WETCHIMP-WSL: intercomparison of wetland methane emissions models over West Siberia


¹School of Earth and Space Exploration, Arizona State University, Tempe, AZ, USA
²Canadian Centre for Climate Modelling and Analysis, Environment Canada, Victoria, BC, Canada
³National Institute for Environmental Studies, Tsukuba, Japan
⁴Max Planck Institute for Meteorology, Hamburg, Germany
⁵Climate and Environmental Physics, Physics Institute, University of Bern, Bern, Switzerland
⁶Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland
⁷Department of Life Sciences, Imperial College, Silwood Park Campus, Ascot, UK
⁸International Center for Climate and Global Change Research and School of Forestry and Wildlife Sciences, Auburn University, Auburn, AL, USA
⁹Department of Earth, Atmospheric, and Planetary Sciences, Purdue University, West Lafayette, IN, USA
¹⁰Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, CO, USA
¹¹Earth Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA
¹²City College of New York, City University of New York, New York, NY, USA
¹³Institute of Botany, University of Hohenheim, Stuttgart, Germany
¹⁴Moscow State University, Moscow, Russia
¹⁵Institute of Forest Science, Russian Academy of Sciences, Uspenskoe, Russia
¹⁶Laboratory of Computational Geophysics, Tomsk State University, Tomsk, Russia
¹⁷Yugra State University, Khanty-Mansyisk, Russia
¹⁸Oak Ridge National Laboratory, Oak Ridge, TN, USA
¹⁹A. M. Obukhov Institute of Atmospheric Physics, Russian Academy of Sciences, Moscow, Russia
²⁰Kazan Federal University, Kazan, Russia
²¹Department of Geography, University of Exeter, Exeter, UK
²²Department of Geosciences, University of Massachusetts, Amherst, MA, USA
²³Institute of Earth Surface Dynamics, University of Lausanne, Lausanne, Switzerland

Correspondence to: T. J. Bohn (theodore.bohn@asu.edu)

Received: 10 December 2014 – Published in Biogeosciences Discuss.: 30 January 2015
Revised: 29 April 2015 – Accepted: 30 April 2015 – Published: 3 June 2015

Abstract. Wetlands are the world’s largest natural source of methane, a powerful greenhouse gas. The strong sensitivity of methane emissions to environmental factors such as soil temperature and moisture has led to concerns about potential positive feedbacks to climate change. This risk is particularly relevant at high latitudes, which have experienced pronounced warming and where thawing permafrost could potentially liberate large amounts of labile carbon over the next 100 years. However, global models disagree as to the magnitude and spatial distribution of emissions, due to uncertainties in wetland area and emissions per unit area and a scarcity of in situ observations. Recent intensive field campaigns across the West Siberian Lowland (WSL) make this an ideal region over which to assess the per-
formance of large-scale process-based wetland models in a high-latitude environment. Here we present the results of a follow-up to the Wetland and Wetland CH₄ Intercomparison of Models Project (WETCHIMP), focused on the West Siberian Lowland (WETCHIMP-WSL). We assessed 21 models and 5 inversions over this domain in terms of total CH₄ emissions, simulated wetland areas, and CH₄ fluxes per unit wetland area and compared these results to an intensive in situ CH₄ flux data set, several wetland maps, and two satellite surface water products. We found that (a) despite the large scatter of individual estimates, 12-year mean estimates of annual total emissions over the WSL from forward models (5.34 ± 0.54 Tg CH₄ yr⁻¹), inversions (6.06 ± 1.22 Tg CH₄ yr⁻¹), and in situ observations (3.91 ± 1.29 Tg CH₄ yr⁻¹) largely agreed; (b) forward models using surface water products alone to estimate wetland areas suffered from severe biases in CH₄ emissions; (c) the interannual time series of models that lacked either soil thermal physics appropriate to the high latitudes or realistic emissions from unsaturated peatlands tended to be dominated by a single environmental driver (inundation or air temperature), unlike those of inversions and more sophisticated forward models; (d) differences in biogeochemical schemes across models had relatively smaller influence over performance; and (e) multiyear or multidecade observational records are crucial for evaluating models’ responses to long-term climate change.

1 Introduction

Methane (CH₄) emissions from high-latitude wetlands are an important component of the global climate system. CH₄ is an important greenhouse gas, with approximately 34 times the global warming potential of carbon dioxide (CO₂) over a century time horizon (IPCC, 2013). Globally, wetlands are the largest natural source of CH₄ emissions to the atmosphere (IPCC, 2013). Because wetland CH₄ emissions are highly sensitive to soil temperature and moisture conditions (Saarnio et al., 1997; Friberg et al., 2003; Christensen et al., 2003; Moore et al., 2011; Glagolev et al., 2011; Sabrekov et al., 2014), there is concern that they will provide positive feedback to future climate warming (Gedney et al., 2004; Eliseev et al., 2008; Ringeval et al., 2011). This risk is particularly important in the world’s high latitudes because they contain nearly half of the world’s wetlands (Lehner and Döll, 2004) and because the high latitudes have been and are forecast to continue experiencing more rapid warming than elsewhere (Serreze et al., 2000; IPCC, 2013). Adding to these concerns is the potential liberation (and possible conversion to CH₄) of previously frozen, labile soil carbon from thawing permafrost over the next century (Christensen et al., 2004; Schuur et al., 2008; Koven et al., 2011; Schaefer et al., 2011).

Process-based models are crucial for increasing our understanding of the response of wetland CH₄ emissions to climate change. Large-scale biogeochemical models, especially those embedded within earth system models, are particularly important for estimating the magnitudes of feedbacks to climate change (e.g., Gedney et al., 2004; Eliseev et al., 2008; Koven et al., 2011). However, as shown in the global Wetland and Wetland CH₄ Intercomparison of Models Project (WETCHIMP; Melton et al., 2013; Wania et al., 2013), there was wide disagreement among large-scale models as to the magnitude of global and regional wetland CH₄ emissions, in terms of both wetland areas and CH₄ emissions per unit wetland area. These discrepancies were due in part to the large variety of schemes used for representing hydrological and biogeochemical processes, in part to uncertainties in model parameterizations, and in part to the sparseness of in situ observations with which to evaluate model performance (Melton et al., 2013).

In addition to these challenges on the global scale, the unique characteristics of high-latitude environments pose further problems for biogeochemical models. For example, much of the northern land surface is underlain by permafrost, which impedes drainage (Smith et al., 2005) and stores ancient carbon (Koven et al., 2011) via temperature-dependent constraints on carbon cycling (Schuur et al., 2008). Similarly, peat soils and winter snowpack can thermally insulate soils (Zhang, 2005; Lawrence and Slater, 2008, 2010), dampening their sensitivities to interannual variability in climate. Several commonly used global biogeochemical models (e.g., Tian et al., 2010; Hopcroft et al., 2011; Hodson et al., 2011; Kleinen et al., 2012) lack representations of some or all of these processes.

The prevalence of peatlands in the high-latitudes poses further challenges to modeling (Frolking et al., 2009). Peatlands are a type of wetland containing deep deposits of highly porous, organic-rich soil, formed over thousands of years under waterlogged and anoxic conditions, which inhibit decomposition (Gorham, 1991; Frolking et al., 2011). Within the porous soil, the water table is often only a few centimeters below the surface, leading to anoxic conditions and CH₄ emissions even when no surface water is present (Saarnio et al., 1997; Friberg et al., 2003; Glagolev et al., 2011). This condition can lead to an underestimation of wetland area when using satellite surface water products as inputs to wetland methane emissions models. In addition, trees and shrubs are found with varying frequency in peatlands (e.g., Shimoyama et al., 2003; Efremova et al., 2014), interfering with the detection of inundation. Furthermore, the water table depth within a peatland is typically heterogeneous, varying on the scale of tens of centimeters as a function of microtopography (hummocks, hollows, ridges, and pools; Eppinga et al., 2008). Models vary widely in their representations of wetland soil moisture conditions, ranging from schemes that do not explicitly consider the water table position (e.g., Hodson et al., 2011) to a single uniform water table depth for
each grid cell (e.g., Zhuang et al., 2004) to more sophisticated schemes that allow for sub-grid heterogeneity in the water table (e.g., Bohn et al., 2007, 2013; Ringeval et al., 2010; Riley et al., 2011; Kleinen et al., 2012; Stocker et al., 2014; Subin et al., 2014). Finally, peatland soils can be highly acidic and nutrient-poor, and much of the available carbon substrate can be recalcitrant (Clymo et al., 1984; Dorrepaal et al., 2009). While some models attempt to account for the effects of soil chemical conditions such as pH, redox potential, and nutrient limitation (e.g., Zhuang et al., 2004; Riley et al., 2011; Sabrekov et al., 2013; Spahni et al., 2013), not all do.

Given the potential problems of parameter uncertainty and equifinality (Tang and Zhuang, 2008; van Huyssteden et al., 2009) and computational limitations when wetland components are embedded within global climate models, it is important to determine which model features are necessary to simulate high-latitude peatlands accurately and to constrain parameter values with observations. Until recently, the evaluation of large-scale wetland CH$_4$ emissions models has been difficult, due to the sparseness of in situ and atmospheric CH$_4$ observations. However, observations from the West Siberian Lowland (WSL) now offer the opportunity to assess model performance, thanks to recent intensive field campaigns (Glagolev et al., 2011), aircraft profiles (Umezawa et al., 2012), tall-tower observations (Sasakawa et al., 2010; Winderlich et al., 2010), and high-resolution wetland inventories (Sheng et al., 2004; Peregon et al., 2008, 2009).

Our primary goal in this study is to determine how well current global large-scale models capture the dynamics of high-latitude wetland CH$_4$ emissions. To this end, we assess the performance of 21 large-scale wetland CH$_4$ emissions models over West Siberia, relative to in situ and remotely sensed observations as well as inverse models. We examine both spatial and temporal accuracy, including seasonal and interannual variability, and estimate the relative influences of environmental drivers on model behaviors. We identify the dominant sources of error and the model features that may have caused them. Finally, we make recommendations as to which model features are necessary for accurate simulations of high-latitude wetland CH$_4$ emissions and which types of observations would help improve future efforts to constrain model behaviors.

2 Methods

2.1 Spatial domain

The West Siberian Lowland (WSL) occupies approximately 2.5 million km$^2$ in northern central Eurasia, spanning from 50 to 75° N and 60 to 95° E (Fig. 1a). This region is bounded on the west by the Ural Mountains; on the east by the Yenisei River and the Central Siberian Plateau; on the north by the

Figure 1. Map of the West Siberian Lowland (WSL). (a) Limits of domain (brown) and peatland distribution (cyan), taken from Sheng et al. (2004); lakes of area > 1 km$^2$ (blue) taken from Lehner and Döll (2004); permafrost zone boundaries after Kremenetski et al. (2003); CH$_4$ sampling sites from Glagolev et al. (2011), denoted by red circles. (b) Dominant land cover at 25 km derived from MODIS-MOD12Q1 500 m land cover classification (Friedl et al., 2010).
Arctic Ocean; and on the south by the Altai Mountains and the grasslands of the Eurasian Steppe (Sheng et al., 2004). The WSL contains most of the drainage areas of the Ob’ and Irtysh rivers, as well as the western tributaries of the Yenisei River, all of which drain into the Arctic Ocean. Permafrost in various forms (continuous, discontinuous, isolated, and sporadic) covers more than half of the area of the WSL, from the Arctic Ocean south to approximately 60° N, with continuous permafrost occurring north of 67° N (Kremenetski et al., 2003). The region’s major biomes (Fig. 1b) consist of the treeless tundra north of 66° N, approximately coincident with continuous permafrost; the taiga forest belt between 55 and 66° N; and the grasslands of the steppe south of 55° N.

Wetlands occupy 600,000 km², or about 25% of the land area of the WSL, primarily in the taiga and tundra zones (Sheng et al., 2004). The vast majority of these wetlands are peatlands, which have peat depths ranging from 50 cm to over 5 m and which comprise a total soil carbon pool of 70 Pg C (Sheng et al., 2004). Numerous field studies have documented strong methane emissions from these peatlands, particularly those south of the southern limit of permafrost (e.g., Sabrekov et al., 2014; Sasakawa et al., 2012; Glagolev et al., 2011, 2012; Friiborg et al., 2003; Shimoyama et al., 2003; Panikov and Dedysh, 2000). Permanent water bodies, ranging in size from lakes 100 km² in area to pools only a few meters across, are comingled with wetlands throughout the domain (Lehner and Döll, 2004; Repo et al., 2007; Eppinga et al., 2008). Notable concentrations of lakes are found (a) north of the Ob’ River between 61 and 64° N and 68 and 80° E; (b) west of the confluence of the Ob’ and Irtysh rivers between 59 and 61° N and 64 and 70° E; and (c) on the Yamal Peninsula north of 68° N.

Because the vegetative and soil conditions vary substantially across the domain, we have divided it into two halves of approximately equal size along 61° N latitude. The region north of this line contains permafrost, while the region south of the line is essentially permafrost-free.

2.2 Terminology

Estimating wetland CH₄ emissions over large scales requires accurately delineating the wetland area over which CH₄ emissions can occur. Unfortunately, “wetland” definitions vary within the scientific community (Mitsch and Gosselink, 2000). For the purposes of estimating CH₄ emissions, the key characteristics include anoxia and available labile carbon substrate; therefore, we will adopt the definition proposed by Canada’s National Wetlands Working Group (Tarnocai et al., 1988): land that is saturated with water for long enough to promote wetland or aquatic processes as indicated by poorly drained soils, hydrophytic vegetation, and various kinds of biological activity which are adapted to a wet environment.

Because permanent, deep (> 2 m) open-water bodies are subject to additional processes (e.g., allochthonous carbon inputs, wind-driven mixing of the water column; Pace et al., 2004), we will exclude them from our definition. Unfortunately, explicit observations of lake depths are lacking for all but the deepest lakes; therefore, we will instead use an area threshold (1 km²) to identify permanent lakes. This definition of wetlands therefore includes all peatlands (inundated or not), seasonally inundated non-peatland soils (e.g., river floodplains), and small ponds or lakes but excludes rivers and large lakes.

We define “surface water” as all freshwater above the soil surface, i.e., the superset of inundation, lakes, and rivers. We define “inundation” as temporary (present for less than 1 year) standing water above the soil surface; “lakes” as permanent water bodies (present for more than 1 year) exceeding 1 km² in area; and “rivers” as channels that carry turbulent water. Surface water therefore includes areas that do not emit large amounts of CH₄, such as rivers, and also excludes some CH₄-emitting areas such as non-inundated peatlands.

For models, we will use the term “CH₄-producing area” to refer to the area over which CH₄ production is simulated, which might not coincide exactly with the areas of actual or simulated wetlands.

2.3 Observations and inversions

Table 1 lists the various observations and inversions that we used in this study. We considered four wetland map products over the WSL, all of which have been used in high-latitude wetland carbon studies. Two of them are regional maps specific to the WSL: Sheng et al. (2004), denoted by “Sheng2004”, and Peregon et al. (2008), denoted by “Peregon2008”. Both Sheng 2004 and Peregon2008 used the 1 : 2500 000-scale map of Romanova (1977); Peregon2008 was entirely based on the Romanova map, while Sheng2004 used the Romanova map north of 65° N and used the 1 : 100 000-scale maps of Markov (1971) and Matukhin and Danilov (2000) elsewhere. Both of these maps delineate the extents of peatlands, including ponds and lakes smaller than 1 km² in area. The Sheng2004 product additionally includes a separate layer delineating lakes larger than 1 km². The Peregon2008 product distinguishes between various wetland subtypes (e.g., sphagnum- or sedge-dominated bogs and high palsa mires). The third map is the Northern Circumpolar Soil Carbon Database (NCSCD; Tarnocai et al., 2009), an inventory of carbon-rich soils, including peatlands, within the Arctic permafrost region. Models that have used this database have taken the Histel and Histosol delineations to be synonymous with peatlands. The fourth map is the wetland layer (GLWD-3, excluding the rivers and lakes of area > 1 km² of layers GLWD-1 and GLWD-2) of the Global Lakes and Wetland Database (GLWD; Lehner and Döll, 2004), in which wetland extents are the union of polygons from four different global databases.

Two global time-varying surface water products derived from remote-sensing observations were also examined in this study: the Global Inundation Extent from Multi-Satellites (GIEMS; Prigent et al., 2007; Papa et al., 2010), derived
Table 1. Observations and inversions used in this study.

<table>
<thead>
<tr>
<th>Name</th>
<th>Reference</th>
<th>Description</th>
<th>Temporal domain</th>
<th>Temporal resolution</th>
<th>Spatial domain</th>
<th>Spatial resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wetland maps</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peregon2008</td>
<td>Peregon et al. (2008)</td>
<td>Wetland map of WSL based on digitization of regional map of Romanova et al. (1977). Wetland types identified by remote sensing and field validation.</td>
<td>Second half of 20th century</td>
<td>Static map</td>
<td>Western Siberia</td>
<td>1 : 2,500,000</td>
</tr>
<tr>
<td>Northern Circumpolar Soil Carbon Database</td>
<td>Tarnocai et al. (2009)</td>
<td>Map of wetlands across the northern circumpolar permafrost region. Over the WSL, based on maps of Fridland (1988) and Naumov (1993).</td>
<td>Second half of 20th century</td>
<td>Static map</td>
<td>Northern circumpolar permafrost region</td>
<td>1 : 2,500,000</td>
</tr>
<tr>
<td>Global Lake and Wetland Database (GLWD)</td>
<td>Lehner and Döll (2004)</td>
<td>Global lake and wetland map. Wetlands were the union of four global data sets.</td>
<td>Second half of 20th century</td>
<td>Static map</td>
<td>Global</td>
<td>1 : 1,000,000</td>
</tr>
<tr>
<td>Surface Water</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Global Inundation Extent from Multi-Satellites (GIEMS)</td>
<td>Papa et al. (2010)</td>
<td>Remote-sensing inundation product based on visible (AVHRR) and active (SSM/I) and passive (ERS) microwave sensors.</td>
<td>1993–2004</td>
<td>Daily, aggregated monthly</td>
<td>Global</td>
<td>25 km equal-area grid, aggregated to 0.5° x 0.5°</td>
</tr>
<tr>
<td>Surface Water Microwave Product Series (SWAMPS)</td>
<td>Schroeder et al. (2010)</td>
<td>Remote-sensing inundation product based on active (SeaWinds-on-QuikSCAT, ERS, and ASCAT) and passive (SSM/I, SSMIS) microwave sensors.</td>
<td>1992–2013</td>
<td>Daily, aggregated monthly</td>
<td>Global</td>
<td>25 km equal-area grid, aggregated to 0.5° x 0.5°</td>
</tr>
<tr>
<td>CH₄ Inventory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glagolev2011</td>
<td>Glagolev et al. (2011)</td>
<td>In situ flux sampling along transect spanning West Siberia, 2006–2010; statistical model of fluxes as function of wetland type applied to map of Peregon et al. (2008).</td>
<td>2006–2010</td>
<td>Monthly climatology</td>
<td>Western Siberia</td>
<td>0.5° x 0.5°</td>
</tr>
<tr>
<td>CH₄ Inversions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bloom2010</td>
<td>Bloom et al. (2010)</td>
<td>Global optimization of relationship between SCIAMACHY atmospheric CH₄ concentrations (Bovensmann et al., 1999), NCEP/NCAR surface temperatures (Kalnay et al., 1996), and GRACE gravity anomalies (Tapley et al., 2004).</td>
<td>2003–2007</td>
<td>Annual</td>
<td>Global</td>
<td>3° x 3°</td>
</tr>
<tr>
<td>Winderlich2012</td>
<td>Winderlich (2012), Schuldt et al. (2013)</td>
<td>Regional inversion over West Siberia, with Kaplan (2002) as the wetland prior.</td>
<td>2009</td>
<td>Monthly climatology</td>
<td>Regional</td>
<td>1° x 1° resolution for both prior and coefficients over WSL</td>
</tr>
</tbody>
</table>
from visible and near-infrared (AVHRR) and active (SSM/I) and passive (ERS) microwave sensors over the period 1993–2004, and the Surface Water Microwave Product Series (SWAMPS; Schroeder et al., 2010), derived from active (SeaWinds-on-QuikSCAT, ERS, and ASCAT) and passive (SSM/I, SSM/I/ S. AMSR-E) microwave sensors over the period 1992–2013. For both products, surface water area fractions \( F_w \) were aggregated from their native 25 km equal-area grids to a 0.5° × 0.5° geographic grid and from daily to monthly temporal resolution, for consistency with model results.

For \( CH_4 \) emissions, our primary reference for in situ observations was the estimate of Glagolev et al. (2011), which we will refer to as “Glagolev2011”. The Glagolev2011 product consists of both a database of over 2000 individual chamber observations from representative landforms at each of 36 major sites over the period 2006–2010 (Fig. 1a) and a map of long-term average emissions created by applying the mean observed emissions to the wetlands of the PeregIno2008 map as a function of wetland type. It is worth noting that the Glagolev2011 product is currently undergoing a revision based on higher-resolution maps, which will lead to a substantial increase in annual emissions from the taiga zone, due to a larger spatial extent of high-emitting wetland types (Glagolev et al., 2013). Possible changes to emissions in the tundra zone (in the northern half of the WSL) are not yet known. We consider this product’s large uncertainty in our evaluation of model predictions.

We also considered emissions estimates from five inversions. Two of them were regional: “Kim2011” (Kim et al., 2011) and “Winderlich2012” (Winderlich, 2012; Schuldt et al., 2013). Kim et al. (2011) used an earlier version of Glagolev2011 (Glagolev et al., 2010) at a 1° × 1° resolution as their prior distribution for wetland emissions within the atmospheric transport model NIES-TM (Maksyutov et al., 2008) over the period 2002–2007. Kim et al. (2011) derived 12 climatological average monthly (spatially uniform) coefficients for wetland emissions to optimize atmospheric \( CH_4 \) concentrations over the WSL relative to observed \( CH_4 \) concentrations obtained by aircraft sampling at two locations in the WSL. Winderlich (2012) used the Kaplan (2002) wetland inventory for prior wetland emissions, within the global inversion system TM3-STILT (Rödenbeck et al., 2009; Trusilova et al., 2010) for the year 2009. Winderlich (2012) derived 12 monthly coefficients for wetland emissions, uniquely for each point in a 1° × 1° grid, to optimize atmospheric \( CH_4 \) concentrations over the WSL relative to the concentrations measured at the Zotino Tall Tower Observatory and three other \( CH_4 \) tower observation sites (Demyan-skoe, Igrim, and Karasevoe) located between 58 and 63° N.

The other inversions we considered were global: the “Reference” and “Kaplan” versions of the Bousquet et al. (2011) inversion, denoted by “Bousquet2011R” and “Bousquet2011K”, respectively, and the estimate of Bloom et al. (2010), denoted by “Bloom2010”. Bousquet et al. (2011) used the Laboratoire de Météorologie Dynamique general circulation model (LMDZ; Hauglustaine et al., 2004) atmospheric transport model on a 3.75° × 2.5° grid to estimate monthly \( CH_4 \) emissions at a 1° × 1° resolution for the period 1993–2009, optimizing atmospheric concentrations of several gases, including \( CH_4 \), relative to global surface observation networks, for both inversions. The Matthews and Fung (1987) emissions inventory was the prior for wetland emissions in the Bousquet2011R inversion, while the Kaplan (2002) emissions were the prior for the Bousquet2011K inversion. In both cases, a single, spatially uniform set of monthly coefficients was derived for each of 11 large regions of the globe. The region containing the WSL was boreal Asia (in which the WSL makes up the majority of the wetlands). Consequently, spatial patterns in estimated emissions at the scale of 1° × 1° were identical to those of the prior emissions; only the regional total emissions were constrained by the inversions. The 17-year record length of the Bousquet2011 inversions made them appealing candidates for investigating the sensitivities of emissions to interannual variability in environmental drivers. Bloom et al. (2010) did not use an atmospheric transport model, but rather optimized the parameters in a simple model relating observed atmospheric \( CH_4 \) concentrations from the Scanning Imaging Absorption Spectrometer for Atmospheric Chemistry (SCIAMACHY; Bovensmann et al., 1999) on the Envisat satellite to observed surface temperatures from the National Center for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR) weather analyses (Kalnay et al., 1996) and gravity anomalies from the Gravity Recovery and Climate Experiment satellite (GRACE; Tapley et al., 2004), under the assumption that gravity anomalies are indicative of large-scale surface and near-surface water anomalies. The Bloom2010 inversion covered the period 2003–2007, at a 3° × 3° resolution.

2.4 Models

Among the participating models (Table 2) were those of the WETCHIMP study (Melton et al., 2013; Wania et al., 2013) that contributed \( CH_4 \) emissions estimates: CLM4Me (Riley et al., 2011), DLEM (Tian et al., 2010, 2011a, b, 2012), IAP-RAS (Mokhov et al., 2007; Eliseev et al., 2008), LPJ-Bern (Spahn et al., 2011, Zürcher et al., 2013), LPJ-WHYMe (Wania et al., 2009a, b, 2010), LPJ-WSL (Hodson et al., 2011), ORCHIDEE (Ringeval et al., 2010), SDGVM (Hocroft et al., 2011), and UW-VIC (denoted by “UW-VIC (GIEMS)”; Bohn et al., 2013). In addition, we analyzed several other models. “UW-VIC (SWAMPS)” is another instance of UW-VIC with surface water calibrated to match the SWAMPS product. VISIT (Ito and Inatomi, 2012) contributed four configurations using different combinations of wetland maps and methane models: “VISIT (GLWD)” and “VISIT (Sheng)” used the Cao (1996) methane model with the GLWD and Sheng2004 wetland maps, respectively,
Table 2. Participating models and their relevant hydrologic features.

<table>
<thead>
<tr>
<th>Model</th>
<th>Full name</th>
<th>Reference</th>
<th>Configuration(^1)</th>
<th>Period</th>
<th>Observational constraints on CH(_4)-producing areas</th>
<th>Unnatural emissions(^6)</th>
<th>Water table(^8)</th>
<th>Organic soil(^7)</th>
<th>Soil freeze–thaw(^8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLM-Me</td>
<td>Community Land Model v. 4 Methane</td>
<td>Riley et al. (2011)</td>
<td>CLM-Me</td>
<td>1993–2004</td>
<td>GIEMS</td>
<td>S(^a)</td>
<td>Yes</td>
<td>Uniform</td>
<td>Yes</td>
</tr>
<tr>
<td>DLEM2</td>
<td>Dynamic Land Ecosystem Model v. 2</td>
<td>Tian et al. (2010, 2011a, b, 2012)</td>
<td>DLEM2</td>
<td>1993–2004</td>
<td>GIEMS</td>
<td>S</td>
<td>Yes</td>
<td>Uniform</td>
<td>Yes</td>
</tr>
<tr>
<td>ORCHIDEE</td>
<td>Organizing Carbon and Hydrology in Dynamic Ecosystems</td>
<td>Ringeval et al. (2010)</td>
<td>ORCHIDEE</td>
<td>1993–2004</td>
<td>HIROS(^e)</td>
<td>HIROS(^c)</td>
<td>T</td>
<td>Yes</td>
<td>TOPMODEL</td>
</tr>
<tr>
<td>SDGVM</td>
<td>Sheffield Dynamic Global Vegetation Model</td>
<td>Hopcroft et al. (2011)</td>
<td>SDGVM</td>
<td>1993–2004</td>
<td>--</td>
<td>--</td>
<td>ETOPO2(^d)</td>
<td>T</td>
<td>Yes</td>
</tr>
<tr>
<td>UW-VIC</td>
<td>Variable Infiltration Capacity – University of Washington</td>
<td>Bohn et al. (2013)</td>
<td>UW-VIC (GEIWS)</td>
<td>1993–2004</td>
<td>GEIWS</td>
<td>SRTh(^f), ASTER(^f)</td>
<td>Sheng2004</td>
<td>M,M</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VISIT (GLWD)</td>
<td>VISIT (GLWD)</td>
<td>1993–2010</td>
<td>--</td>
<td>--</td>
<td>Sheng2004</td>
<td>M,M</td>
<td>Yes</td>
</tr>
</tbody>
</table>

---

1 Configuration short name identifying both the model and the parameter or feature settings for a particular simulation; for models that contributed only a single simulation, the configuration equals the model name.
2 Surface water: name of time-varying surface water product (if any) used as a constraint on CH\(_4\)-producing area.
3 Topography: name of static wetland map product (if any) used as a constraint on CH\(_4\)-producing area.
4 Map: Name of static wetland map product (if any) used as a constraint on CH\(_4\)-producing area.
5 Code: single-letter code summarizing the types of CH\(_4\)-producing area constraints used (S: surface water only; T: topography with or without surface water constraint; M: static wetland map with or without surface water or topography constraints; M+: subset of M that excludes the NCSCD).
6 Water table: approach used to account for water table depth variations within the grid cell; TOPMODEL: water table depth varies within the grid cell as inferred from topography; GLWD: constant water table depth across the grid cell; n/a: no water table.
7 Soil freeze–thaw: yes or no indicates whether the model accounts for the freezing and thawing of water within the soil column.
8 Source of wetland inundation data (GIS data source: CLM-Me, ORCHIDEE, or SDGVM).

---

\(^{a}\) CLM-Me and ORCHIDEE are listed as S due to tuning or rescaling of inundated areas to match GIEMS, thus destroying contribution of topography.
\(^{b}\) http://cdiac.esd.ornl.gov/ndps/ndp017.html
\(^{c}\) Hydro1K (2013), ETOPO (2006), Sheng2004
\(^{d}\) Amante and Eakins (2009)
\(^{e}\) ETOPO (2006)
\(^{f}\) NASA (2001)
and “VISIT (GLWD-WH)” and “VISIT (Sheng-WH)” replaced the Cao model with the Walter and Heimann (2000) model. LPX-BERN (Shapouri et al., 2013; Stocker et al., 2013, 2014) is a newer version of LPJ-Bern that also contributed four configurations: “LPX-BERN”, which prescribed peatland extent using Pergor2008 and inundation extent using GIEMS; “LPX-BERN (DyPTOP)”, which dynamically predicted the extents of peatlands and inundation; and “LPX-BERN (N)” and “LPX-BERN (DyPTOP-N)”, which additionally simulated interactions between the carbon and nitrogen cycles. DLEM2 is a newer version of DLEM that includes soil thermal physics and lateral matter fluxes (Liu et al., 2013; Pan et al., 2014). LPJ-MPI (Kleinen et al., 2012) is a version of the LPJ model that contains a dynamic peatland model with methane transport by the model of Walter and Heimann (2000). Finally, VIC-TEM-TOPMODEL (Zhu et al., 2014) is a hybrid of UW-VIC (Liang et al., 1994), TEM (Zhuang et al., 2004), and TOPMODEL (Beven and Kirkby, 1979).

The hydrologic and biogeochemical features of these models are listed in Tables 2 and 3, respectively. The models used a variety of approaches to define CH₄-producing areas. To have some consistency across models, the original WETCHIMP study asked participating modelers to use the GIEMS product if their model required wetland extent to be prescribed. Accordingly, some models (DLEM, DLEM2, and LPJ-WSL) used the GIEMS surface water product exclusively to prescribe (time-varying) CH₄-producing areas; these are denoted by the code “S” in Table 2.

Several models (CLM4Me, LPJ-MPI, LPX-BERN (DyPTOP), LPX-BERN (DyPTOP-N), ORCHIDEE, SDGVM, and VIC-TEM-TOPMODEL) predicted surface water and CH₄-producing areas dynamically using topographic information and the TOPMODEL (Beven and Kirkby, 1979) distributed water table approach (in which the area over which the water table is at or above the soil surface can be interpreted to correspond to surface water extent); these models are denoted by a “T” in Table 2. For these models, the CH₄-producing area is the area in which labile soil carbon is sufficiently warm and anoxic for methanogenesis to occur, including both surface water and any non-inundated land with sufficiently shallow water table depths. LPJ-MPI and LPX-BERN (DyPTOP and DyPTOP-N) prognostically determined peatland area as a function of long-term soil moisture conditions; their CH₄-producing areas thus included peatlands (inundated or not) as well as completely saturated or inundated mineral soils. Because the other T models’ CH₄-producing areas had no explicit limits, those teams reported approximations of the models’ true CH₄-producing areas: CLM4Me, ORCHIDEE, and VIC-TEM-TOPMODEL reported their surface water areas; and SDGVM reported the area for which the water table was above a threshold depth, with the threshold chosen to minimize the global rms error between this area and GIEMS. Additionally, both CLM4Me and ORCHIDEE tied their surface water areas to the long-term mean of GIEMS: CLM4Me did so by calibration and ORCHIDEE did so by rescaling its surface water areas. Thus, we have placed these two models in the C category in Table 2.

Finally, the remaining models (IAP-RAS, LPJ-Bern, LPJ-WHyMe, LPX-BERN, LPX-BERN (N), both UW-VIC configurations, and all four VISIT configurations) used wetland maps, either alone or in combination with topography and surface water products, to inform their wetland schemes; these are denoted by “M” in Table 2. In most cases, the wetland maps were used to determine the maximum extent of the CH₄-producing area, within which inundated area and water table depths would vary in time. In contrast, LPJ-Bern, LPX-BERN, and LPX-BERN (N) allowed inundated area (specified by GIEMS) to sometimes exceed the static map-based peatland area; in such cases, it was assumed that the excess inundation occurred in mineral soils. Thus, the CH₄-producing area included peatlands and inundated mineral soils. LPJ-Bern additionally allowed CH₄ production in areas of “wet mineral soil” (in which soil moisture content was greater than 95 % of water-holding capacity) and included this in the total CH₄-producing area.

Models’ hydrologic approaches varied in other ways as well. Some (IAP-RAS and LPJ-WSL) did not include explicit water table depth formulations for estimating emissions in unsaturated (non-inundated) wetlands; IAP-RAS assumed all wetlands were completely saturated, and LPJ-WSL only considered unsaturated wetlands implicitly, using soil moisture as a proxy. Most of the other models used a TOPMODEL approach to relate the distribution of water table depths across the grid cell to topography (generally on a 1 km scale). However, LPJ-WHyMe, UW-VIC (GIEMS), and UW-VIC (SWAMPS) determined water table depth distributions within peatlands from assumed proportions of microtopographic landforms (e.g., hummocks and lawns) on the (horizontal) scale of meters. UW-VIC explicitly handled lakes by treating lakes and peatlands as a single system, spanning the total area of lakes and peatlands which was given by the Sheng et al. (2004) data set and within which surface water area varied dynamically. Areas of permanent surface water over the period 1949–2010 were considered to be lakes and were excluded from methane emissions estimates.

Models also varied in their soil thermal physics schemes. Most models used a one-dimensional heat diffusion scheme to determine the vertical profile of soil temperatures, but VISIT used a linear interpolation between current air temperature (at the soil surface) and annual average air temperature (at the bottom of the soil column). Several models (DLEM, LPJ-MPI, LPJ-WSL, and SDGVM) did not consider the water-ice phase change and therefore did not model permafrost. While IAP-RAS contained a permafrost scheme, it was driven by seasonal and annual summaries of meteorological forcings and used simple analytic functions to estimate the seasonal evolution and vertical profile of soil temperatures. Additionally, DLEM and LPJ-WSL did not con-
sider the insulating effects of organic (peat) soil. In contrast, UW-VIC modeled permafrost, peat soils, and the dynamics of surface water, including lake ice cover and evaporation, thereby adding another factor that influences soil temperatures.

Models also varied in their biogeochemical schemes (Table 3). Most represented methane production as a function of soil temperature, water table depth (except for IAP-RAS and LPJ-WSL), and the availability of carbon substrate. Most (except for IAP-RAS and LPJ-WSL) explicitly accounted for the oxidation of methane above the water table; and most accounted for some degree of plant-aided transport. Some models (LPJ-Bern, LPJ-MPI, LPJ-WhYMe, and LPX-BERN) represented methane production as either a constant or soil-moisture-dependent fraction of aerobic respiration. Some models (DLEM, DLEM2, and VIC-TEM-TOPMODEL) imposed additional dependences on soil pH and oxidation state. Models differed in the pathways and availability of carbon substrate; some models (UW-VIC, VIC-TEM-TOPMODEL, VISIT (GLWD-WH), and VISIT (Sheng-WH)) related carbon substrate availability to net primary productivity (NPP) as a proxy for root exudates; some (CLM4Me, IAP-RAS, LPJ-MPI, LPJ-WSL, ORCHIDEE, SDGVM, VISIT (GLWD), and VISIT (Sheng)) related carbon substrate to the content and residence times of various soil carbon reservoirs; and others (DLEM, DLEM2, LPJ-Bern, LPJ-WhYMe, all four LPX-BERN configurations) drew carbon substrate from a combination of both root exudates and soil carbon (or dissolved organic carbon, in the case of DLEM and DLEM2). CLM4Me and two configurations of LPX-BERN simulated interactions between the carbon and nitrogen cycles. Several models (all versions of LPJ and LPX, ORCHIDEE, and SDGVM) included dynamic vegetation components. Some models (LPJ-Bern, LPJ-MPI, LPJ-WhYMe, LPX-BERN, and UW-VIC) accounted for the inhibition of NPP of some plant species under saturated soil moisture conditions. Finally, models employed a variety of methods, alone or in combination (Table 3), to select parameter values, including taking the median of literature values, optimizing emissions to match in situ observations from representative sites regionally (e.g., UW-VIC optimized parameter values to match the Glagolev2011 data set in the WSL) or globally, or optimizing total global emissions to match various estimates from inversions.

<table>
<thead>
<tr>
<th>Model</th>
<th>$R_{\text{anaerobic}}/R_{\text{aerobic}}$</th>
<th>C substrate source</th>
<th>pH</th>
<th>Redox state</th>
<th>Dynamic vegetation</th>
<th>Nitrogen–carbon cycle interaction</th>
<th>Saturated NPP inhibition</th>
<th>Parameter selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLM4Me</td>
<td>Variable</td>
<td>Cpool</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Optimized to various sites</td>
</tr>
<tr>
<td>DLEM</td>
<td>Variable</td>
<td>NPP and Cpool</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Optimized to various sites</td>
</tr>
<tr>
<td>DLEM2</td>
<td>Variable</td>
<td>NPP and Cpool</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Optimized to various sites</td>
</tr>
<tr>
<td>IAP-RAS</td>
<td>a/a</td>
<td>Cpool</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Literature; scaled to global total</td>
</tr>
<tr>
<td>LPJ-Bern</td>
<td>Constant</td>
<td>NPP and Cpool</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Optimized to various sites; scaled to global total</td>
</tr>
<tr>
<td>LPJ-MPI</td>
<td>Constant</td>
<td>Cpool</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Literature</td>
<td>Literature; scaled to global total</td>
</tr>
<tr>
<td>LPJ-WhYMe</td>
<td>Constant</td>
<td>NPP and Cpool</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Literature</td>
</tr>
<tr>
<td>LPJ-WSL</td>
<td>Constant</td>
<td>Cpool</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Literature</td>
</tr>
<tr>
<td>LPX-BERN</td>
<td>Constant</td>
<td>NPP and Cpool</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Optimized to various sites; scaled to global total</td>
</tr>
<tr>
<td>LPX-BERN (DyPTOP)</td>
<td>Constant</td>
<td>NPP and Cpool</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Optimized to various sites; scaled to global total</td>
</tr>
<tr>
<td>LPX-BERN (N)</td>
<td>Constant</td>
<td>NPP and Cpool</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Optimized to various sites; scaled to global total</td>
</tr>
<tr>
<td>ORCHIDEE</td>
<td>Variable</td>
<td>Cpool</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Literature and optimized to various sites</td>
</tr>
<tr>
<td>SDGVM</td>
<td>Variable</td>
<td>Cpool</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Literature</td>
</tr>
<tr>
<td>UW-VIC (GIEMS)</td>
<td>Variable</td>
<td>NPP</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Optimized to sites in Glagolev2011</td>
</tr>
<tr>
<td>UW-VIC (SWAMPS)</td>
<td>Variable</td>
<td>NPP</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Optimized to sites in Glagolev2011</td>
</tr>
<tr>
<td>VIC-TEM-TOPMODEL</td>
<td>Variable</td>
<td>NPP</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Optimized to various sites</td>
</tr>
<tr>
<td>VISIT(GLWD)</td>
<td>Variable</td>
<td>Cpool</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Literature</td>
</tr>
<tr>
<td>VISIT(GLWD-WH)</td>
<td>Variable</td>
<td>NPP</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Literature</td>
</tr>
<tr>
<td>VISIT(Sheng)</td>
<td>Variable</td>
<td>Cpool</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Literature</td>
</tr>
<tr>
<td>VISIT(Sheng-WH)</td>
<td>Variable</td>
<td>NPP</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Literature</td>
</tr>
</tbody>
</table>

1 $R_{\text{anaerobic}}/R_{\text{aerobic}}$: from the ratio of anaerobic to aerobic respiration is handled in the model (constant: ratio is held constant; variable: ratio varies either as an explicit function of environmental conditions or as the result of separate governing equations for aerobic and anaerobic respiration; a/a: not applicable).

2 C substrate source: Cpool: soil carbon pool; NPP: root exudates, in proportion to net primary productivity.

3 pH: indicates whether soil pH influences CH$_4$ emissions.

4 Redox state: indicates whether soil redox status influences CH$_4$ emissions.

5 Dynamic vegetation: indicates whether vegetation species abundance change in response to environmental conditions.

6 Nitrogen–carbon cycle interaction: indicates whether interactions between the nitrogen and carbon cycles influence CH$_4$ emissions.

7 Saturated NPP inhibition: indicates whether NPP decreases under wet soil conditions for any plant species.

8 Parameter selection: method of choosing parameter values (literature: values chosen from ranges reported in literature; optimized: values chosen to minimize the difference between simulated and observed values, either of CH$_4$ fluxes at selected sites or of global atmospheric CH$_4$ concentrations).
2.5 Model simulations

To be consistent with WETCHIMP’s transient simulation (“Experiment 2-trans”, Wania et al., 2013), we focused our analysis on the period 1993–2004, although several non-WETCHIMP models provided data from 1993–2010. All models used the CRUNCEP gridded meteorological forcings (Viovy and Ciais, 2011) as a common input. Model-specific inputs are described in Wania et al. (2013).

Model outputs (monthly CH$_4$ emissions (average g CH$_4$ month$^{-1}$ m$^{-2}$ over the grid cell area) and monthly CH$_4$-producing area (km$^2$)) were analyzed at a 0.5° × 0.5° spatial resolution (resampled from native resolution as necessary).

Due to large seasonal variations in CH$_4$-producing areas, our analysis focused on June–July–August (JJA) averages of area and CH$_4$ emissions, since it is during these months that the majority of the year’s methane is emitted across all models (areas in other seasons would not be representative of annual CH$_4$ emissions). Similarly, in analyzing interannual variability in CH$_4$ emissions, we focused on JJA CH$_4$ emissions, which dominate the annual total and have stronger correlations with JJA environmental factors (such as air temperature, precipitation, or inundation) than annual CH$_4$ emissions have with annual average environmental factors. We also computed growing season CH$_4$ “intensities” (average JJA CH$_4$ emissions per unit JJA CH$_4$-producing area).

2.6 Data access

All data used in this study, including observational products, inversions, and forward model results, are available from WETCHIMP-WSL (2015).

3 Results

3.1 Average annual total emissions

As shown in Fig. 2 and Table S1 in the Supplement, 12-year mean estimates (±standard error on the mean) of annual total emissions over the WSL from forward models (5.34 ± 0.54 Tg CH$_4$ yr$^{-1}$), inversions (6.06 ± 1.22 Tg CH$_4$ yr$^{-1}$), and observations (3.91 ± 1.29 Tg CH$_4$ yr$^{-1}$) largely agreed, despite large scatter in individual estimates. Model estimates ranged from 2.42 (LPX-BERN (DyPTOP-N)) to 11.19 Tg CH$_4$ yr$^{-1}$ (IAP-RAS). The Glagolev2011 estimate was substantially lower than the mean of the models, corresponding to the 36th percentile of the distribution of model estimates. However, the potential upward revision of Glagolev2011 (Sect. 2.2) would move it to a substantially higher percentile of their distribution. Inversions yielded a similarly large range of estimates: 3.08 (Kim2011) to 9.80 Tg CH$_4$ yr$^{-1}$ (Winderlich2012). Despite their large spread, 15 out of...
the 17 forward models fell within the range of inversion estimates. Here we have excluded the “WH” configurations of VISIT and the configurations of LPX-BERN for which nitrogen–carbon interaction was turned off, due to their similarities to their counterparts that were included. The wide variety in the relative proportions of CH$_4$ emitted from the south and north halves of the domain, with the southern contribution ranging from 13 to 69% (right-hand column in Fig. 2), indicates a lack of agreement on which types of wetlands and climate conditions are producing the bulk of the region’s CH$_4$.

### 3.2 Differences among observational data sets

The large degree of disagreement among observational data sets is worth addressing before using them to evaluate the models. Important differences are evident among wetland maps (Fig. 3). Sheng2004 and Peregon2008 are extremely similar, in part because they both used the map of Romanova (1977) north of 65° N. Both of these data sets show wetlands distributed across most of the WSL, with large concentrations south of the Ob’ River (55–61° N, 70–85° E), east of the confluence of the Ob’ and Irtysh rivers (57–62° N, 65–70° E), and north of the Ob’ River (61–66° N, 70–80° E). In comparison, the GLWD map entirely lacks wetlands in the tundra region north of 67° N and shows additional wetland area in the northeast (64–67° N, 70–90° E). The NCSCD is substantially different from the other three maps. Owing to its focus on permafrost soils, it completely excludes the extensive wetlands south of the southern limit of permafrost (approximately 60° N). Given the numerous field studies documenting these productive southern wetlands (Sect. 2.1), the NCSCD seems to be inappropriate for studies that extend beyond permafrost.

The two surface water products (GIEMS and SWAMPS) also exhibit large differences. While they both agree that the surface water area fraction ($F_\text{w}$) is most extensive in the central region north of the Ob’ River (61–64° N), GIEMS gives areal extents that are 3–6 times those of SWAMPS. Outside of this central peak, GIEMS $F_\text{w}$ drops off rapidly to nearly 0 in most places (particularly in the forested region south of the Ob’ River, which may be due to difficulties in detecting inundation under vegetative canopy and/or reduced sensitivity where the open-water fraction is less than 10%; Prigent et al., 2007), while SWAMPS maintains low levels of $F_\text{w}$ through-
out most of the WSL. Along the Arctic coastline, SWAMPS shows high $F_w$, which may indicate contamination of the signal by the ocean. In both data sets, $F_w$ exhibits some similarity with the distribution of lakes and rivers (Fig. 1), illustrating the inclusion of non-wetlands in these surface water products.

Among the CH$_4$ data sets (Fig. 4), a clear difference can be seen between the spatial distributions of Glagolev2011 and Kim2011 (both of which assign the majority of emissions to the region south of the Ob’ River, between 55 and 60° N); and Winderlich2012 and Bousquet2011K (both of which assign the majority of emissions to the central region north of the Ob’ River, between 60 and 65° N). We discuss possible reasons for this discrepancy in Sect. 4.3. The global inversions (Bousquet2011R and K, and Bloom2010) have coarser spatial resolution than the regional inversions of Kim2011 and Winderlich2012. Bousquet2011R and K have similar distributions between 60 and 65° N, but Bousquet2011R has relatively stronger emissions between 57 and 60° N and weaker emissions between 65 and 67° N; in this respect, Bousquet2011R is intermediate between Glagolev2011 and Winderlich2012. Finally, Bloom2010 exhibits relatively little spatial variability in emissions, likely due to its use of GRACE observations as a proxy for wetland inundation and water table conditions.

### 3.3 Primary drivers of model spatial uncertainty

The wide disagreement among models is plainly evident in Fig. 5, which plots average JJA CH$_4$ emissions versus average JJA CH$_4$-producing areas for the WSL as a whole (top left), the south (bottom left), and the north (bottom right). A series of lines (“spokes”) passing through the origin, with slopes of integer multiples of 1 g CH$_4$ m$^{-2}$ month$^{-1}$, allows comparison of spatial average intensities (CH$_4$ emissions per unit CH$_4$-producing area). All points along a given line have the same intensity but different CH$_4$-producing areas. We have included the Glagolev2011–Peregon2008 CH$_4$–area estimate (denoted by a black star) and the mean of the inversions (denoted by a grey star) for reference. We set the area coordinate for the inversions to Peregon2008 because (a) the wetland area was not available for all inversions and (b) Peregon2008 is a relatively accurate estimate of wetland area. JJA CH$_4$ emissions, JJA wetland or CH$_4$-producing areas, and JJA intensities, for all models, observations, and inversions,
Figure 5. Model estimates of JJA $\text{CH}_4$ emissions ($\text{Tg CH}_4 \text{ month}^{-1}$) and JJA wetland or $\text{CH}_4$-producing area ($10^3 \text{ km}^2$): for the entire WSL (top left) and the southern (bottom left) and northern (bottom right) halves, for the period 1993–2004. Lines passing through the origin, with slopes of integer multiples of $1 \text{ g CH}_4 \text{ m}^{-2} \text{ month}^{-1}$, allow a comparison of spatial average intensities ($\text{CH}_4$ emissions per unit $\text{CH}_4$-producing area). Circles denote models that used satellite surface water products alone (corresponding to code S in Table 2) to delineate wetlands. Triangles denote models that used topographic information, with or without surface water products (corresponding to code T in Table 2). Squares denote models that used wetland maps with or without topography or surface water products (corresponding to code M in Table 2).

Table 4. Estimates of June–July–August $\text{CH}_4$ emissions from subsets of the participating models, over the entire WSL and its southern ($< 61^\circ \text{N}$) and northern halves, for the period 1993–2004. Biases were computed with respect to the Glagolev2011–Peregon2008 estimates.

<table>
<thead>
<tr>
<th>Subset</th>
<th>Average Jun–Jul–Aug $\text{CH}_4$ emissions ($\text{Tg CH}_4 \text{ month}^{-1}$)</th>
<th>Average Jun–Jul–Aug contributing area ($10^3 \text{ km}^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WSL</td>
<td>South</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>Bias</td>
</tr>
<tr>
<td>I</td>
<td>1.10</td>
<td>0.14</td>
</tr>
<tr>
<td>T</td>
<td>1.42</td>
<td>0.46</td>
</tr>
<tr>
<td>M</td>
<td>1.32</td>
<td>0.36</td>
</tr>
<tr>
<td>M+</td>
<td>1.30</td>
<td>0.34</td>
</tr>
</tbody>
</table>

are listed in Table S1. Over the entire WSL (Fig. 5, top left), the scatter in model estimates of $\text{CH}_4$ emissions results from scatter in both area (ranging from 200 000 to 1200 000 km$^2$) and intensity (ranging from 1 to 8 g CH$^4$ m$^{-2}$ month$^{-1}$), with no clear relationship between the two.

However, a strong area-driven bias is evident in the south (Fig. 5, bottom left). Although the mean modeled $\text{CH}_4$ emission rate (0.58 Tg CH$^4$ month$^{-1}$) is fairly close to both Glagolev2011 (0.67 Tg CH$^4$ month$^{-1}$) and the mean of inversions (0.60 Tg CH$^4$ month$^{-1}$), the distribution of model estimates is substantially skewed, with most models’ estimates falling well below both Glagolev2011 and the mean of the inversions. Glagolev2011’s estimate corresponds to the 81st percentile of the model CH$^4$ distribution; the expected upward revision of Glagolev2011 (Sect. 2.2; exact JJA amount not yet known) would only raise that percentile. The mean of the inversions corresponds to the 76th percentile. Similarly, the models substantially underestimate the CH$^4$-producing area, with Peregon2008 occupying the 83rd percentile of the model distribution. On the other hand,
Table 5. Spatial correlations between simulated average annual CH$_4$ emissions and GIEMS surface water area fraction (F$_w$).

<table>
<thead>
<tr>
<th>Model</th>
<th>Correlation</th>
<th>Model</th>
<th>Correlation</th>
<th>Model</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLM4Me</td>
<td>0.69</td>
<td>LPJ-WHyMe</td>
<td>0.45</td>
<td>UW-VIC (GIEMS)</td>
<td>0.44</td>
</tr>
<tr>
<td>DLEM</td>
<td>0.70</td>
<td>LPJ-WSL</td>
<td>0.97</td>
<td>UW-VIC (SWAMPS)</td>
<td>0.11</td>
</tr>
<tr>
<td>DLEM2</td>
<td>0.21</td>
<td>LPX-BERN (N)</td>
<td>0.41</td>
<td>VIC-TEM-TOPMODEL</td>
<td>0.41</td>
</tr>
<tr>
<td>IAP-RAS</td>
<td>-0.03</td>
<td>LPX-BERN (DyPTOP-N)</td>
<td>0.28</td>
<td>VISIT (GLWD)</td>
<td>0.62</td>
</tr>
<tr>
<td>LPJ-Bern</td>
<td>0.56</td>
<td>ORCHIDEE</td>
<td>0.61</td>
<td>VISIT (Sheng)</td>
<td>0.65</td>
</tr>
<tr>
<td>LPJ-MPI</td>
<td>0.01</td>
<td>SDGVM</td>
<td>0.09</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6. Mean CH$_4$ emissions from LPX-BERN, 1993–2010, for the entire WSL and the south and north halves of the domain.

<table>
<thead>
<tr>
<th>Configuration</th>
<th>Mean [Tg CH$_4$ yr$^{-1}$]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WSL</td>
</tr>
<tr>
<td>LPX-BERN</td>
<td>3.81</td>
</tr>
<tr>
<td>LPX-BERN (DyPTOP)</td>
<td>3.17</td>
</tr>
<tr>
<td>LPX-BERN (N)</td>
<td>3.08</td>
</tr>
<tr>
<td>LPX-BERN (DyPTOP-N)</td>
<td>2.44</td>
</tr>
</tbody>
</table>

Differences

<table>
<thead>
<tr>
<th>Differences</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>LPX-BERN (N) – LPX-BERN</td>
<td>-0.73</td>
</tr>
<tr>
<td>LPX-BERN (DyPTOP-N) – LPX_BERN (DyPTOP)</td>
<td>-0.73</td>
</tr>
<tr>
<td>LPX-BERN (DyPTOP) – LPX-BERN</td>
<td>-0.64</td>
</tr>
<tr>
<td>LPX-BERN (DyPTOP-N) – LPX-BERN (N)</td>
<td>-0.64</td>
</tr>
</tbody>
</table>

the model intensity distribution is much less biased, with Glagolev2011 corresponding to the 47th percentile. Even a doubling of Glagolev2011’s intensity would place it at only the 69th percentile of the model distribution, a smaller bias than for area. Thus, the area bias is the major driver of CH$_4$ bias in the south. In comparison, the north (Fig. 5, bottom right) is relatively unbiased.

Model inputs and formulations played a key role in determining CH$_4$-producing area biases. Statistics of model performance relative to Glagolev2011–Peregon2008, categorized by the wetland codes in Table 2, are listed in Table 4. The models that used satellite surface water products alone (denoted by circles in Fig. 5 and the code S in Table 2) estimated the lowest CH$_4$-producing areas in the south, with a bias of $-270,000$ km$^2$ and standard deviation of $31,000$ km$^2$. Additionally, two models (LPJ-Bern and LPJ-WHyMe) from the M group (denoted by squares in Fig. 5 and the code M in Table 2) also yielded low areas, due to their use of the NCSCD map, which omitted non-permafrost wetlands. The “M+” group, consisting of all M models except those two, exhibited the smallest bias and second-smallest standard deviation ($-31,000$ and $34,000$ km$^2$, respectively). Models that determined CH$_4$-producing area dynamically using topographic data but without the additional input of wetland maps (denoted by triangles in Fig. 5 and the code T in Table 2) yielded nearly as small a bias as the M+ group ($-42,000$ km$^2$) but had the largest scatter (standard deviation of $173,000$ km$^2$) of the groups. The fact that two of the S models (CLM4Me and ORCHIDEE) supplied CH$_4$-producing areas that excluded non-inundated methane-emitting wetlands had little effect on the results, since their total CH$_4$ emissions (which included non-inundated emissions) also suffered from a large negative bias ($-0.45$ Tg CH$_4$ yr$^{-1}$, or $-67\%$).

Examining the spatial distributions of annual CH$_4$ (Fig. 6) and JJA CH$_4$-producing areas (Fig. 7) shows why the use of surface water data alone results in poor model performance. Among the models from the S group (CLM4Me, DLEM, DLEM2, LPJ-WSL, and ORCHIDEE), the spatial distributions of both CH$_4$ emissions and CH$_4$-producing area tend to be strongly correlated with GIEMS (See Table 5 for correlations), which exhibits very low surface water areas south of the Ob’ River, despite the large expanses of wetlands there (Sect. 3.2). Similarly, the low emissions of LPJ-WHyMe and LPJ-Bern in the south can be explained by their use of the NCSCD map, which only considered peatlands (Histels and Histosols) within the circumpolar permafrost zones (which only occur north of $60\degree$ N). For LPJ-WHyMe, these permafrost peatlands were the only type of wetland modeled (i.e., the model domain only included the circumpolar permafrost zones), so LPJ-WHyMe’s emissions were almost nonexistent in the south. LPJ-Bern also used the NCSCD’s Histels and Histosols to delineate peatlands but additionally simulated methane dynamics in wet or inundated mineral soils outside the permafrost zone. While this allowed LPJ-Bern to make emissions estimates in the south, the much
lower porosities of mineral soils resulted in larger sensitivities of water table depth to evaporative loss than those of peat soils. These drier soils led to net CH$_4$ oxidation in much of the south.

Aside from area-driven biases, a large degree of intensity-driven scatter is evident in both the south and north. Indeed, the underestimation of areas in the south, accompanied by resulting reductions in CH$_4$ emissions, partially compensated for some of the intensity-driven scatter there. However, some of the more extreme intensities were arguably the result of area biases, in that some of the global wetland models (CLM4Me, IAP-RAS, LPJ-Bern, and LPJ-WHyMe) scaled their intensities to match their global total emissions with those of global inversions, which could result in local biases if their wetland maps suffered from either global or local bias (which was true of these models). Interestingly, several models yielded estimates similar to those of the two regionally optimized UW-VIC simulations, implying that the regional optimization did not confer a distinct advantage on UW-VIC.

Nitrogen limitation influenced intensity in LPX-BERN, the one model that included it. Although we did not plot results from the two LPX-BERN configurations that lacked nitrogen–carbon interactions in Fig. 5, we compare results from all four LPX-BERN configurations in Table 6. In LPX-BERN (N) and LPX-BERN (DyPTOP-N), the nitrogen limitation imposed by nitrogen–carbon interactions substantially reduced NPP, relative to LPX-BERN and LPX-BERN (DyPTOP), leading to a reduction of mean annual CH$_4$ emissions of approximately 20% over the entire WSL over the period 1993–2010. This reduction was slightly larger than the difference in emissions between simulations using the Sheng2004 map to prescribe peatland area (LPX-BERN and LPX-BERN (N)) and simulations using the DyPTOP method to determine peatland extent dynamically (LPX-BERN (DyPTOP) and LPX-BERN (DyPTOP-N)). In addition, the reduction in emissions due to nitrogen limitation was concentrated in the northern half of the domain, in contrast to the reduction due to dynamic peatland extent, which was concentrated in the southern half of the domain. Nitrogen limitation also reduced trends in CH$_4$ emissions over the entire WSL over the period 1993–2010, through reductions in soil carbon accumulation rates. However, both these trends and their reductions were very small (< 0.5% per year in most cases) and statistically insignificant over the study period.
3.4 Model temporal uncertainty and major environmental drivers

3.4.1 Average seasonal cycles

Models demonstrated general agreement on the shape of the seasonal cycle of emissions (Fig. 8, top left) and intensities (Fig. 8, bottom right), despite wide disagreement on the shape and timing of the seasonal cycle of the CH\textsubscript{4}-producing area (Fig. 8, bottom left). The regional inversions (Kim2011 and Winderlich2012) agreed on a July peak for CH\textsubscript{4}, although Winderlich2012 suggested a noticeably larger contribution from cold season months than the others (which is plausible, given reports of non-zero winter emissions; Rinne et al., 2007; Kim et al., 2007; Panikov and Dedysh, 2000). In contrast, both Bousquet inversions peaked in August. Unlike the other three inversions, the Bousquet2011R inversion had negative emissions (net oxidation) in either May or June.
of almost every year of its record. These negative emissions were widespread, throughout not only the WSL but the entire boreal Asia region, and cast doubt on the accuracy of their seasonal cycle. Turning to the surface water products (Fig. 8, bottom left), GIEMS and SWAMPS displayed quite different shapes in their seasonal cycles of surface water extent: GIEMS exhibited a sharp peak in June and SWAMPS displayed a broad, flat maximum from June through September. In fact, SWAMPS had a similar shape to GIEMS south of about $64^\circ$N; the broad peak for the WSL as a whole was the result of late-season peaks further north.

Most models’ CH$_4$ emissions peaked in July, in agreement with the regional inversions. A few models peaked in June: CLM4Me, DLEM2, LPJ-MPI, VISIT (GLWD), and VISIT (Sheng). Correspondingly early peaks in intensity can explain the early peaks in the DLEM2 and the VISIT simulations, indicating either early availability of carbon substrate in the soil or rapid soil warming (the latter is likely for VISIT, given its linearly interpolated soil temperatures). In contrast, LPJ-MPI’s early peak in emissions was the result of an early (May) peak in CH$_4$-producing area, which, in turn, was the result of early snowmelt. Two models (LPJ-BERN and UW-VIC (GIEMS)) peaked in August. LPJ-Bern’s late peak resulted from a late peak in wet mineral soil intensity, despite an exceptionally late (October) peak in CH$_4$-producing area.

The late peak of UW-VIC (GIEMS) corresponded to a late peak in intensity, implying either late availability of carbon substrate (due to inhibition of NPP under inundation) or delayed warming of the soil (due to excessive insulation by peat or surface water).

Aside from the above cases, the relative agreement among models on a July peak in CH$_4$ emissions comes despite wide variation in seasonal cycles of the CH$_4$-producing area. For example, DLEM’s CH$_4$-producing area held steady at its maximum extent from April through November, and VIC-TEM-TOPMODEL’s CH$_4$-producing area peaked in August, possibly due to low evapotranspiration or runoff rates. Some of the discrepancies in CH$_4$-producing area seasonality arose from several models using static maps to define some or all wetland areas (Sects. 2.3 and 2.4). These differences matter little to the seasonal cycle of CH$_4$ emissions, in part because of the similarity between the seasonal cycles of inundated area and water table depths within the static CH$_4$-producing areas and in part because of the nearly universal strong correlation at seasonal timescales between simulated intensities and near-surface air temperature (so that cold-season CH$_4$-producing areas have little influence over emissions).
3.4.2 Interannual variability

At multiyear timescales (shown for the period 1993–2010 in Fig. 9), models’ and inversions’ total annual CH$_4$ emissions displayed a wide range of interannual variability, even after accounting for the effects of differences in intensity. Values of the coefficient of variation (CV) for models over the period 1993–2004 ranged from 0.069 (LPX-BERN (N)) to 0.338 (UW-VIC (GIEMS)) with a mean of 0.169 (Table 7). While Bousquet2011K’s CV of 0.160 fell near the mean model CV, Bousquet2011R’s CV of 0.446 was 25% larger than the largest model CV, and over twice the second-largest model CV. Bousquet2011R’s high variability was due in part to a peak in CH$_4$ emissions in 2002 followed by a large drop in emissions between 2002 and 2004, actually becoming negative (net CH$_4$ oxidation) in 2004 before continuing at a much lower mean value from 2005 to 2009. This peak and decline coincide with a similar peak and decline in...


<table>
<thead>
<tr>
<th></th>
<th>WSL</th>
<th>CRU T JJA</th>
<th>CRU P JJA</th>
<th>SWAMPS JJA</th>
<th>GIEMS JJA</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRU T JJA</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRU P JJA</td>
<td>-0.10</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SWAMPS JJA</td>
<td>0.14</td>
<td>0.66</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GIEMS JJA</td>
<td>-0.11</td>
<td>0.44</td>
<td>0.68</td>
<td>1.00</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>S</th>
<th>CRU T JJA</th>
<th>CRU P JJA</th>
<th>SWAMPS JJA</th>
<th>GIEMS JJA</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRU T JJA</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRU P JJA</td>
<td>-0.28</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SWAMPS JJA</td>
<td>-0.12</td>
<td>0.44</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GIEMS JJA</td>
<td>-0.10</td>
<td>0.22</td>
<td>0.87</td>
<td>1.00</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>CRU T JJA</th>
<th>CRU P JJA</th>
<th>SWAMPS JJA</th>
<th>GIEMS JJA</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRU T JJA</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRU P JJA</td>
<td>-0.06</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SWAMPS JJA</td>
<td>0.32</td>
<td>0.60</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GIEMS JJA</td>
<td>-0.05</td>
<td>0.34</td>
<td>0.61</td>
<td>1.00</td>
<td></td>
</tr>
</tbody>
</table>

The relative strengths of the correlations between models’ CH$_4$ emissions and drivers varied widely, as shown in the scatterplots in Fig. 12. Over the entire WSL (top left) as well as the south and north halves (bottom left and right), the low correlation between $T_{\text{air}}$ and $F_w$ led to consistent trade-offs in the correlations between simulated emissions and $T_{\text{air}}$ (x axis) or $F_w$ (y axis). Some models (all four LPX-BERN simulations, all four VISIT simulations, IAP-RAS, ORCHIDEE, and SDGVM) had correlations with $T_{\text{air}}$ that were greater than 0.7 in one or both halves of the domain; since this means that $T_{\text{air}}$ would explain the majority of CH$_4$ variance in a linear model, we have denoted them as “$T_{\text{air}}$-dominated”. Other models (DLEM, LPJ-WSL, DLEM2, and LPJ-MPI) were “$F_w$-dominated” in one or both halves of the domain. For the other models and inversions, no driver explained the majority of the variance. A few models had small enough contributions from one or the other driver for the resulting correlations to be negative, due to the small negative correlation between $T_{\text{air}}$ and $F_w$. Neither of the two Bousquet2011 inversions exhibited strong correlations with either $F_w$ or $T_{\text{air}}$, which might imply that models also should not exhibit strong correlations with one driver.

Indeed, the overarching pattern in the model correlations was that models that lacked physical and biochemical formulations appropriate to the high latitudes exhibited stronger correlations with inundation or air temperature than either the inversions or more sophisticated models. One characteristic that most of the $F_w$-dominated models (except for DLEM2) have in common is that they lack soil thermal formulations that account for soil freeze–thaw processes; conversely, most of the non-$F_w$-dominated models do have such formulations. In addition, inundated fractions of DLEM, DLEM2, and LPJ-WSL were explicitly driven by GIEMS $F_w$. Unlike the other three models, LPJ-MPI does account for the thermal effects $F_w$ (Fig. 10) and precipitation (Fig. 11). Several models (notably LPJ-MPI, LPJ-WHYMe, LPJ-WSL, DLEM, and VIC-TEM-TOPMODEL), as well as Bousquet2011K, mirrored this drop to varying degrees, but none dropped as much in proportion to their means or became negative. In contrast, Bloom2010, spanning only the period 2003–2007, exhibited extremely little interannual variability, perhaps due to its use of GRACE as a proxy for inundated area and water table depth.

To investigate the influence of various climate drivers on CH$_4$ emissions, we computed the individual correlations between the JJA CH$_4$ emissions and the following JJA drivers: CRU air temperature ($T_{\text{air}}$), CRU precipitation (P), GIEMS $F_w$, and SWAMPS $F_w$, for forward models and the two Bousquet2011 inversions, over the period 1993–2004 (Table S2). Here we included four additional model configurations that we did not show in previous sections: VISIT (GIEMS-WH), VISIT (SHENG-WH), LPX-BERN, and LPX-BERN-DyPTOP. The two drivers yielding the highest correlations with JJA CH$_4$ emissions were JIA CRU $T_{\text{air}}$ and JIA GIEMS $F_w$. These two drivers also exhibited nearly zero correlation with each other over the WSL and the south and north halves (Table 8). Because variations in water table position are driven by the same hydrologic factors (snowmelt, rainfall, evaportranspiration, and drainage) that drive variations in $F_w$, correlation with $F_w$ should serve as a general measure of the influence of both surface and sub-surface moisture conditions on methane emissions, even for models that were not explicitly driven by $F_w$. Therefore, we chose to examine model behavior in terms of correlations with JIA CRU $T_{\text{air}}$ and JIA GIEMS $F_w$. As an aside, this choice was not an endorsement of GIEMS over SWAMPS (which yielded qualitatively similar results to GIEMS); it simply resulted in better separation among models.
of peat soils, which might explain LPJ-MPI’s low (slightly negative) correlation with air temperature.

Some of the $T_{\text{air}}$-dominated models also lack sophisticated soil thermal physics. VISIT’s strong correlation with $T_{\text{air}}$ can be explained by the fact that its soil temperature scheme is a simple linear interpolation between current air temperature at the surface and annual average air temperature at the bottom of the soil column; as a result, VISIT’s soil temperature has a 1.0 correlation with air temperature. Comparing the WH configurations of VISIT to the default configurations, the model of Walter and Heimann (2000) had a lower correlation with air temperature than the Cao (1996) model. SDGVM also lacks soil freeze–thaw dynamics. IAP-RAS assumes all wetlands are completely saturated and holds their areas constant in time; as a result, its CH$_4$ emissions have no dependence on soil moisture or $F_w$ but a strong dependence on air temperature. LPX-BERN’s high correlation with air temperature is the result of a relative insensitivity of CH$_4$ emissions to water table depth, but at present there are too few sites with multi-year observations in the region to determine whether this low sensitivity is reasonable. Nitrogen–carbon interaction (LPX-BERN (N) and LPX-BERN (DyPTOP-N)) appeared to have only a minor effect on LPX-BERN’s interannual variability in the north but led to a slight reduction in correlation with $T_{\text{air}}$ in the south. Finally, UW-VIC (GIEMS) had small negative correlations with both $T_{\text{air}}$ and $F_w$ in the north, likely the result of its surface water formulation. UW-VIC’s surface water dynamics had been initially calibrated using the SWAMPS product: the much larger surface water extents of GIEMS in the north resulted in substantially deeper surface water, with corresponding insulating effects, greater evaporative cooling, and longer residence times, thus lowering correlations with both observed $F_w$ and $T_{\text{air}}$. The large difference in behavior between UW-VIC (GIEMS) and UW-VIC (SWAMPS) implies that the differences arising from optimizing surface water dynamics to different products far outweighed the differences between UW-VIC and other models in their selection of biogeochemical parameters.

4 Discussion

4.1 Long-term means and spatial distributions

The most striking finding, in terms of long-term means and spatial distributions, was the substantial bias in CH$_4$ emissions that resulted from using satellite surface water products or inaccurate wetland maps to delineate wetlands. Surface water is an important component of wetland models, but it clearly is a poor proxy for wetland extent at high latitudes because it both excludes the large expanses of strongly emitting non-inundated peatlands that exist there (Sect. 2.1) that were missed by GIEMS and underrepresented by SWAMPS and erroneously includes the high concentrations of large lakes there (e.g., Lehner and Döll, 2004), which do not necessarily emit methane at the same rates or via the same carbon cycling processes as wetlands (e.g., Walter et al., 2006; Pace et al., 2004). The practical difficulties in detecting inundation under forest canopies with visible or high-frequency microwave sensors (e.g., Sippel and Hamilton, 1994) compound these problems. In the case of the WSL, equating wetlands with surface water not only caused underestimation of total CH$_4$ emissions but also led to the attribution of the majority of the region’s emissions to the permafrost zone in the north. This issue is not unique to the WSL, as the collocation of permafrost, lakes, and inundation is present throughout the high latitudes (Tarnocai et al., 2009; Lehner
Figure 12. Influence of interannual variations in surface water area fraction ($F_{sw}$) on model CH$_4$ emissions (expressed as correlation between JJA GIEMS $F_{sw}$ and JJA CH$_4$) vs. influence of air temperature ($T_{air}$) on model CH$_4$ emissions (expressed as correlation between JJA CRU $T_{air}$ and JJA CH$_4$), for the entire WSL (top) and the southern and northern halves of the domain (bottom). $F_{sw}$-Dominated and $T_{air}$-Dominated denote correlation thresholds above which surface water area or air temperature, respectively, explain more than 50% of the variance in CH$_4$ emissions. Circles denote models that used satellite surface water products alone (corresponding to code S in Table 2) to delineate wetlands. Triangles denote models that used topographic information, with or without surface water products (corresponding to code T in Table 2). Squares denote models that used wetland maps with or without topography or surface water products (corresponding to code M in Table 2).

and Döll, 2004; Brown et al., 1998). Indeed, in their analysis of the Hudson Bay Lowland (HBL), Melton et al. (2013) found that three of the four lowest emissions estimates were from S models (CLM4Me, DLEM, and LPJ-WSL), although whether this was due to a bias in area was not examined. Given present concerns over the potential liberation of labile carbon from thawing permafrost over the next century (Koven et al., 2011), it is crucial to avoid under- or overestimating emissions from permafrost wetlands.

It is therefore important for modelers – both forward and inverse – to use accurate wetland maps such as Peregon et al. (2008), Sheng et al. (2004), or Lehner and Döll (2004) in their model development, whether as a static input parameter or as a reference for evaluating prognostically computed CH$_4$-producing areas, and to account for the existence of non-inundated portions within these wetlands in which methane emissions have a dependence on water table depth. Maps such as Tarnocai et al. (2009) may be inappropriate unless restricting simulations to permafrost wetlands. Ideally, modelers would be able to draw on a global version of the high-resolution map of Peregon et al. (2008) that not only delineates wetlands but also identifies the major subtypes (e.g., sphagnum-dominated or sedge-dominated, as in Lupascu et al., 2012) to which different methane emissions parameters could potentially be applied. When using surface water products to constrain simulated inundated extents, modelers must be sure either to mask out permanent lakes and large rivers, using a data set such as GLWD (Lehner and Döll, 2004) or MOD44W (Carroll et al., 2009), or better, to implement carbon cycling processes that are appropriate to these forms of surface water.

4.2 Temporal variability, environmental drivers, and model features

Another notable finding was that models that lacked physical and biochemical formulations appropriate to the high latitudes exhibited more extreme correlations with $F_{sw}$ or air temperature than either inversions or more sophisticated models. In other words, high-latitude biogeoophysical processes – specifically, soil freeze–thaw, the insulating effects of snow and peat, and relationships between emissions and water table depth in peatlands – make a substantial difference
to the sensitivities of emissions to environmental drivers, at least over the 12-year period of this study. Even if we do not fully trust the Bousquet2011 inversions, it seems reasonable to assume that the models that simulate high-latitude-specific processes are more likely to be correct in this regard than the other models. These sensitivities have a bearing on models’ responses to potential future climate change (e.g., Riley et al., 2011; Koven et al., 2011).

Thus, it appears that the following model features are desirable for reliable simulations of boreal wetlands:

- realistic soil thermal physics, including freeze–thaw dynamics. Most of the models that were highly correlated with one driver (LPJ-WSL, DLEM, LPJ-MPI, VISIT, and SDGVM) lacked this feature.

- accurate representations of peat soils. Again, many of the models with high correlations with one driver (LPJ-WSL, DLEM, VISIT, and SDGVM) lacked this feature.

- realistic representations of unsaturated (non-inundated) peatlands, including the dependence of CH₄ emissions on water table depth. LPJ-WSL, an Fw-dominated model, effectively set non-inundated CH₄ emissions to 0 because it did not simulate wetlands outside of the time-varying GIEMS surface water area. At the other extreme, IAP-RAS, a Tair-dominated model, treated all wetlands in their static map as if they were saturated, thereby eliminating the contribution of soil moisture variability. The relative insensitivity of LPX-BERN's emissions to water table position similarly reduced the contribution of soil moisture variability, although there are too few observations to say whether this is unreasonable.

Other model features either made relatively little difference in this study or were severely underrepresented but warrant further investigation. This is especially true of biogeochemical processes. For example, whether models contained dynamic vegetation (phenology and/or community composition) or dynamic peatland (peat accumulation and loss) components did not affect performance. However, our 12-year study period was likely too short to see the effects of these features. Changes in vegetation community composition may become more important in end-of-century projections (e.g., Alo and Wang, 2008; Kaplan and New, 2006). In particular, recent studies (Koven et al., 2011; Ringeval et al., 2011; Riley et al., 2011) have found a “wetland feedback”, in which vegetation growth in response to future climate change can lower water tables and reduce inundated extents via increased evapotranspiration. This drying effect reduces end-of-century CH₄ emissions from an approximate doubling of current rates without the feedback to only a 20–30% increase with the feedback. Similarly, hydrologic and chemical changes in peat soils, in response to disturbances such as permafrost thaw or drainage for mining or agricultural purposes, may be important in end-of-century projections (e.g., Strack et al., 2004). However, to properly assess the accuracy of dynamic vegetation or peatland schemes and their effects on CH₄ emissions, a longer historical study period, along with longer observational records (including observations of species compositions and soil carbon densities) would be necessary.

Other features may warrant further study. Replacing the Cao (1996) model with the model of Walter and Heimann (2000) modestly lowered VISIT’s otherwise extreme correlation with Tair. It is not clear if this is an inherent difference between the two formulations or just an artifact of their parameter values in VISIT, but it might imply that the Walter and Heimann model is more appropriate for applications at high latitudes. Similarly, nitrogen–carbon interaction had a substantial latitude-dependent effect on mean CH₄ emissions for LPX-BERN (Table 6). Again, the size of the effect could be model-dependent, and potential impacts on sensitivities to climate change might become more apparent over a longer analysis period.

Some of the scatter in model sensitivities to drivers may come from differences in the values of parameters related to methane production, methane oxidation, and plant-aided transport, which recent studies (Riley et al., 2011; Berrittella and van Husisden, 2011) have found to be particularly influential over wetland CH₄ emissions. The investigation of these parameters over the WSL in a model intercomparison can be difficult due to the many large differences among model formulations. As shown in Sects. 3.3 and 3.4.2, the methods of biogeochemical parameter selection had far less influence over the model results than the presence or absence of major features such as sophisticated soil thermal physics. Such a comparison would require the examination of a subset of the models that have sufficiently similar snow, soil, and water table formulations in order to isolate the effects of microbial and vegetative parameters.

Other features that were not investigated here could have potentially large impacts on the response of high-latitude wetlands to future climate change. One such feature is acclimatization, in which soil microbial communities gradually adapt to the long-term mean soil temperature. This feature has been explored in the ORCHIDEE model (Koven et al., 2011; Ringeval et al., 2010), where it greatly reduced the response of wetland CH₄ emissions to long-term temperature changes. Unfortunately, the version of ORCHIDEE used in this study and in the original WETCHIMP study (Melton et al., 2013; Wania et al., 2013) did not use acclimatization. Acclimatization likely would lower ORCHIDEE’s correlation with Tair over timescales long enough for changes in the long-term mean to be as large as interannual anomalies. Another feature explored by Koven et al. (2011) is the liberation of ancient labile carbon stored in permafrost. As with dynamic vegetation, a robust evaluation of these effects would require a much longer study period.
4.3 Future needs for observations and inversions

The wide disagreement among estimates from observations and inversions hampers our ability to assess model performance. Given the large influence that wetland maps can have on emissions estimates (not only in the WSL, but over larger areas, as shown by Petrescu et al., 2010), care must be taken to select appropriate maps. Ideally, global satellite or map products such as the GLWD (which omitted the northernmost wetlands in the WSL) should be validated against more intensively ground-truthed regional maps, such as Sheng2004 and Peregon2008, where such maps exist. Similarly, resolving the discrepancies between the GIEMS and SWAMPS remote-sensing surface water products would require verification against independent observations.

The large discrepancy between the spatial distributions of emissions from Glagolev2011 and Kim2011 (concentrated in the south) and Winderlich2012 and Bousquet2011K (concentrated in the north) may be due to several factors. First, the inversions’ posterior estimates reflect their prior distributions: Kim2011 used an earlier version of Glagolev2011 (Glagolev et al., 2010) as its prior, while Winderlich2012 and Bousquet2011K both used the Kaplan (2002) distribution as their prior. Second, different types and locations of observations were used: Glagolev2011 was based on in situ chamber measurements of CH₄ fluxes, 80% of which were obtained south of the Ob River, while Winderlich2012 was based on atmospheric CH₄ concentrations observed at towers near or north of the Ob River. Third, observations were not taken from the same years. Finally, the Winderlich2012 wetland CH₄ emissions may have been influenced by assumed emission rates from fossil fuel extraction and biomass burning, which were not adjusted during the inversion. Efforts like the revision of Glagolev2011 will certainly help in resolving some discrepancies, but all estimates would benefit from incorporating observations over long time periods and wider areas to reduce uncertainties in their long-term means.

The global inversions were also subject to uncertainties. For example, while the Bousquet2011 inversions imply that wetland CH₄ emissions in the WSL are not strongly correlated with either Fₒ or air temperature, the Bousquet2011 inversions’ temporal behaviors must be evaluated with caution. The reference inversion’s coefficient of variability (CV), which resulted in net negative annual emissions over the WSL in 2004, was substantially higher than the highest model CV. Bousquet et al. (2006) noted that their inversions were more sensitive to the interannual variability of wetland emissions than to their mean; accordingly, it is possible that the Bousquet2011 inversions underestimated the long-term mean, thereby raising the CV. Another possibility is that the monthly coefficients that optimized total emissions over all of boreal Asia were not optimal over the WSL alone, since the environmental drivers interacting with wetlands elsewhere may not have been in phase with those in the WSL. A further possibility, given credence by the reference inversion’s consistent net negative emissions over all of boreal Asia in May and June, is that errors in other components of the inversion (e.g., atmospheric OH concentrations, methane oxidation rates, background methane concentrations advected from elsewhere) influenced wetland emissions. Finally, other methane sources that were not accounted for in the inversion might have been attributed to wetlands, for example, geological CH₄ seeps (Etiope et al., 2008), leaks from gas pipelines (Ulmishek, 2003), or lakes (Walter et al., 2006).

At the other extreme, the Bloom2010 product exhibited almost no spatial or temporal variability. This might be an artifact of using GRACE data as a proxy for wetland inundation and water table levels. The spatiotemporal accuracy of Bloom2010 must also be questioned, given that it did not use an atmospheric transport model or account for methane oxidation in the atmosphere. Thus, while Bloom2010 provided a useful estimate of long-term mean emissions, it was less helpful in constraining model responses to climate drivers.

Another general limitation of inversions and observations, distinct from estimates of long-term mean emissions, is the lack of sufficiently long periods of record to assess model sensitivities to environmental drivers and climate change. The Bousquet2011 inversions and the SWAMPS surface water product are long enough to begin to address this issue on the global scale, but the Bousquet2011 inversions are not optimized for the WSL. Regional inversions such as Kim2011 and Winderlich2012, which might offer more spatially accurate estimates for the WSL than the Bousquet2011 inversions, only offer a single year of posterior emissions. Long records of in situ observations of CH₄ emissions and the factors that most directly influence these emissions (e.g., soil temperature and water table depth) only exist in a handful of locations (e.g., the Bakchar Bog in the WSL; Panikov and Dedysh, 2000; Friborg et al., 2003; Glagolev et al., 2011). Indeed, the paucity of in situ records limited our ability to evaluate LPX-BERN’s relatively low sensitivity to water table depth. Year-round observations would also be helpful, as winter emissions are sparsely sampled (Rinne et al., 2007; Kim et al., 2007; Panikov and Dedysh, 2000) and inversions disagree as to the magnitude of winter emissions (Fig. 8). The recent implementation of tower networks in the WSL (Sasakawa et al., 2010; Winderlich et al., 2010) show some promise in this regard, as their observations are both multiyear and year-round. More comprehensive observations of emissions from non-wetland methane sources such as seeps, pipe leaks, and lakes, most of which have so far not been accounted for in inversions (although pipe leaks are now being considered; Berchet et al., 2014), would be beneficial in increasing the accuracy of inversions.
5 Conclusions

We compared CH$_4$ emissions from 21 large-scale wetland models, including the models from the WETCHIMP project, to 5 inversions and several observational data sets of CH$_4$ emissions, surface water area, and total CH$_4$-producing area over the West Siberian Lowland (WSL) over the period 1993–2004. Despite the large scatter of individual estimates, mean estimates of annual total emissions over the WSL from forward models (5.34 ± 0.54 Tg CH$_4$ yr$^{-1}$), inversions (6.06 ± 1.22 Tg CH$_4$ yr$^{-1}$), and observations (3.91 ± 1.29 Tg CH$_4$ yr$^{-1}$) largely agreed. However, it was clear that reliance on satellite surface water products alone to delineate wetlands caused substantial biases in long-term mean CH$_4$ emissions over the region. Models and inversions largely agreed on the timing of the seasonal cycle of emissions over the WSL, but some outliers in the timing of peaks in the simulated inundated area indicated potential inaccuracies in simulating the timing of snowmelt and drainage rates. Models and inversions also displayed a wide range of interannual variability: the CV of the Boussquet2011 reference inversion was more than twice the CVs of all but one model, while the CV of the Bloom2010 inversion was essentially 0. Summer CH$_4$ emissions from the Boussquet2011 inversions exhibited only weak correlations with summer air temperature or inundation. Models that accounted for soil thermal physics and realistic methane–soil moisture relationships similarly tended to have low to moderate correlations with both inundation and air temperature, due in part to the competing influences of temperature and moisture, and in part to the insulating effects of snow and peat soils. In contrast, models lacking these formulations tended to be either inundation- or temperature-dominated (either inundation or air temperature, due in part to the insulating effects of snow and peat soils at high latitudes). In contrast, due to (a) misclassifying large areas of high-latitude peatlands that can emit methane when the water table is below the surface; (b) often including permanent water bodies, whose carbon cycling dynamics can be substantially different from those of wetlands; and (c) difficulties in detecting inundation under forest canopies. To improve the accuracy of global wetland map products may require combining information from satellite products and canonical maps.

Models should implement realistic soil thermal physics and snow schemes and account for the presence of peat soils at high latitudes.

- Models should implement realistic soil thermal physics and snow schemes and account for the presence of peat soils at high latitudes.

- Multiyear and multidecade observational and inversion products are crucial for assessing whether model simulations capture the correct sensitivities of wetland CH$_4$ emissions to environmental drivers.

The Supplement related to this article is available online at doi:10.5194/bg-12-3321-2015-supplement.

Author contributions. T. J. Bohn and J. R. Melton jointly conceived and designed this study with input from J. O. Kaplan. J. R. Melton provided the results from the original WETCHIMP models. T. J. Bohn, A. Ito, T. Kleinen, R. Spahni, B. D. Stocker, B. Zhang, and X. Zhu provided results for the other non-WETCHIMP models. R. Schroeder and K. C. McDonald provided the SWAMPS data set. M. V. Glagolev provided the Glagolev et al. (2011) data set and information on fossil methane seeps. S. Maksyutov provided the Kim et al. (2011) inversion and the Peregon et al. (2008) wetland map. V. Brovkin provided the Winderlich (2012) inversion. T. J. Bohn processed and reformatted results of all models, observations, and inversions and analyzed the results. T. J. Bohn and J. R. Melton collaborated on all figures. T. J. Bohn prepared the manuscript with contributions from all coauthors.

Acknowledgements. The WETCHIMP project received support for collaboration from the COST Action TERRABITES (ES0804). We thank Dennis P. Lettenmaier at the University of Washington and NASA grant NNX11AR16G from the Northern Eurasian Earth Science Partnership Initiative (NEEESPI) for the use of computational resources. T. J. Bohn was supported by grant 1216037 from the US National Science Foundation (NSF) Science, Engineering and Education for Sustainability (SEES) Post-Doctoral Fellowship program. J. R. Melton was supported by a National Science and Engineering Research Council of Canada (NSERC) visiting Post-Doctoral Fellowship. T. Kleinen was supported by the German Ministry of Education and Research (CarboPerm-Project, BMBF Grant No. 03G0836C). R. Spahni was supported by the Swiss National Science Foundation and by the European Commission through the FP7 project Past4Future (grant No. 243908). B. D. Stocker was supported by ERC FP7 project EMBRACE (grant No. 282672). H. Tian and B. Zhang were supported by US National Science Foundation (EaSM: AGS-1243220). Q. Zhuang and X. Zhu were supported by the US Department of Energy with project No. DE-SC0007007. M. V. Glagolev was supported by a grant in accordance with the Tomsk State University Competitiveness Improvement Program. S. Maksyutov was supported by Grant A-1202 of Environment Research and Technology Development Fund (ERTDF), Ministry of Environment Japan. S. N. Denisov and A. V. Eliseev were supported by the Russian President grant NSH-3894.2014.4 and by the Russian Foundation for Basic Research grant 15-05-02457. A. Gallego-Sala was supported by a Natural
References


Amante, C. and Eakins, B. W.: ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis, NOAA Technical Memorandum NOS NGDC-24, National Geophysical Data Center, NOAA, doi:10.7289/V5C8276M (last access: April 2013), 2009.


Efremova, T. T., Sekretenko, O. P., Avrova, A. F., and Efremov, S. P.: Spatial structure of acid properties of litter in the suc-


Hydro1K: HYDRO1K: Website, available at: https://ita.cr.usgs.gov/HYDRO1K (last access: 1 April 2013), 2013.


Kim, Y., Ueyama, M., Nakagawa, F., Tsunogai, U., Harazono, Y., and Tanaka, N.: Assessment of winter fluxes of CO2 and CH4 in boreal forest soils of central Alaska estimated by the profile method and the chamber method: a diagnosis of methane emis-


Riley, W. J., Subin, Z. M., Lawrence, D. M., Swenson, S. C., Torn, M. S., Meng, L., Mahowald, N. M., and Hess, P.: Barriers to predicting changes in global terrestrial methane fluxes: analyses using CLM4Me, and methane biogeochemistry model integrated in


tions, Montreal, Quebec, 413–427, 1988.

Tarnocai, C., Canadell, J. G., Schuur, E. A. G., Kuhry, P., Mazhi-

Tian, H., Xu, X., Liu, M., Ren, W., Zhang, C., and Lu, C.: Spa-


Trusilova, K., Rödenbeck, C., Gerbig, C., and Heimann, M.: Tech-
nical Note: A new coupled system for global-to-regional down-

Ulmishek, G. F.: Petroleum geology and resources of the West Siberian Basin, Russia, US Department of the Interior, US Ge-

Umezawa, T., Machida, T., Aoki, S., and Nakazawa, T.: Contri-

sion model based on temperate and arctic wetland sites, Biogeo-

p529/iov/cruncep/, last access: 1 September 2011.

Walter, B. P. and Heimann, M.: A process-based, climate-sensitive model to derive methane emissions from natural wetlands: Ap-


Wania, R., Ross, I., and Prentice, I. C.: Integrating peat-

Wania, R., Ross, I., and Prentice, I. C.: Implementation and evalua-


Winderlich, J.: Setup of a CO2 and CH4 measurement system in Central Siberia and modeling of its results, Technical Reports 26, Max-Planck-Institut für Biogeochemie, Jena, p. 120, 2012.


Zhu, X., Zhuang, Q., Lu, X., and Song, L.: Spatial scale-dependent land-atmosphere methane exchanges in the northern high lat-

ry: A retrospective analysis with a process-based biogeo-