of the soil. The comparison of TSL to SFI exposes weaknesses in many model’s representations of permafrost processes. The structure of the land model, for example, the insulating ability of its snow scheme or the discretization of the soil column, can play a significant role in determining permafrost extent. Similar behavior is seen across all RCP scenarios for each individual CMIP5 model when comparing change in permafrost area via TSL and SFI, indicating that land model responses are consistent regardless of climate trajectory.

Despite large differences in the absolute permafrost area, the relationship between the decrease in permafrost area and the warming air temperature over the present-day permafrost region can be similar between many models and is approximately linear for all models. Spatial and temporal differences in the pattern of climate change can impact this permafrost–climate sensitivity as evidenced by the spectrum of sensitivity under either the SFI or KUD models. Averaged across all methods and models under RCP8.5, permafrost area sensitivity is $1.67 \times 10^6$ km$^2$ C$^{-1}$ but the sensitivity spans a wide range (0.75–2.32 $\times 10^6$ km$^2$ C$^{-1}$).

Using projected change in climate relative to present-day climate (i.e., bias corrected climate trajectories rather than absolute climate trajectories direct from the CGCMs), we have estimated potential trajectories for sustainable near-surface permafrost out to 2099. When measured over the present-day permafrost region, the range of warming for a given century-scale RCP scenario across all CMIP5 models can be as large as the average warming, resulting in quite different permafrost regimes by the end of the twenty-first century. Nonetheless, robust results of future permafrost sustainability are evident, particularly in the core CMIP5 experiments. Conservatively, under RCP4.5 near-surface permafrost is unlikely to be sustainable in the present-day discontinuous zone and under RCP8.5 existence will be reduced to the Canadian Archipelago, Russian Arctic coast, and east Siberian uplands. Projections based on SFI or KUD provide useful insights to future permafrost distribution, but it is preferable for Earth system models to include the full physics of subsurface processes in order to allow for interactions with hydrology, vegetation, and biogeochemical cycling and thus correctly account for and quantify feedbacks related to permafrost degradation on the local and global climate system.

Acknowledgments. Support from NSF via ARC-1041081, ARC-0902057, and ARC-0901962 is gratefully acknowledged. This research was also supported by funding from the U.S. Department of Energy, Office of Biological and Environmental Research as part of its Climate Change Prediction Program, and Cooperative Agreement DE-FC03-97ER62402/A010. Thanks to Charlie Koven for useful discussions, the CMIP5 model groups for their data, model developers for information about model structures, and three anonymous reviewers for aiding clarity.

REFERENCES

online at http://www.cnrm.meteo.fr/surfex/IMG/pdf/surfex_scientific_documentation.pdf.]


Yoshikawa, K., and L. Hinzman, 2003: Shrinking thermokarst ponds and groundwater dynamics in discontinuous permafrost


