Evaluation of wetland methane emissions across North America using atmospheric data and inverse modeling

S. M. Miller[1,*]**, R. Commane**[2]**, J.R. Melton**[3]**, A. E. Andrews**[4]**, J. Benmergui**[2]**, E.J. Dlugokencky**[4]**, G. Janssens-Maenhout**[5]**, A.M. Michalak**[6]**, C. Sweeney**[7]**, and D.E.J. Worthy**[8]

[1]Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA, USA
[2]School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA
[3]Climate Processes Section, Environment Canada, Victoria, Canada
[4]Global Monitoring Division, Earth System Research Laboratory, National Oceanic and Atmospheric Administration, Boulder, CO, USA
[5]Institute for Environment and Sustainability, European Commission Joint Research Centre, Ispra, Italy
[6]Department of Global Ecology, Carnegie Institution for Science, Stanford, CA, USA
[7]Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, CO, USA
[8]Climate Research Division, Environment and Climate Change Canada, Toronto, Canada
[*]now at: Department of Global Ecology, Carnegie Institution for Science, Stanford, CA, USA

Correspondence to: S. M. Miller (scot.m.miller@gmail.com)

Abstract

Existing estimates of methane (CH_4) fluxes from North American wetlands vary widely in both magnitude and distribution. In light of these differences, this study uses atmospheric CH_4 observations from the US and Canada to analyze seven different bottom-up, wetland CH_4 estimates reported in a recent model comparison project. We first use synthetic data to explore whether wetland CH_4 fluxes are detectable at atmospheric observation sites. We find that the observation network can detect aggregate wetland fluxes from both eastern and western Canada but generally not from the US. Based upon these results, we then use real data and inverse modeling results to analyze the magnitude, seasonality, and spatial distribution of each model estimate. The magnitude of Canadian fluxes in many models is larger than indicated by atmospheric observations. Many models predict a seasonality that is narrower than implied by inverse modeling results, possibly indicating an over-sensitivity to air or soil temperatures. The LPJ-Bern and SDGVM models have a geographic distribution that is most consistent with atmospheric observations, depending upon the region and season. These models utilize land cover maps or dynamic modeling to estimate wetland coverage while most other models rely primarily on remote sensing inundation data.

1 Introduction

CH_4 fluxes from wetlands play a critical role in global climate change. CH_4 is the second-most important long-lived greenhouse gas, and the radiative forcing of the current atmospheric burden is approximately 26 % of carbon dioxide (Butler, 2014). Wetlands are possibly the largest single
5 source of this gas to the atmosphere and account for roughly 30 % of global emissions (Kirschke et al., 2013).

Despite the important role of wetland CH_4 fluxes in climate change, existing estimates of this source differ on the magnitude, seasonality, and spatial distribution of fluxes, from regional to global scales. In fact, a recent global model comparison project named WETCHIMP (Wetland
10 and Wetland CH_4 Inter-comparison of Models Project) found large differences among existing CH_4 wetland models (Fig. 1, Melton et al., 2013; Wania et al., 2013). For example, existing estimates of maximum global wetland coverage differ by about a factor of 6 – from 4.1×10^6 to $26.9 \times 10^6 \, \mathrm{km}^2$. Furthermore, estimates of global natural wetland fluxes range from 92–$264 \, \mathrm{Tg} \, CH_4 \, \mathrm{yr}^{-1}$. The relative magnitude of these uncertainties increases at sub-global spatial
15 scales. As a case in point, CH_4 estimates for Canada's Hudson Bay Lowlands (HBL) range from 0.2 to $11.3 \, \mathrm{Tg} \, CH_4 \, \mathrm{yr}^{-1}$. These disagreements in current CH_4 estimates do not bode well for scientists' abilities to accurately predict future changes in wetland fluxes due to climate change (Melton et al., 2013).

A number of studies have used chamber measurements of CH_4 to parameterize or evaluate
20 biogeochemical CH_4 models (e.g., Livingston and Hutchinson, 2009). These measurements usually encompass fluxes from a relatively small area, and fluxes can often vary greatly with landscape heterogeneity at these spatial scales (Waddington and Roulet, 1996; Hendriks et al., 2010). CH_4 data collected in the atmosphere sees the cumulative effect of CH_4 fluxes across a broader region (e.g., Winderlich et al., 2010; Pickett-Heaps et al., 2011; Miller et al., 2014).
25 Hence, atmospheric data can provide a unique tool for evaluating existing CH_4 flux estimates across different countries or continents.

The present study compares the WETCHIMP CH_4 flux estimates against atmospheric CH_4 data and inverse modeling results from 2007–2008 through two sets of analyses. First, we con-

struct a set of synthetic data experiments to understand whether the atmospheric CH_4 observation network can detect CH_4 fluxes from wetlands. We also explore the factors that might prevent the network from detecting wetland fluxes. To answer these questions, we utilize a model selection procedure based upon the Bayesian Information Criterion (BIC) (Sect. 2.2 Shiga et al.,
5   2014; Fang et al., 2014; Fang and Michalak, 2015). This procedure determines whether wetland fluxes from different regions and seasons are necessary to describe variability in synthetic atmospheric CH_4 observations. Based on these synthetic experiments, we conduct a second set of analyses using real atmospheric data and inverse modeling results. We use this data to analyze the magnitude, seasonal cycle, and spatial distribution of each WETCHIMP CH_4 estimate.
10 We investigate these questions over the US and Canada using CH_4 data collected from towers and regular aircraft flights operated by NOAA and its partners and from towers operated by Environment Canada.

2 Methods

This section first describes the atmospheric CH_4 data and the atmospheric model that allows
15 direct comparison between the data and various flux estimates. Subsequent sections describe both the synthetic and real data experiments outlined in the introduction (Sect. 1).

2.1 Data and atmospheric model

The present study utilizes atmospheric CH_4 observations from aircraft and tower platforms across the US and Canada, a total of 14 703 observations from 2007–2008. These observation
20 sites include four towers operated by Environment Canada and 10 towers in the US operated by NOAA and its partners. Observations at the NOAA towers consist of daily flasks (occasionally weekly), and observations at the Environment Canada sites are continuous measurements. As in Miller et al. (2014), we use afternoon averages of this continuous data. In addition to these towers, we utilize observations from 17 regular NOAA aircraft monitoring locations in the US
25 and Canada (Fig. 2). We incorporate aircraft data up to 2500m altitude as was done in Miller

et al. (2013). Observations above that height are usually representative of the free troposphere with limited sensitivity to surface fluxes. These observations and the associated model runs (described below) are the same as those in Miller et al. (2013) and Miller et al. (2014).

We then employ an atmospheric transport model to relate CH_4 fluxes at the Earth's surface
5 to atmospheric concentrations at the observation sites. The modeling approach here combines the Weather Research and Forecasting (WRF) meteorological model and a particle-following model known as STILT, the Stochastic Time-Inverted Lagrangian Transport model (e.g., Lin et al., 2003; Nehrkorn et al., 2010; Hegarty et al., 2013). WRF-STILT generates a set of footprints; these footprints quantitatively estimate the sensitivity of each observation to fluxes at
10 each surface location (with units of ppb per unit surface flux). We multiply the footprints by a flux model and add this product to an estimate of the "background" concentration – the CH_4 concentration of air entering the North American regional domain. We estimate this background concentration using CH_4 observations collected near or over the Pacific Ocean and in the high Arctic, a setup described in detail by Miller et al. (2013) and Miller et al. (2014). The resulting
15 modeled concentrations can be compared directly against atmospheric CH_4 observations. The observations, WRF-STILT runs, background concentrations, and uncertainties in the modeling framework are described in greater detail in the Supplement, Miller et al. (2013), and Miller et al. (2014).

Using this setup, we can evaluate predicted CH_4 concentrations using the WETCHIMP flux
20 estimates (Fig. 1) against observed atmospheric concentrations. The WETCHIMP project was designed to compare simulated wetland distributions and modeled CH_4 fluxes at multi-year, continental scales (Melton et al., 2013; Wania et al., 2013). The project entailed several sets of model runs, but Melton et al. (2013) primarily reported on one set of runs – runs for 1901–2009 that used the same observed climate and CO_2 concentration datasets to force all models.
25 Each CH_4 model utilized its own parameterization for wetland area and distribution. We use the outputs from this set of model runs in the present study. Of the WETCHIMP models, seven provide a flux estimate on a suitable time step for boreal North America and six provide an estimate for temperate North America. These models include CLM4Me (Riley et al., 2011), DLEM (Tian et al., 2010), LPJ-Bern (Spahni et al., 2011), LPJ-WHyMe (Wania et al., 2010),

5

LPJ-WSL (Hodson et al., 2011), ORCHIDEE (Ringeval et al., 2010), and SDGVM (Singarayer et al., 2011). All flux model outputs used from the WETCHIMP study have a temporal resolution of one month. These models are described in Melton et al. (2013); Wania et al. (2013), and the Supplement.

2.2 Synthetic data experiments

We assess the ability of the CH_4 observation network to detect wetland fluxes and use a model selection framework adapted from the BIC. A model selection framework can sort through a large number of potential explanatory variables and will choose the smallest set of variables that best describe the dataset of interest (e.g., Ramsey and Schafer, 2012). In the current setup, we generate synthetic atmospheric CH_4 observations. The model selection framework then indicates whether a wetland model and/or an anthropogenic emissions inventory are necessary to describe variability in these observations. In this way, model selection can indicate the sensitivity of the observation network to wetland CH_4 fluxes.

We use a form of the BIC that has been adapted for use within a geostatistical inverse modeling framework. This setup has previously been used to select either bottom-up models or environmental drivers of CO_2 and CH_4 fluxes (e.g., Mueller et al., 2010; Yadav et al., 2010; Gourdji et al., 2012; Miller et al., 2013, 2014; Shiga et al., 2014; Fang et al., 2014; Fang and Michalak, 2015). The implementation here mirrors that of Fang et al. (2014), Shiga et al. (2014), and Fang and Michalak (2015):

$$\text{BIC} = \underbrace{\ln|\mathbf{\Psi}| + (\mathbf{z} - \mathbf{HX}\boldsymbol{\beta})^{\text{T}}\mathbf{\Psi}^{-1}(\mathbf{z} - \mathbf{HX}\boldsymbol{\beta})}_{\text{negative log-likelihood}} + \underbrace{p\ln(n)}_{\text{penalty term}} \tag{1}$$

The first two terms in Eq. (1) are the negative log-likelihood, a measure of how well the model fits the data. The last term penalizes a particular model based upon the number of explanatory variables (p). The best combination or candidate model has the lowest BIC score.

The variable $\mathbf{z}$ ($n \times 1$) represents the observations minus background concentrations, $\mathbf{H}$ ($n \times m$) the footprints, and $\mathbf{\Psi}$ ($n \times n$) a covariance matrix derived from an atmospheric inversion framework. The variable m refers to the total number of flux or emissions grid boxes

6

in both space and time. These variables are based upon two existing inverse modeling studies by Miller et al. (2013) and Miller et al. (2014) (refer to the Supplement). The matrix $\mathbf{X}$ ($m \times p$) contains p explanatory variables. In the current setup, $\mathbf{X}$ can include a wetland flux estimate and/or individual emissions sources from an anthropogenic inventory. $\boldsymbol{\beta}$ ($p \times 1$) is a set of co-efficients that scale the variables in $\mathbf{X}$. We set these coefficients to one in the the synthetic data experiments. As a result, the model selection framework cannot reproduce wetland fluxes by simply upscaling anthropogenic emissions sources that might have a similar distribution to wetlands.

The first experiments described here use synthetic atmospheric CH_4 data. We generate the synthetic data using one of the WETCHIMP models and the anthropogenic emissions estimates from Miller et al. (2013) and Miller et al. (2014). For consistency among the synthetic datasets, we scale the annual HBL CH_4 budget in each WETCHIMP model to match the overall magnitude estimated by several top-down studies (Pickett-Heaps et al., 2011; Miller et al., 2014; Wecht et al., 2014). We then multiply these fluxes by the footprints ($\mathbf{H}$) and add error that is randomly generated from the covariance matrix ($\boldsymbol{\Psi}$). This covariance matrix represents errors in atmospheric transport and in the measurements – collectively referred to as model-data mismatch. This matrix also represents uncertainties in the prior flux estimate. In a geostatistical inverse model, this prior flux model is given by $\mathbf{X}\boldsymbol{\beta}$ (refer to the supplement for more detail).

We divide the WETCHIMP wetland fluxes into four regions (Fig. 2) and four seasons (DJF, MAM, JJA, and SON). The model selection framework then chooses variables that are necessary to reproduce the synthetic data, variables that include EDGAR and the 16 wetland flux maps. The penalty term in Eq. 1 increases as we add wetland maps or add EDGAR to the $\mathbf{X}$ matrix. Each variable added to $\mathbf{X}$ will increase the penalty term by $\ln(n)$; an additional variable must improve the log-likelihood by more than this penalty term to be chosen by model selection.

We then run this framework 1000 times, generating new synthetic data each time, and calculate the percentage of all trials in which the model selection chooses a wetland model. The 1000 repeats are needed due to the random or stochastic nature of the synthetic data experiment; the results of the model selection can vary slightly, depending on the particular random errors that

we generate based upon the covariance matrix ($\mathbf{\Psi}$). This procedure ensures that the model selection results are not the output of a single realization. We then report on how frequently each of the 16 wetland flux maps is chosen. If a wetland flux map is chosen with high frequency, then a wetland flux map is necessary to describe variability in the synthetic CH_4 observations, and

5 the synthetic observation network can detect aggregate wetland CH_4 fluxes from the given region and season. This setup mirrors that of Shiga et al. (2014), who employed a model selection framework to explore the detectability of anthropogenic CO_2 emissions.

We also explore why the synthetic CH_4 observations may not be able to detect wetland fluxes. We run a series of case studies and in each case remove a different confounding factor that might

10 prevent the network from detecting wetland CH_4 fluxes. In one case, we remove anthropogenic emissions. In subsequent cases, we remove model-data mismatch errors and/or prior flux errors. In each case, we re-run the model selection experiment and examine whether the results improve when each of these confounding factors is removed.

2.3 Real data experiments

15 This paper subsequently compares the spatial distribution, magnitude, and seasonality of each WETCHIMP estimate against real atmospheric CH_4 observations, using the synthetic experiments to guide the analysis.

We first explore the spatial distribution of the WETCHIMP flux estimates. We modify the model selection setup in Sect. 2.2 to focus on the spatial distribution of each estimate using

20 a procedure developed by Fang et al. (2014) and Fang and Michalak (2015). Instead of fixing the coefficients ($\boldsymbol{\beta}$) to one, we instead estimate the coefficients using real atmospheric CH_4 observations. We also include an intercept term that can vary by month; the intercept for each month is represented by a vector of ones in the matrix $\mathbf{X}$, and this intercept is included as part of each candidate model for $\mathbf{X}$. We then run model selection using real observations. As

25 a result of this setup, a wetland model is not necessary to reproduce either the magnitude or seasonality of the atmospheric CH_4 data; the model selection framework can simply scale the intercept term or scale EDGAR to reproduce the magnitude or seasonality of the observations. The spatial distribution of wetland fluxes, however, can only come from a wetland model. The

8

model selection procedure will only choose a wetland model if the spatial distribution of that model describes sufficient, additional variability in the observations (e.g., Fang et al., 2014).

Model selection can therefore indicate which WETCHIMP models have the best spatial distribution relative to the atmospheric observations; any WETCHIMP model chosen by model selection has a spatial distribution that improves model-data fit, and the model improves that fit more than the penalty term in Eq. 1. A negative result does not necessarily indicate that a WETCHIMP model has a poor spatial distribution. In that case, the observations may not be very sensitive to the spatial distribution of fluxes for the given region or given season. Similarly, the spatial distribution in a WETCHIMP model may improve model-data fit but not by more than the penalty term in Eq. 1. By contrast, a positive result indicates that a WETCHIMP model likely has a particularly good spatial distribution. As in Sect. 2.2, we divide the wetland fluxes into four sub-continental regions and four seasons. The Supplement describes this setup in greater detail.

We then analyze the magnitude and seasonality of the WETCHIMP fluxes using a number of model-data time series. We model CH_4 concentrations at a number of US and Canadian observation sites using the WRF-STILT model, the WETCHIMP estimates, and the EDGAR v4.2FT2010 emissions inventory (Olivier and Janssens-Maenhout, 2012; European Commission, Joint Research Centre (JRC)/Netherlands Environmental Assessment Agency (PBL), 2013). We average the observations and model output at the monthly scale and then compare the magnitude of these model estimates for each month against the averaged observations.

We further compare the seasonality of existing bottom-up models against the seasonality of a recent inverse modeling estimate by Miller et al. (2014). We plot the monthly budgets for both the bottom-up models and the inversion estimate, and we plot the monthly CH_4 budget as a fraction of the annual total.

Note that inter-annual variability in existing CH_4 flux models is small relative to the differences among these models; as a result, conclusions from the 2 year study period (2007–2008) likely hold for other years. For example, the inter-annual variability in the total US/Canadian budget is ± 7.3–9.7% (standard deviation), depending upon the model in question (Note that

9

LPJ-Bern has even larger inter-annual variation due to an issue with model spin-up described in Wania et al. (2013)).

3 Results and discussion: synthetic experiments

The synthetic experiments presented here explore the limits of existing atmospheric data for constraining wetland fluxes. If atmospheric observations are to constrain wetland CH_4 fluxes, those observations must be able to detect wetland CH_4 fluxes above errors in the transport model and above other emissions sources such as fossil fuels and agriculture.

The four columns in Fig. 3a display the results from an individual season in each of four geographic regions. In this experiment, the synthetic CH_4 observations can detect aggregate wetland CH_4 fluxes from Eastern Canadian wetlands in greater than 75 % of all trials for the summer and fall seasons. In the eastern US, the model selection framework chooses a wetland model in 25–50 % of all trials in multiple different seasons. By contrast, the synthetic CH_4 data are least sensitive to wetland fluxes in the western US, and the model selection framework chooses wetland fluxes from that region in fewer than 25 % of all trials irrespective of the season. This result may be due, in part, to the scant wetlands and sparse atmospheric observations in much of the west.

The results also vary by season. Of any region, the atmospheric CH_4 network is best able to constrain fluxes across multiple seasons in eastern Canada. The largest wetland fluxes in the WETCHIMP models are in Ontario and Quebec (Fig. 1). It is therefore unsurprising that the network is best able to detect wetland fluxes in that region, even though there are relatively few observation sites in the area. In other regions, the atmospheric CH_4 network is less sensitive to wetlands during the winter, fall, and spring shoulder seasons.

We run several additional model selection experiments to explore why the synthetic observations may not always be able to detect wetland CH_4 fluxes (Fig. 3b-e). We remove anthropogenic emissions from the synthetic dataset for the experiment in Fig. 3b. We remove all model data mismatch errors in Fig. 3c; model-data mismatch encompasses errors in atmospheric transport and in the measurements. Subsequently, we remove all errors due to the prior flux estimate

10

in Fig. 3d. In Fig. 3e, we remove both types of errors. In each case, we re-run the model selection experiment to see if the sensitivity of the atmospheric CH_4 network to wetland fluxes improves.

Anthropogenic emissions have only a modest effect on the results in specific regions and seasons. In case (b) without anthropogenic emissions, the results improve by $\sim$25–50% in the fall and spring shoulder seasons for several geographic regions.

By contrast, the model-data mismatch and prior flux errors have a much larger effect on the model selection results. The results improve incrementally across many regions and seasons when we remove model-data mismatch errors in case (c). The results improve across the spring, summer, and fall seasons and improve across all four geographic regions. However, the magnitude of this improvement is never more than 25%. Model-data mismatch errors are likely dominated by errors in modeled atmospheric transport. These results imply that transport errors play an incremental yet pervasive role in the utility of the atmospheric observations.

The prior flux errors have the largest effect on the results, particularly during the warmest seasons. In case (d), the results show great improvement during fall, spring, and summer and show little improvement during winter or in the western US. In the setup here, the prior flux uncertainties scale with the seasonal magnitude of the fluxes. When we remove the prior flux errors, the results concomitantly show the greatest improvement in seasons that have larger overall CH_4 fluxes. These results indicate that the prior estimate greatly impacts the utility of the atmospheric CH_4 observations. A geostatistical inverse model can leverage any combination of land surface maps, meteorological maps, and/or anthropogenic inventory estimates in the inversion prior. These maps or estimates are incorporated into the X matrix in Eq. 1. If accurate maps or estimates are not available, then the prior flux errors will be large, and the model selection framework will be less likely to choose any particular variable. If these maps or estimates have high explanatory power, then the prior flux errors will be small, and the model selection framework will be more likely to choose any one variable. As a result, the detectability of wetland CH_4 fluxes partly depends on the availability of land surface or meteorological data that matches those fluxes. The atmospheric network can better differentiate wetland CH_4 fluxes from other CH_4 sources when accurate prior information can guide that identification.

11

Case (e) (no model-data mismatch errors and no errors in the prior flux estimate) shows large, ubiquitous improvements; the model selection chooses a wetland model 100% of the time in almost all regions and seasons. The results for Eastern Canada during winter are the exception. In winter, the wetland model cannot always explain enough variability in the synthetic observations to overcome the BIC penalty term in Eq. 1.

The density of the atmospheric CH_4 network may also play a role in these results. Wetlands in the Eastern US are sparse relative to eastern Canada, but the higher density of observations in the Eastern US may contribute to a moderate success rate $(25-50\%)$ for that region. Recent and planned network expansions in the eastern US and in Canada could play a key role in future efforts to constrain wetland fluxes across these regions.

Overall, the synthetic experiment results indicate that the observation network cannot detect wetland fluxes from the US (i.e., model selection has a success rate <50%). Across Canada, the results are more promising (i.e., near 100% success rate in some regions/seasons), despite the relative sparsity of the observation network there.

4 Results and discussion: comparisons with atmospheric data and inverse modeling results

4.1 Spatial distribution of the fluxes

We compare the spatial distribution of the WETCHIMP flux estimates against CH_4 data from the atmospheric observation network. To this end, we use a version of the model selection framework that chooses wetland models based upon their spatial distribution (Fang et al., 2014; Fang and Michalak, 2015). WETCHIMP models that are chosen by the framework have a spatial distribution that is more consistent with atmospheric observations relative to those that are not selected.

The results of this model selection analysis are displayed in Table 1. This table lists the regions and seasons that had a success rate $> 50\%$ in the synthetic data experiment; the atmospheric CH_4 network is most sensitive to wetland CH_4 fluxes in those regions and seasons.

Two of the WETCHIMP models were chosen by the model selection framework – LPJ-Bern (in eastern Canada) and SDGVM (in eastern and western Canada). The spatial distribution of these models improve the model-data fit more than the penalty term in Eq. 1.

The LPJ-Bern and SDGVM models have several unique spatial characteristics that could explain these results. Over eastern Canada, LPJ-Bern and SDGVM concentrate the large fluxes in the HBL. Other models, by contrast, often distribute the fluxes more broadly across Ontario and Quebec or put the largest fluxes in Ontario outside of the HBL. In western Canada, SDGVM distributes fluxes across northern, boreal Saskatchewan and Alberta.

The LPJ-Bern and SDGVM models share another common characteristic: both model wetland area independently instead of relying solely on remote sensing inundation datasets. LPJ-WSL, ORCHIDEE, DLEM, and CLM4Me use remote sensing inundation datasets like GIEMS (Global Inundation Extent from Multi-Satellites, Prigent et al., 2007) to construct a wetland map. Other models, like LPJ-Bern and LPJ-WHyMe also use land cover maps and/or land surveys to estimate wetland (or at least CH_4-producing) area. SDGVM estimates this area dynamically as a function of soil moisture (Melton et al., 2013; Wania et al., 2013). Wetland maps generated using these different approaches show substantial differences. Remote sensing datasets estimate relatively high levels of inundation in regions of Canada that are not forested or have many small lakes (see further discussion in Melton et al., 2013; Bohn et al., 2015). By contrast, modeling approaches that dynamically generate wetland area or use land cover maps assign more wetlands over regions with high water tables but little surface water as seen by remote sensing based inundation datasets. As a result of these differences, models like LPJ-Bern assign more wetlands and CH_4 fluxes in the HBL relative to other regions of eastern Canada.

Of note, LPJ-Bern and LPJ-WhyMe have many structural model similarities but predict relatively different spatial distributions of CH_4 fluxes. The latter estimates fluxes that are more broadly distributed across Quebec and Labrador. LPJ-WHyMe only simulates fluxes from high latitude peatlands and uses an estimated peatland distribution from Tarnocai et al. (2009); this distribution extends across Quebec and Labrador. LPJ-Bern, by contrast, includes fluxes from non-peatland regions and applies a smaller scaling factor to peatland fluxes relative to LPJ-

13

WHyMe Wania et al. (2013). As a result, the fluxes in LPJ-Bern have a spatial distribution that is different from the peatland map and also different from LPJ-WHyMe.

4.2 Flux magnitude

We next compare the magnitude of predicted concentrations using the WETCHIMP models against atmospheric observations at individual locations. Unlike previous sections that utilized model selection, this section employs several model-data time series, displayed in Fig. 4. This model estimate consists of several components: the background (in green) is the estimated background concentration of CH_4 in clean air before entering the model domain as in Miller et al. (2013, 2014). The estimated contribution of anthropogenic emissions from EDGAR v4.2FT2010 is added to this background (in red). The contribution of wetland fluxes from the WETCHIMP models is then added to the previous inputs, and the sum of all components (blue lines) can be compared directly against measured concentrations.

The various WETCHIMP flux estimates produce very different modeled concentrations at the atmospheric observation sites (Fig. 4). Overall, modeled concentrations with the WETCHIMP fluxes usually exceed the CH_4 measurements during summer. At Chibougamau, Fraserdale, and Park Falls in early summer, all seven WETCHIMP models predict CH_4 concentrations that equal or exceed the observations. The ORCHIDEE, LPJ-WHyMe, and LPJ-Bern models always exceed the measurements during summer while DLEM and SDGVM better match the observations at these sites. Notably, a number of previous studies report that the EDGAR inventory may underestimate US anthropogenic CH_4 emissions (e.g., Kort et al., 2008; Miller et al., 2013; Wecht et al., 2014; Turner et al., 2015). If EDGAR underestimate emissions, then the WETCHIMP models would be an even larger overestimate relative to the atmospheric data.

Many models appear to overestimate the magnitude of fluxes across boreal North America, but this result does not necessarily imply that these models have underestimated fluxes elsewhere. CH_4 models that estimate the largest fluxes across boreal North America do not always compensate with smaller fluxes in other regions of the globe. For example, the ORCHIDEE model not only estimates large fluxes over North America but also estimates higher fluxes over the tropics than any other model (Melton et al., 2013).

14

4.3 Seasonal cycle

Bottom-up CH_4 flux estimates show variable features when compared to atmospheric observations, and the seasonal cycle of these estimates is no exception. Figure 5 compares the seasonal cycle of the existing estimates over Canada's HBL. Eastern Canada is one of the largest wetland
5 regions in North America (Fig. 1), and nearby atmospheric observation sites see a much larger CH_4 enhancement from wetlands relative to other regions (Fig. 4 and S4).

In this region, the bottom-up estimates diverge on the seasonal cycle of fluxes. Most estimates predict peak fluxes in July or August, though two variations of the LPJ model predict seasonal peaks in September and October – LPJ-WHyMe and LPJ-Bern, respectively. LPJ-WHyMe is
10 a module inside of LPJ-Bern, a possible explanation for the similar seasonal cycle in these two models. Differences among models are also notable during the fall and spring seasons. For example, fluxes in June account for anywhere between 6 and 21 % of the annual CH_4 budget, depending upon the model. Fluxes in October account for between 1 and 23 % of the annual budget (Fig. 5b).

15 Figure 5 also displays the seasonality of an inverse modeling estimate from Miller et al. (2014) for comparison. That estimate incorporates observations from Chibougamau, Quebec, and Fraserdale, Ontario, atmospheric measurement sites that are strongly influenced by fluxes from the HBL. Differences between this inverse modeling estimate and the WETCHIMP models often exceed the 95 % confidence interval of the inverse model. The WETCHIMP estimates
20 are often comparable to Miller et al. (2014) in magnitude during fall and spring months but exceed the inverse modeling estimate in summer months (Fig. 5a). On whole, the WETCHIMP models have a narrower relative seasonal cycle than the inverse modeling estimate (Fig. 5b). That estimate assigns a greater portion of the annual budget to the fall and spring shoulder seasons.

25 Additional top-down studies exist for the HBL, but those studies use a seasonal cycle drawn from an existing bottom-up model and do not estimate the seasonal cycle independently from CH_4 observations (Pickett-Heaps et al., 2011; Wecht et al., 2014; Turner et al., 2015). By comparison, a recent inverse modeling study of the Western Siberian Lowlands found parallel results

15

for that region – existing models also predict a seasonal cycle that is narrower than the seasonality implied by atmospheric observations (Winderlich, 2012; Bohn et al., 2015).

Numerous possible explanations could underly differences in the seasonal cycle of CH_4 fluxes. For example, the temperature threshold for CH_4 production may be too high in some models. Relative to summer months, the bottom-up models predict small fluxes during fall/spring months when air temperatures are near freezing but soils are still unfrozen (Fig. S3). According to estimates from the North American Regional Reanalysis (NARR) (Mesinger et al., 2006), surface soils in the HBL (0 and 10 cm depth) begin to thaw in April and are largely unfrozen in May (Fig. S3). In the fall, surface soils (0 cm depth) begin to freeze in November, but deeper soils (10 and 40 cm) remain largely unfrozen until December. Compared to the bottom-up models, the inverse modeling estimate predicts a wider seasonal window, a result that would be consistent with dates of deep soil freeze and thaw.

5 Conclusions

A recent model comparison study revealed wide differences among several estimates of wetland CH_4 fluxes – differences at global to regional scales. In the first component of this study, we use a synthetic data experiment to understand whether the atmospheric observation network can detect wetland CH_4 fluxes. We find that the network can reliably identify aggregate wetland fluxes from both eastern and western Canada. The network can detect wetland fluxes from the eastern US in a smaller fraction of trials and rarely from the western US. This analysis also accounts for distracting signals in the atmosphere from anthropogenic sources or simulated atmospheric transport errors.

In a second component of the study, we analyze each bottom-up CH_4 model from the WETCHIMP study using real atmospheric data. We find that the LPJ-Bern and SDGVM models have spatial distributions that are most consistent with atmospheric observations, depending upon the region and season of interest. In addition, almost all models overestimate the magnitude of wetland CH_4 fluxes when compared against atmospheric data at individual observation sites. The model ensemble may also estimate a seasonal cycle for eastern Canada that is too narrow (i.e.,

places too much of the total annual flux in the summer relative to the fall and spring shoulder seasons).

The results of this paper suggest possible pathways to improve future top-down estimates of wetland CH_4 fluxes. The ability of the atmospheric observation network to detect wetland fluxes depends in large part upon the prior flux model. In a geostatistical inverse model, this model can incorporate land surface maps – wetland maps, estimates of land surface processes, and maps of anthropogenic emissions sources. This information plays a large role in whether atmospheric observations can detect wetland fluxes; the observations can more adeptly identify wetland fluxes when accurate land surface maps are available to guide that identification. By contrast, atmospheric transport and measurement errors (i.e., model-data mismatch errors) have a ubiquitous but smaller effect on the utility of atmospheric CH_4 observations.

The results presented here also hold a number of suggestions for future bottom-up modeling efforts:

1. Spatial distribution: Bottom-up estimates that use surface water inundation as the only proxy for wetland area do not perform as well relative to atmospheric observations. Bottom-up models that use satellite inundation data should incorporate additional tools like wetland mapping or dynamic modeling to capture wetlands covered by vegetation.

2. Magnitude: Existing top-down studies that use a diverse array of in situ and satellite CH_4 observations show good agreement on the magnitude of CH_4 fluxes from the Hudson Bay Lowlands (HBL) region (e.g., Pickett-Heaps et al., 2011; Miller et al., 2014; Wecht et al., 2014; Turner et al., 2015). These studies could be used to calibrate the magnitude of future bottom-up estimates, at least over the HBL where CH_4 observations provide a strong constraint on wetland fluxes.

3. Seasonal cycle: Bottom-up models do not show consensus on the seasonal cycle of wetland fluxes across Canada. Few top-down studies estimate the seasonal cycle independently using atmospheric observations. Additional top-down studies would indicate the range of seasonal cycle estimates that are consistent with atmospheric observations, particularly studies that use a diverse set of atmospheric models and/or diverse observa-

17

tional datasets. These efforts could help reconcile differences in the seasonal cycle among bottom-up models and between bottom-up models and the few, existing top-down studies.

These steps will hopefully lead to better convergence among wetland CH_4 estimates for North America.

5 **The Supplement related to this article is available online at doi:10.5194/bgd-0-1-2016-supplement.**

References

Andrews, A. E., Kofler, J. D., Trudeau, M. E., Williams, J. C., Neff, D. H., Masarie, K. A., Chao, D. Y., Kitzis, D. R., Novelli, P. C., Zhao, C. L., Dlugokencky, E. J., Lang, P. M., Crotwell, M. J., Fischer, M. L., Parker, M. J., Lee, J. T., Baumann, D. D., Desai, A. R., Stanier, C. O., De Wekker, S. F. J.,
20 Wolfe, D. E., Munger, J. W., and Tans, P. P.: CO_2, CO, and CH_4 measurements from tall towers in the NOAA Earth System Research Laboratory's Global Greenhouse Gas Reference Network: instrumentation, uncertainty analysis, and recommendations for future high-accuracy greenhouse gas monitoring efforts, Atmos. Meas. Tech., 7, 647–687, doi:10.5194/amt-7-647-2014, 2014.

Bohn, T. J., Melton, J. R., Ito, A., Kleinen, T., Spahni, R., Stocker, B. D., Zhang, B., Zhu, X.,
25 Schroeder, R., Glagolev, M. V., Maksyutov, S., Brovkin, V., Chen, G., Denisov, S. N., Eliseev, A. V.,

18

Gallego-Sala, A., McDonald, K. C., Rawlins, M.A., Riley, W. J., Subin, Z. M., Tian, H., Zhuang, Q., and Kaplan, J. O.: WETCHIMP-WSL: intercomparison of wetland methane emissions models over West Siberia, Biogeosciences, 12, 3321–3349, doi:10.5194/bg-12-3321-2015, 2015.

Butler, J.: The NOAA annual greenhouse gas index (AGGI), http://www.esrl.noaa.gov/gmd/aggi/, 2014.

European Commission, Joint Research Centre (JRC)/Netherlands Environmental Assessment Agency (PBL): Emission Database for Global Atmospheric Research (EDGAR), release EDGARv4.2 FT2010, available at: http://edgar.jrc.ec.europa.eu (last access: 14 June 2015), 2013.

Fang, Y. and Michalak, A. M.: Atmospheric observations inform CO_2 flux responses to enviroclimatic drivers, Global Biogeochem. Cy., 29, GB005034, doi:10.1002/2014GB005034, 2015.

Fang, Y., Michalak, A. M., Shiga, Y. P., and Yadav, V.: Using atmospheric observations to evaluate the spatiotemporal variability of CO_2 fluxes simulated by terrestrial biospheric models, Biogeosciences, 11, 6985–6997, doi:10.5194/bg-11-6985-2014, 2014.

Gourdji, S. M., Mueller, K. L., Yadav, V., Huntzinger, D. N., Andrews, A. E., Trudeau, M., Petron, G., Nehrkorn, T., Eluszkiewicz, J., Henderson, J., Wen, D., Lin, J., Fischer, M., Sweeney, C., and Michalak, A. M.: North American CO_2 exchange: inter-comparison of modeled estimates with results from a fine-scale atmospheric inversion, Biogeosciences, 9, 457–475, doi:10.5194/bg-9-457-2012, 2012.

Hegarty, J., Draxler, R. R., Stein, A. F., Brioude, J., Mountain, M., Eluszkiewicz, J., Nehrkorn, T., Ngan, F., and Andrews, A.: Evaluation of Lagrangian particle dispersion models with measurements from controlled tracer releases, J. Appl. Meteorol. Clim., 52, 2623–2637, 2013.

Hendriks, D. M. D., van Huissteden, J., and Dolman, A. J.: Multi-technique assessment of spatial and temporal variability of methane fluxes in a peat meadow, Agr. Forest Meteorol., 150, 757–774, doi:10.1016/j.agrformet.2009.06.017, 2010.

Hodson, E. L., Poulter, B., Zimmermann, N. E., Prigent, C., and Kaplan, J. O.: The El Niño–Southern Oscillation and wetland methane interannual variability, Geophys. Res. Lett., 38, L08810, doi:10.1029/2011GL046861, 2011.

Kirschke, S., Bousquet, P., Ciais, P., Saunois, M., Canadell, J. G., Dlugokencky, E. J., Bergamaschi, P., Bergmann, D., Blake, D. R., Bruhwiler, L., Cameron-Smith, P., Castaldi, S., Chevallier, F., Feng, L., Fraser, A., Heimann, M., Hodson, E. L., Houweling, S., Josse, B., Fraser, P. J., Krummel, P. B., Lamarque, J.-F., Langenfelds, R. L., Le Quere, C., Naik, V., O'Doherty, S., Palmer, P. I., Pison, I., Plummer, D., Poulter, B., Prinn, R. G., Rigby, M., Ringeval, B., Santini, M., Schmidt, M., Shindell, D. T., Simpson, I. J., Spahni, R., Steele, L. P., Strode, S. A., Sudo, K., Szopa, S., van der Werf, G. R.,

Voulgarakis, A., van Weele, M., Weiss, R. F., Williams, J. E., and Zeng, G.: Three decades of global methane sources and sinks, Nat. Geosci., 6, 813–823, doi:10.1038/ngeo1955, 2013.

Kort, E. A., Eluszkiewicz, J., Stephens, B. B., Miller, J. B., Gerbig, C., Nehrkorn, T., Daube, B. C., Kaplan, J. O., Houweling, S., and Wofsy, S. C.: Emissions of CH_4 and N_2O over the United States and Canada based on a receptor-oriented modeling framework and COBRA-NA atmospheric observations, Geophys. Res. Lett., 35, L18808, doi:10.1029/2008gl034031, 2008.

Lin, J., Gerbig, C., Wofsy, S., Andrews, A., Daube, B., Davis, K., and Grainger, C.: A near-field tool for simulating the upstream influence of atmospheric observations: the Stochastic Time-Inverted Lagrangian Transport (STILT) model, J. Geophys. Res.-Atmos., 108, 4493, doi:10.1029/2002jd003161, 2003.

Livingston, G. and Hutchinson, G.: Enclosure-based measurement of trace gas exchange: applications and sources of error, chap. 2, in: Methods in Ecology, Wiley, London, 14–51, 2009.

Melton, J. R., Wania, R., Hodson, E. L., Poulter, B., Ringeval, B., Spahni, R., Bohn, T., Avis, C. A., Beerling, D. J., Chen, G., Eliseev, A. V., Denisov, S. N., Hopcroft, P. O., Lettenmaier, D. P., Riley, W. J., Singarayer, J. S., Subin, Z. M., Tian, H., Zürcher, S., Brovkin, V., van Bodegom, P. M., Kleinen, T., Yu, Z. C., and Kaplan, J. O.: Present state of global wetland extent and wetland methane modelling: conclusions from a model inter-comparison project (WETCHIMP), Biogeosciences, 10, 753–788, doi:10.5194/bg-10-753-2013, 2013.

Mesinger, F., Dimego, G., Kalnay, E., Mitchell, K., Shafran, P. C., Ebisuzaki, W., Jovi, D., Woollen, J., Rogers, E., Berbery, E. H., Ek, M. B., Fan, Y., Grumbine, R., Higgins, W., Li, H., Lin, Y., Manikin, G., Parrish, D., and Shi, W.: North American regional reanalysis, B. Am. Meteorol. Soc., 87, 343–360, doi:10.1175/BAMS-87-3-343, 2006.

Miller, S. M., Wofsy, S. C., Michalak, A. M., Kort, E. A., Andrews, A. E., Biraud, S. C., Dlugokencky, E. J., Eluszkiewicz, J., Fischer, M. L., Janssens-Maenhout, G., Miller, B. R., Miller, J. B., Montzka, S. A., Nehrkorn, T., and Sweeney, C.: Anthropogenic emissions of methane in the United States, P. Natl. Acad. Sci. USA, 110, 20018–20022, doi:10.1073/pnas.1314392110, 2013.

Miller, S. M., Worthy, D. E. J., Michalak, A. M., Wofsy, S. C., Kort, E. A., Havice, T. C., Andrews, A. E., Dlugokencky, E. J., Kaplan, J. O., Levi, P. J., Tian, H., and Zhang, B.: Observational constraints on the distribution, seasonality, and environmental predictors of North American boreal methane emissions, Global Biogeochem. Cy., 28, 146–160, doi:10.1002/2013GB004580, 2014.

Mueller, K. L., Yadav, V., Curtis, P. S., Vogel, C., and Michalak, A. M.: Attributing the variability of eddy-covariance CO_2 flux measurements across temporal scales using geostatistical regression for a mixed northern hardwood forest, Global Biogeochem. Cy., 24, doi:10.1029/2009GB003642, 2010.

Nehrkorn, T., Eluszkiewicz, J., Wofsy, S. C., Lin, J. C., Gerbig, C., Longo, M., and Freitas, S.: Coupled Weather Research and Forecasting-Stochastic Time-Inverted Lagrangian Transport (WRF-STILT) model, Meteorol. Atmos. Phys., 107, 51–64, doi:10.1007/s00703-010-0068-x, 2010.

Olivier, J. and Janssens-Maenhout, G.: CO_2 emissions from fuel combustion, 2012 edn., chap. III, in: Greenhouse-Gas Emissions, International Energy Agency, Paris, III.1 – III.51, 2012.

Pan, L. L., Bowman, K. P., Atlas, E. L., Wofsy, S. C., Zhang, F., Bresch, J. F., Ridley, B. A., Pittman, J. V., Homeyer, C. R., Romashkin, P., and Cooper, W. A.: The stratosphere-troposphere analyses of regional transport 2008 experiment, B. Am. Meteorol. Soc., 91, 327–342, doi:10.1175/2009BAMS2865.1, 2010.

Pickett-Heaps, C. A., Jacob, D. J., Wecht, K. J., Kort, E. A., Wofsy, S. C., Diskin, G. S., Worthy, D. E. J., Kaplan, J. O., Bey, I., and Drevet, J.: Magnitude and seasonality of wetland methane emissions from the Hudson Bay Lowlands (Canada), Atmos. Chem. Phys., 11, 3773–3779, doi:10.5194/acp-11-3773-2011, 2011.

Prigent, C., Papa, F., Aires, F., Rossow, W. B., and Matthews, E.: Global inundation dynamics inferred from multiple satellite observations, 1993–2000, J. Geophys. Res.-Atmos., 112, D12107, doi:10.1029/2006JD007847, 2007.

Ramsey, F. and Schafer, D.: The Statistical Sleuth: A Course in Methods of Data Analysis, Cengage Learning, Boston, 3rd edn., 2012.

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, a methane biogeochemistry model integrated in CESM, Biogeosciences, 8, 1925–1953, doi:10.5194/bg-8-1925-2011, 2011.

Ringeval, B., de Noblet-Ducoudré, N., Ciais, P., Bousquet, P., Prigent, C., Papa, F., and Rossow, W. B.: An attempt to quantify the impact of changes in wetland extent on methane emissions on the seasonal and interannual time scales, Global Biogeochem. Cy., 24, GB2003, doi:10.1029/2008GB003354, 2010.

Shiga, Y. P., Michalak, A. M., Gourdji, S. M., Mueller, K. L., and Yadav, V.: Detecting fossil fuel emissions patterns from subcontinental regions using North American in situ CO_2 measurements, Geophys. Res. Lett., 41, 4381–4388, doi:10.1002/2014GL059684, 2014.

Singarayer, J. S., Valdes, P. J., Friedlingstein, P., Nelson, S., and Beerling, D. J.: Late Holocene methane rise caused by orbitally controlled increase in tropical sources, Nature, 470, 82–85, doi:10.1038/nature09739, 2011.

21

Spahni, R., Wania, R., Neef, L., van Weele, M., Pison, I., Bousquet, P., Frankenberg, C., Foster, P. N., Joos, F., Prentice, I. C., and van Velthoven, P.: Constraining global methane emissions and uptake by ecosystems, Biogeosciences, 8, 1643–1665, doi:10.5194/bg-8-1643-2011, 2011.

Tarnocai, C., Canadell, J. G., Schuur, E. A. G., Kuhry, P., Mazhitova, G., and Zimov, S.: Soil organic carbon pools in the northern circumpolar permafrost region, Global Biogeochem. Cy., 23, doi:10.1029/2008GB003327, 2009.

Tian, H., Xu, X., Liu, M., Ren, W., Zhang, C., Chen, G., and Lu, C.: Spatial and temporal patterns of CH_4 and N_2O fluxes in terrestrial ecosystems of North America during 1979–2008: application of a global biogeochemistry model, Biogeosciences, 7, 2673–2694, doi:10.5194/bg-7-2673-2010, 2010.

Turner, A. J., Jacob, D. J., Wecht, K. J., Maasakkers, J. D., Lundgren, E., Andrews, A. E., Biraud, S. C., Boesch, H., Bowman, K. W., Deutscher, N. M., Dubey, M. K., Griffith, D. W. T., Hase, F., Kuze, A., Notholt, J., Ohyama, H., Parker, R., Payne, V. H., Sussmann, R., Sweeney, C., Velazco, V. A., Warneke, T., Wennberg, P. O., and Wunch, D.: Estimating global and North American methane emissions with high spatial resolution using GOSAT satellite data, Atmos. Chem. Phys., 15, 7049–7069, doi:10.5194/acp-15-7049-2015, 2015.

Waddington, J. and Roulet, N.: Atmosphere-wetland carbon exchanges: scale dependency of CO_2 and CH_4 exchange on the developmental topography of a peatland, Global Biogeochem. Cy., 10, 233–245, doi:10.1029/95GB03871, 1996.

Wania, R., Ross, I., and Prentice, I. C.: Implementation and evaluation of a new methane model within a dynamic global vegetation model: LPJ-WHyMe v1.3.1, Geosci. Model Dev., 3, 565–584, doi:10.5194/gmd-3-565-2010, 2010.

Wania, R., Melton, J. R., Hodson, E. L., Poulter, B., Ringeval, B., Spahni, R., Bohn, T., Avis, C. A., Chen, G., Eliseev, A. V., Hopcroft, P. O., Riley, W. J., Subin, Z. M., Tian, H., van Bodegom, P. M., Kleinen, T., Yu, Z. C., Singarayer, J. S., Zürcher, S., Lettenmaier, D. P., Beerling, D. J., Denisov, S. N., Prigent, C., Papa, F., and Kaplan, J. O.: Present state of global wetland extent and wetland methane modelling: methodology of a model inter-comparison project (WETCHIMP), Geosci. Model Dev., 6, 617–641, doi:10.5194/gmd-6-617-2013, 2013.

Wecht, K. J., Jacob, D. J., Frankenberg, C., Jiang, Z., and Blake, D. R.: Mapping of North American methane emissions with high spatial resolution by inversion of SCIAMACHY satellite data, J. Geophys. Res.-Atmos., 119, 7741–7756, doi:10.1002/2014JD021551, 2014.

Winderlich, J.: Setup of a CO_2 and CH_4 measurement system in Central Siberia and modeling of its results, PhD thesis, University of Hamburg, Hamburg, Germany, available at: http://ediss.sub.uni-hamburg.de/volltexte/2012/5533/pdf/Dissertation.pdf (last access: 14 June 2015), 2012.

Winderlich, J., Chen, H., Gerbig, C., Seifert, T., Kolle, O., Lavrič, J. V., Kaiser, C., Höfer, A., and Heimann, M.: Continuous low-maintenance CO_2/CH_4/H_2O measurements at the Zotino Tall Tower Observatory (ZOTTO) in Central Siberia, Atmospheric Measurement Techniques, 3, 1113–1128, doi:10.5194/amt-3-1113-2010, 2010.

5 Yadav, V., Mueller, K. L., Dragoni, D., and Michalak, A. M.: A geostatistical synthesis study of factors affecting gross primary productivity in various ecosystems of North America, Biogeosciences, 7, 2655–2671, doi:10.5194/bg-7-2655-2010, 2010.

Table 1. Spatial flux patterns chosen by the model selection framework.

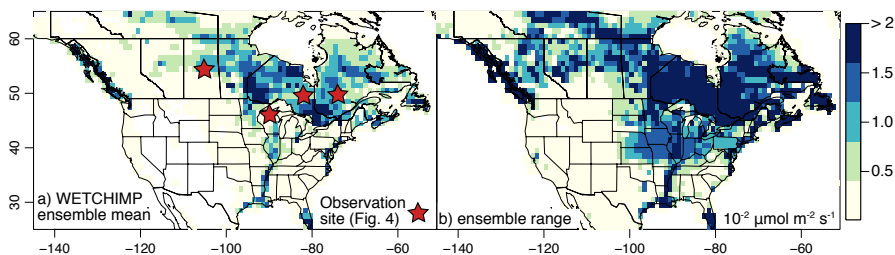| Region	Season	Chosen models
E. Canada	summer	LPJ-Bern, SDGVM
E. Canada	fall	LPJ-Bern
W. Canada	summer	SDGVM

Figure 1. Mean of the annual methane fluxes estimated by the WETCHIMP models **(a)** and the range of fluxes estimated by the ensemble **(b)**. Note that the range in estimates is larger than the mean. The fluxes shown above are the average flux per m^2 of land area, not per m^2 of wetland area.

Figure 2. The US and Canadian atmospheric methane observation network for 2007–2008 (14 703 total observations). Small yellow dots indicate observations from the START08 measurement campaign (Pan et al., 2010). Larger dots indicate tower and aircraft sites with regular observations over the two year period (Andrews et al., 2014). The grey background delineates the four regions used in the synthetic data experiments (Sect. 2.2).
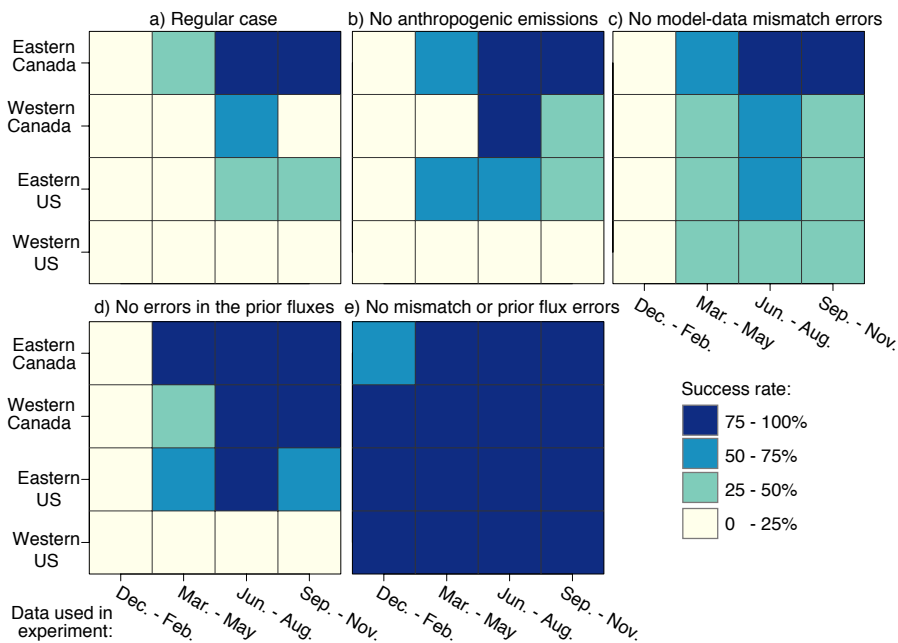
Figure 3. This figure displays the results of the synthetic data experiments. These experiments examine whether the observation network can detect aggregate wetland CH_4 fluxes. The figure shows the percentage of trials that are successful. Darker shades indicate that the network is more sensitive to wetland fluxes in the given region and season. Panel **(a)** shows the results for the standard setup while panels **(b-e)** show the results of several test cases in which anthropogenic emissions or different errors are set to zero.
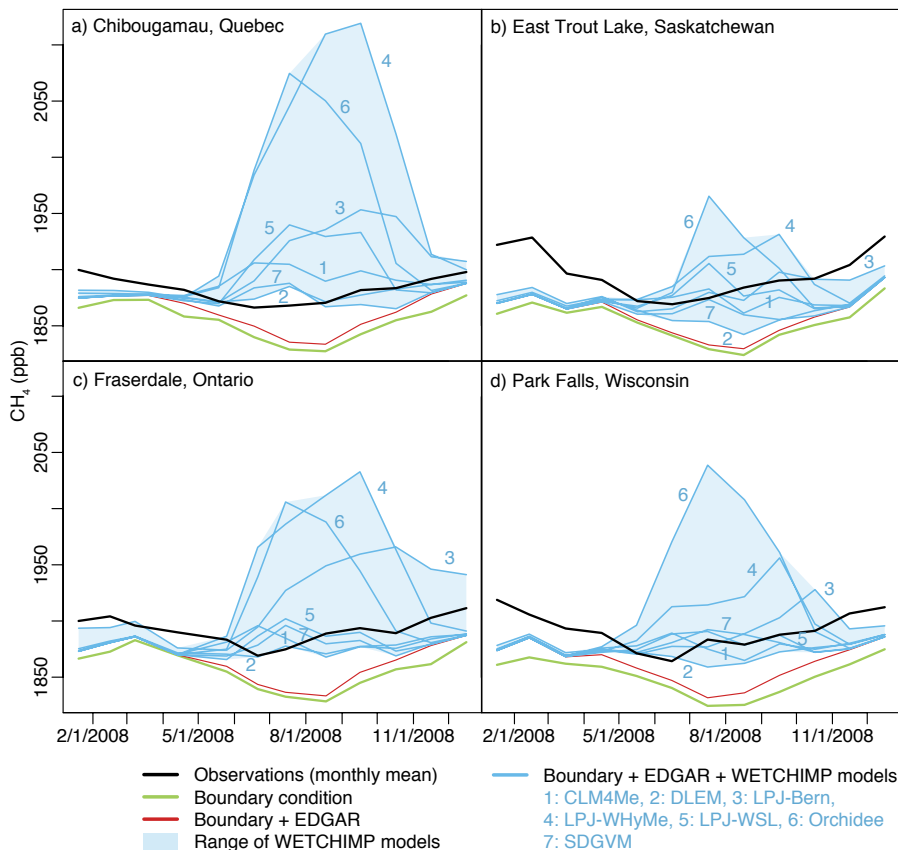
27

Figure 4. These time series compare atmospheric methane measurements at several observation sites against model estimates using the WETCHIMP ensemble and the EDGAR v4.2FT2010 anthropogenic emissions inventory. Refer to Fig. S4 for model-data time series at additional sites, particularly sites that are distant from large wetlands.
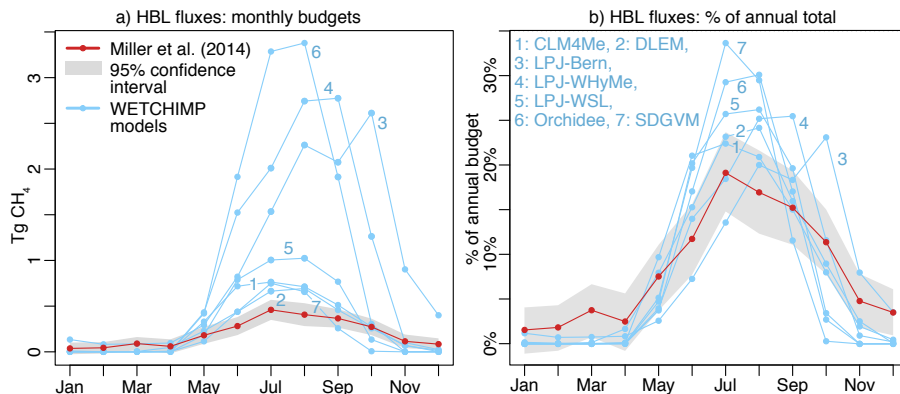
Figure 5. The seasonal cycle in methane fluxes estimated for the HBL (50–60° N, 75–96° W). We include both the WETCHIMP estimates and an inverse modeling estimate from Miller et al. (2014). Panel **(a)** displays the monthly budget from each estimate while **(b)** displays each month as a percentage of the annual budget estimated by a given model.