

RESEARCH ARTICLE

10.1002/2013GB004580

Key Points:

- Boreal methane emissions differ from existing inventories
- Estimate methane from the Hudson Bay Lowlands of 1.8 ± 0.24 Tg C/yr
- A simple methane emissions model best predicts atmospheric methane observations

Supporting Information:

- Readme
- Text S1

Correspondence to:

S. M. Miller,
smiller@fas.harvard.edu

Citation:

Miller, S. M., et al. (2014), Observational constraints on the distribution, seasonality, and environmental predictors of North American boreal methane emissions, *Global Biogeochem. Cycles*, 28, 146–160, doi:10.1002/2013GB004580.

Received 30 JAN 2013

Accepted 16 JAN 2014

Accepted article online 23 JAN 2014

Published online 26 FEB 2014

Observational constraints on the distribution, seasonality, and environmental predictors of North American boreal methane emissions

Scot M. Miller¹, Doug E. J. Worthy², Anna M. Michalak³, Steven C. Wofsy¹, Eric A. Kort⁴, Talya C. Havice¹, Arlyn E. Andrews⁵, Edward J. Dlugokencky⁵, Jed O. Kaplan⁶, Patricia J. Levi¹, Hanqin Tian⁷, and Bowen Zhang⁷

¹Department of Earth and Planetary Sciences, Harvard University, Cambridge, Massachusetts, USA, ²Environment Canada, Toronto, Ontario, Canada, ³Department of Global Ecology, Carnegie Institution for Science, Stanford, California, USA, ⁴Department of Atmospheric, Oceanic and Space Sciences, University of Michigan, Ann Arbor, Michigan, USA, ⁵Global Monitoring Division, NOAA Earth System Research Laboratory, Boulder, Colorado, USA, ⁶Environmental Engineering Institute, Ecole Polytechnique Federale de Lausanne, Lausanne, Switzerland, ⁷School of Forestry and Wildlife Sciences, Auburn University, Auburn, Alabama, USA

Abstract Wetlands comprise the single largest global source of atmospheric methane, but current flux estimates disagree in both magnitude and distribution at the continental scale. This study uses atmospheric methane observations over North America from 2007 to 2008 and a geostatistical inverse model to improve understanding of Canadian methane fluxes and associated biogeochemical models. The results bridge an existing gap between traditional top-down, inversion studies, which typically emphasize total emission budgets, and biogeochemical models, which usually emphasize environmental processes. The conclusions of this study are threefold. First, the most complete process-based methane models do not always describe available atmospheric methane observations better than simple models. In this study, a relatively simple model of wetland distribution, soil moisture, and soil temperature outperformed more complex model formulations. Second, we find that wetland methane fluxes have a broader spatial distribution across western Canada and into the northern U.S. than represented in existing flux models. Finally, we calculate total methane budgets for Canada and for the Hudson Bay Lowlands, a large wetland region (50–60°N, 75–96°W). Over these lowlands, we find total methane fluxes of 1.8 ± 0.24 Tg C yr⁻¹, a number in the midrange of previous estimates. Our total Canadian methane budget of 16.0 ± 1.2 Tg C yr⁻¹ is larger than existing inventories, primarily due to high anthropogenic emissions in Alberta. However, methane observations are sparse in western Canada, and additional measurements over Alberta will constrain anthropogenic sources in that province with greater confidence.

1. Introduction

Atmospheric methane (CH₄) is the second most important long-lived greenhouse gas, and since the preindustrial era, its radiative forcing has increased to 0.507 W m⁻², approximately one third that of CO₂ [Butler, 2012]. Therefore, greenhouse gas reduction strategies and future climate predictions will require accurate estimates of methane emissions. Total global emissions are constrained to approximately $\pm 15\%$ using observations of the global CH₄ burden and rate of increase, combined with an estimate of the CH₄ atmospheric lifetime [e.g., Kirschke et al., 2013]. However, uncertainties in emissions from individual source types can be greater than a factor of 2 [O'Connor et al., 2010; Dlugokencky et al., 2011; Melton et al., 2013]. For example, wetlands likely constitute the largest single source of atmospheric methane, but estimates of global fluxes vary from 60 to 213 Tg C yr⁻¹ (80 to 284 Tg CH₄ yr⁻¹), meaning they comprise anywhere from 14 to 50% of the total budget [e.g., O'Connor et al., 2010; Melton et al., 2013; Bridgham et al., 2013; Kirschke et al., 2013; Ciais et al., 2013, and references therein]. Anthropogenic sources (e.g., fossil fuel extraction and processing, ruminants, and landfills), by comparison, likely account for 50–65% of total global emissions [Ciais et al., 2013; Kirschke et al., 2013]. Uncertainties in methane fluxes are larger at the regional scale; estimates of methane from the Hudson Bay Lowlands (HBL), a large boreal wetland region in Canada, range from 0.28 to 8.5 Tg C yr⁻¹ [Roulet et al., 1992; Worthy et al., 2000; Pickett-Heaps et al., 2011; Melton et al., 2013].

The present study focuses on improving methane flux estimates from boreal wetlands. These regions are a particular concern because of their large soil carbon stocks. Methane fluxes in wetlands occur primarily in waterlogged, anaerobic soil conditions due to the decomposition of organic material by methanogenic Archaea. Boreal and arctic regions are far less productive than many other ecosystems but nonetheless play a vital role in the global carbon cycle. These northerly regions may contain half of all wetlands and soil carbon in the world (~ 1700 Pg C), twice the amount of carbon currently held within the atmosphere [Tarnocai *et al.*, 2009].

Evidence suggests that high-latitude wetlands are already changing due to an evolving climate and that ecosystem changes may accelerate [Tarnocai, 2009; Avis *et al.*, 2011; Schuur *et al.*, 2013]. For example, most studies predict that climate change will increase methane fluxes from boreal and arctic regions; estimates range from 6% to 35% increase in methane fluxes per $^{\circ}\text{C}$ of global temperature increase [e.g., Gedney *et al.*, 2004; Khvorostyanov *et al.*, 2008; O'Connor *et al.*, 2010; Koven *et al.*, 2011; Zhu *et al.*, 2011].

Three factors may explain the large differences among model estimates of boreal methane fluxes. First, models differ in their underlying environmental variables. For example, existing models of global wetland area range from 2.6 to 9×10^6 km² [Petrescu *et al.*, 2010] and have differing spatial distributions (especially over boreal North America [Melton *et al.*, 2013]). Second, models further differ in functional form (see section 2.4), due in part to uncertainties and/or complexity in biophysical methane processes. For example, many models relate maps of soil temperature to wetland methane fluxes using a coefficient known as Q10. This coefficient describes the factor by which a reaction rate increases per 10°C rise in temperature. Estimates of this coefficient range from 1 to 35, largely due to microbial and soil heterogeneity [van Hulzen *et al.*, 1999; Whalen, 2005; O'Connor *et al.*, 2010; Lupascu *et al.*, 2012]. Finally, differences among existing flux models also stem from difficulties extrapolating from plot level to regional scale. Most flux models calibrate to individual wetland sites and extrapolate to regional or global scales [O'Connor *et al.*, 2010; Zhang *et al.*, 2012]. However, small-scale study sites exhibit substantial heterogeneity, and fluxes can vary by an order of magnitude over microtopography on the centimeter scale [Waddington and Roulet, 1996; Comas *et al.*, 2005; Hendriks *et al.*, 2010].

Top-down approaches like inverse modeling provide one means of reducing the wide uncertainty in wetland methane fluxes. Top-down studies use atmospheric methane measurements and meteorological models to improve existing flux estimates at regional [Zhao *et al.*, 2009; Bergamaschi *et al.*, 2010; Villani *et al.*, 2010; Kim *et al.*, 2011] and global [Chen and Prinn, 2006; Bergamaschi *et al.*, 2013; Fraser *et al.*, 2013] scales. Most existing methods emphasize total emissions budgets and provide relatively little information on wetland processes, but two recent publications begin to bridge this gap. Spahni *et al.* [2011] conduct a global-scale inversion that estimates fluxes by wetland type. Pickett-Heaps *et al.* [2011] use atmospheric methane measurements from northern Ontario to assess the magnitude and seasonal structure of a wetland flux model over the HBL. Results imply a premature seasonal onset of fluxes in this model, referred to as the "Kaplan model." The authors suggest removing fluxes from snow-covered regions as one possible solution. In spite of these recent studies, existing top-down approaches provide limited assessment of the underlying environmental variables or the functional form of existing wetland flux models.

The present study moves closer to integrating top-down flux estimates with process-based, bottom-up modeling methods. First, we explore how atmospheric methane measurements can be used to construct and assess biogeochemical process models at continental scale. Second, we use a broad network of measurement sites in Canada and the U.S. to understand the spatial and seasonal distribution of North American boreal wetland fluxes. To achieve these goals, we combine in situ methane measurements across Canada and the United States from 2007 and 2008, a regional atmospheric transport model, and a geostatistical inverse modeling framework.

2. Model and Measurements

The methods sections and subsequent discussion are organized as follows. First, we describe the atmospheric model and measurements (sections 2.1, 2.2, and 2.3). Using this model, we compare two existing wetland flux estimates, Kaplan and DLEM, against atmospheric methane observations. Both flux estimates are described in detail below (sections 2.4). We subsequently use a geostatistical inverse modeling

framework to estimate North American boreal methane fluxes (section 3.1). This flux estimate has two components. The first component, termed the deterministic model, is a combination of environmental predictors (e.g., soil moisture and temperature) that best represent the methane fluxes, as seen through the atmospheric methane observations (section 3.2). The second component, termed the stochastic component, estimates the spatial and/or temporal flux patterns that may be lacking in the environmental predictors and therefore cannot be modeled using the deterministic model. The geostatistical inverse model produces a final best estimate, termed the posterior fluxes, and it is the sum of the deterministic and stochastic components.

2.1. The Regional Atmospheric Model

We simulate in situ methane mixing ratios using STILT, the Stochastic, Time-Inverted, Lagrangian Transport model [Lin *et al.*, 2003]. STILT is a particle model; an ensemble of air-following particles is released from each methane observation site. In this study, a new 500-particle ensemble is initiated for each of the hourly methane measurements. These particles travel backward in time along the wind fields of a meteorology model, in this case for 10 days. STILT further includes stochastic motions that simulate boundary layer turbulence.

Wind fields from the Weather Research and Forecasting model (WRF version 2.1.2) are used to drive STILT trajectories in this study. Nehr Korn *et al.* [2010], Hegarty *et al.* [2013], and the supporting information describe this meteorology in greater detail. The WRF fields used here have a nested resolution, 10 km within 24–48 h of the observation sites and 40 km in more distant regions (see supporting information).

STILT subsequently uses the trajectories to calculate a footprint map. The footprints relate the surface fluxes in North America to the concentration increment seen at the measurement location and have units of mixing ratio per unit surface flux. This footprint is based on the number of particles in a region and their altitudes relative to the planetary boundary layer.

The STILT setup here incorporates fluxes from existing inventories on a $\frac{1}{4}^{\circ}$ by $\frac{1}{6}^{\circ}$ longitude-latitude grid (11 to 65°N and 145 to 51°W).

2.2. Model Boundary Condition

STILT only models emissions over the North American continent. The model therefore requires a boundary condition to represent the concentration of methane in incoming air over the Pacific and Arctic oceans before reaching North American sources. This study uses an empirical boundary curtain that interpolates a variety of trace gas measurements from ground-based sites and aircraft in the NOAA Earth System Research Laboratory Global Monitoring Division's Cooperative Global Air Sampling Network. The resulting boundary curtain varies latitudinally and vertically and has a daily temporal resolution (see the supporting information). The estimated boundary condition value associated with each STILT particle run depends on the ending latitude, altitude, and day of each particle. This boundary value is then added to the modeled methane signal from North American sources. The sum can be directly compared against measured methane mixing ratios at tower sites across Canada and the northern U.S. (e.g., section 3.1).

2.3. Measurements

This study uses observed methane mixing ratios for 2007 to 2008 from five observation sites sensitive to boreal wetland fluxes: hourly measurements from four Canadian observation towers and daily flask measurements from one U.S. tall tower. Sites (from east to west) include Chibougamau, Quebec (CHM, 50°N, 74°W, 30 m above ground level (agl)); Fraserdale, Ontario (FSD, 50°N, 83°W, 40 m agl); Park Falls, Wisconsin (LEF, 46°N, 90°W, 244 m agl); East Trout Lake, Saskatchewan (ETL, 54°N, 104°W, 105 m agl); and Candle Lake, Saskatchewan (CDL, 54°N, 105°W, 30 m agl, 2007 only) (Figure 1).

Small-scale heterogeneities caused by turbulent eddies and incomplete mixing make it difficult to model hourly-scale variability in the in situ data. STILT also has difficulty estimating the very shallow nighttime boundary layer and therefore rarely captures variations in nighttime concentrations. Hence, this study uses afternoon averages of the methane data and model output (1 P.M. to 7 P.M. local time), a total of 2485 observations after averaging.

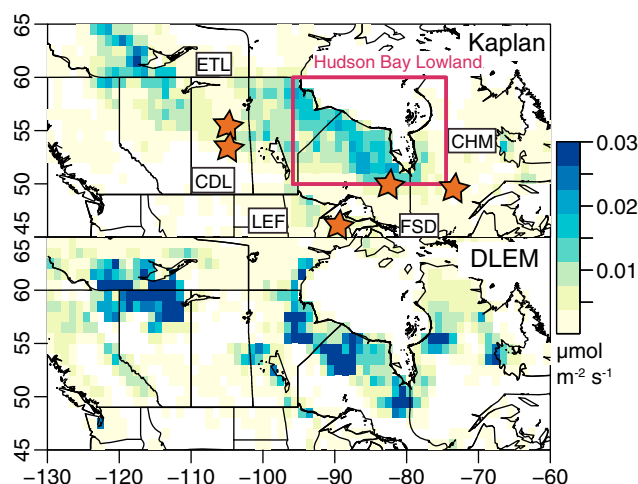


Figure 1. Summer mean wetland fluxes from the Kaplan and DLEM wetland methane models (for July, August, and September, averaged over 2007–2008). Both models estimate similar annual totals for the HBL, but DLEM has a more pronounced summer peak.

2.4. Existing Flux Models

2.4.1. The Kaplan Model

The first inventory used in this study is the Kaplan model, described in *Kaplan* [2002] and *Pickett-Heaps et al.* [2011] (Figure 1). The model has the following functional form:

$$E = \delta b W f(T) \sum_{i=1}^2 \frac{C_i}{\tau_i} \quad (1)$$

$$f(T) = 1.0 \times 10^3 F e^{\frac{-309}{T-227}}$$

where E is the wetland flux expressed here in units of $\mu\text{mol m}^{-2} \text{s}^{-1}$. C_1 and C_2 represent the moles of carbon per unit area in soil and litter, normalized by their respective lifetimes (τ_1 and τ_2). This soil carbon estimate is taken from the Lund-Potsdam-Jena (LPJ) dynamic global vegetation model [*Sitch et al.*, 2003]. W is the maximum possible extent of wetlands in a given grid box (a fraction, from LPJ), and δ is a measure of whether wetlands are actually present ($\delta = 0$ if soil moisture (M) $< 10\%$ and $\delta = 1$ otherwise). An emissions factor b represents the fraction of methane per mole of carbon respired ($b = 3 \times 10^{-2}$). $f(T)$ represents an Arrhenius equation of temperature (in Kelvin). In this equation, F adjusts the inventory based on soil temperature to better match differences between boreal (B) and tropical (T) wetlands (refer to *Pickett-Heaps et al.* [2011]). We build this inventory with soil moisture (for δ) and soil temperature (T) from WRF (the same meteorology used to drive STILT) at a soil depth of 25 cm. This soil depth provided the best match between the Kaplan model and atmospheric methane observations. This setup differs from *Pickett-Heaps et al.* [2011], who used surface skin temperature instead of soil temperature below the ground surface.

The LPJ model outputs used here for wetland coverage and soil carbon are updated from previous studies that also used the LPJ/Kaplan model [e.g., *Bergamaschi et al.*, 2007; *Pickett-Heaps et al.*, 2011]. Among other updates, soil carbon is approximately a factor of 4 lower than in the previous studies listed above. This adjustment matches the LPJ model against upland soil profiles, but the change appears inconsistent with methane observations over boreal wetlands [e.g., *Pickett-Heaps et al.*, 2011]. We readjust the LPJ soil carbon estimate upward by a factor of 4.15 to match the LPJ/Kaplan model in *Pickett-Heaps et al.* [2011]. This previous study compares the Kaplan model against measurements from Frasersdale (FSD), Ontario, and likely better represents high-latitude soil carbon than the new LPJ estimate.

2.4.2. The DLEM Model

DLEM, the Dynamic Land Ecosystem Model, includes more complexity than the Kaplan model described above [*Tian et al.*, 2010] (Figure 1). It models the production of methane in soil pore water (P , expressed here in $\mu\text{mol m}^{-2} \text{s}^{-1}$), and only a fraction of methane produced is released to the atmosphere (E). This fraction depends on a multitude of factors that are discussed in depth by *Tian et al.* [2010]: plant-mediated transport, diffusive flux, ebullition, oxidation by methanotrophy, and oxidation during plant-mediated transport.

Methane production in soil pore water (P) is simpler to describe:

$$P = P_{\max} \times \frac{[\text{DOC}]}{[\text{DOC}] + k} \times f(T) \times f(\text{pH}) \times f(M) \quad (2)$$

where P_{\max} is the maximum possible rate of CH_4 production in soils, a spatially variable parameter [see *Tian et al.*, 2010]. $[\text{DOC}]$ is dissolved organic content, determined by gross primary productivity, litter fall, and soil organic matter decomposition rates [*Tian et al.*, 2012]. k is the half-saturation coefficient, $f(T)$ is the effect of soil temperature, $f(\text{pH})$ is the effect of soil pH, and $f(M)$ is the effect of soil water content. The functions of temperature, pH, and soil moisture have the following forms [*Tian et al.*, 2010]:

$$f(T) = \begin{cases} 0 & T < 268.15 \\ 2.5^{(T-303.15)/10} & 303.15 > T \geq 268.15 \\ 1 & T \geq 303.15 \end{cases} \quad (3)$$

$$f(\text{pH}) = \begin{cases} 0 & \text{pH} < 4.0 \text{ or } \text{pH} \geq 10.0 \\ \frac{1.02}{1+1e6 \times \exp[-2.5\text{pH}]} & 4.0 < \text{pH} < 7.0 \\ \frac{1.02}{1+1e6 \times \exp[-2.5(14.0-\text{pH})]} & 7.0 < \text{pH} < 10.0 \end{cases} \quad (4)$$

$$f(M) = \begin{cases} 0 & M \leq M_{fc} \\ 0.368 \left(\frac{M-M_{fc}}{M_s-M_{fc}} \right)^2 \exp \left[\frac{M-M_{fc}}{M_s-M_{fc}} \right] & M_{fc} < M < M_s \\ 1 & M \geq M_s \end{cases} \quad (5)$$

where M_{fc} is the field capacity, and M_s is the saturated water content of soil. *Tian et al.* [2010] provide a graphical depiction of these functional dependences.

The environmental data for DLEM are derived from a number of sources: meteorological data from North American Regional Reanalysis (NARR) [*Mesinger et al.*, 2006] and land cover/vegetation data from a combination of sources (see *Tian et al.* [2010] for more details).

3. Statistical Framework

3.1. Conceptual Overview

We implement a geostatistical inverse model to infer information about methane fluxes and to assess the environmental drivers in existing wetland models. The statistical approach follows that of *Kitanidis and Vomvoris* [1983], *Michalak et al.* [2004], and *Gourdji et al.* [2012]. The inversion estimates the spatial and temporal distribution of emissions that is most likely given the atmospheric methane measurements and the transport information provided by the atmospheric model.

The inversion first requires a linear expression for the model-measurement framework:

$$\mathbf{z} = \mathbf{H}\mathbf{s} + \boldsymbol{\varepsilon} \quad (6)$$

where \mathbf{s} ($m \times 1$) are the true, unknown fluxes. Unlike the wetland-specific fluxes estimated by Kaplan and DLEM (E), \mathbf{s} encompasses fluxes from all source types. \mathbf{z} is the $n \times 1$ vector of observed mixing ratios minus the estimated boundary condition value (see section 2.2). \mathbf{H} ($n \times m$) are the footprints computed by STILT (section 2.1). Finally, $\boldsymbol{\varepsilon}$ ($n \times 1$) describes model-data mismatch—all errors unrelated to an imperfect emissions estimate (e.g., transport error and aggregation error). This vector is assumed to follow a multivariate normal distribution with a mean of 0:

$$\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{R}) \quad (7)$$

where \mathbf{R} ($n \times n$) is the covariance matrix of these errors.

Using the above framework, the inversion then models the unknown fluxes (\mathbf{s} , equation (6)) using the following structure:

$$\mathbf{s} \sim \mathbf{X}\boldsymbol{\beta} + \mathcal{N}(\mathbf{0}, \mathbf{Q}) \quad (8)$$

Table 1. Auxiliary Data or Predictors Tested for Use in the Deterministic Model^a

Description	Static/Variable	Source Model
Liquid soil moisture (e.g., not frozen) (M)	variable	WRF, NARR ^b
Total soil moisture (liquid + frozen) (M_{Tot})	variable	WRF, NARR
Soil temperature	variable	WRF, NARR
The maximum fraction of a region that could be covered by wetlands (W)	static	LPJ, GIEMS
Soil carbon content (C)	static	LPJ, NCSCD
The estimated distribution of anthropogenic emissions	static	EDGAR v4.2
Smooth tricubic functions over anthropogenic source regions	static	

^aSee sections 2.4 and 3.2.1. The second column (Static/Variable) lists whether the auxiliary data in question are seasonally constant or vary temporally.

^bSoil moisture and temperature are available at multiple vertical soil levels in WRF and NARR: 5, 25, 70, and 150 cm depth in WRF and 0, 10, 40, and 100 cm depth in NARR.

The first component of the statistical model ($\mathbf{X}\beta$) is a weighted least squares regression and is termed the “deterministic model” or “inversion prior” (section 3.2). Each column of \mathbf{X} (dimensions $m \times p$) is a predictor in the weighted regression [e.g., *Gourdji et al.*, 2008, 2012]. In this study, \mathbf{X} includes data sets termed “auxiliary data” (e.g., soil temperature, moisture, and/or an anthropogenic emissions inventory) that help explain the spatial and seasonal distribution of methane fluxes. Additionally, one column of this matrix is constant, equivalent to the intercept of the regression. The regression coefficients (β , dimensions $p \times 1$) are unknown and are estimated in the inversion using the atmospheric methane data.

The second component the geostatistical inverse model, $\mathcal{N}(\mathbf{0}, \mathbf{Q})$, is termed the “stochastic component” or the “spatially correlated residual.” The stochastic component adjusts, at grid scale, the fluxes estimated by the deterministic model. This component, for example, can correct the deterministic model if any environmental data in \mathbf{X} have the incorrect distribution. The covariance matrix \mathbf{Q} (dimensions $m \times m$) describes the variance and the spatiotemporal correlation of the stochastic component. It includes off-diagonal elements that follow an exponential covariance model: any fluxes estimated by the stochastic component will be spatially correlated with a given decorrelation length. This spatial correlation means that the stochastic component can adjust the flux estimate on a fine grid scale relative to the density of atmospheric observations [e.g., *Michalak et al.*, 2004; *Mueller et al.*, 2008; *Villani et al.*, 2010; *Bergamaschi et al.*, 2013; *Miller et al.*, 2013].

The best estimate of a geostatistical inversion is obtained by minimizing a cost function (L) with respect to the methane fluxes (\mathbf{s}) and the coefficients (β) [e.g., *Kitanidis and Vomvoris*, 1983; *Michalak et al.*, 2004]:

$$L_{\mathbf{s}, \beta} = \frac{1}{2}(\mathbf{z} - \mathbf{H}\mathbf{s})^T \mathbf{R}^{-1}(\mathbf{z} - \mathbf{H}\mathbf{s}) + \frac{1}{2}(\mathbf{s} - \mathbf{X}\beta)^T \mathbf{Q}^{-1}(\mathbf{s} - \mathbf{X}\beta) \quad (9)$$

The supporting information discusses further details of the statistical setup. In particular, we implement the inversion with Lagrange multipliers to prevent negative fluxes (see supporting information) [*Miller et al.*, 2014]. Furthermore, we estimate the covariance matrices (\mathbf{R} and \mathbf{Q}) using restricted maximum likelihood estimation (REML) [*Kitanidis*, 1995; *Michalak et al.*, 2004].

We use this statistical framework to estimate monthly methane fluxes (\mathbf{s}) on a 1° by 1° longitude-latitude grid over the years 2007 and 2008, yielding 41,328 total locations in space and time. The geographic domain of the inversion spans from 35 to 65°N latitude and 145 to 51°W longitude.

3.2. The Deterministic Model of Fluxes

The following sections discuss the deterministic model in greater detail.

3.2.1. Auxiliary Environmental Data

We consider a number of auxiliary data sets or predictors for use in the deterministic model. Ultimately, only a selection of these data sets is used in the inversion depending on how well each explains the atmospheric methane data (see section 3.2.2). These data sets include both environmental drivers of wetland fluxes and inventory data on anthropogenic emissions. The full array of possible data sets for \mathbf{X} are shown in Table 1. These include meteorological data from WRF (used in this version of Kaplan model [*Nehrkorn et al.*, 2010]) and NARR (used in DLEM [*Mesinger et al.*, 2006]). We consider soil carbon estimates from the LPJ model (used in Kaplan [*Sitch et al.*, 2003; *Pickett-Heaps et al.*, 2011]) and the Northern Circumpolar Soil Carbon Database (NCSCD) [*Tarnocai et al.*, 2009; *Hugelius et al.*, 2013]. Wetland coverage estimates include model output from

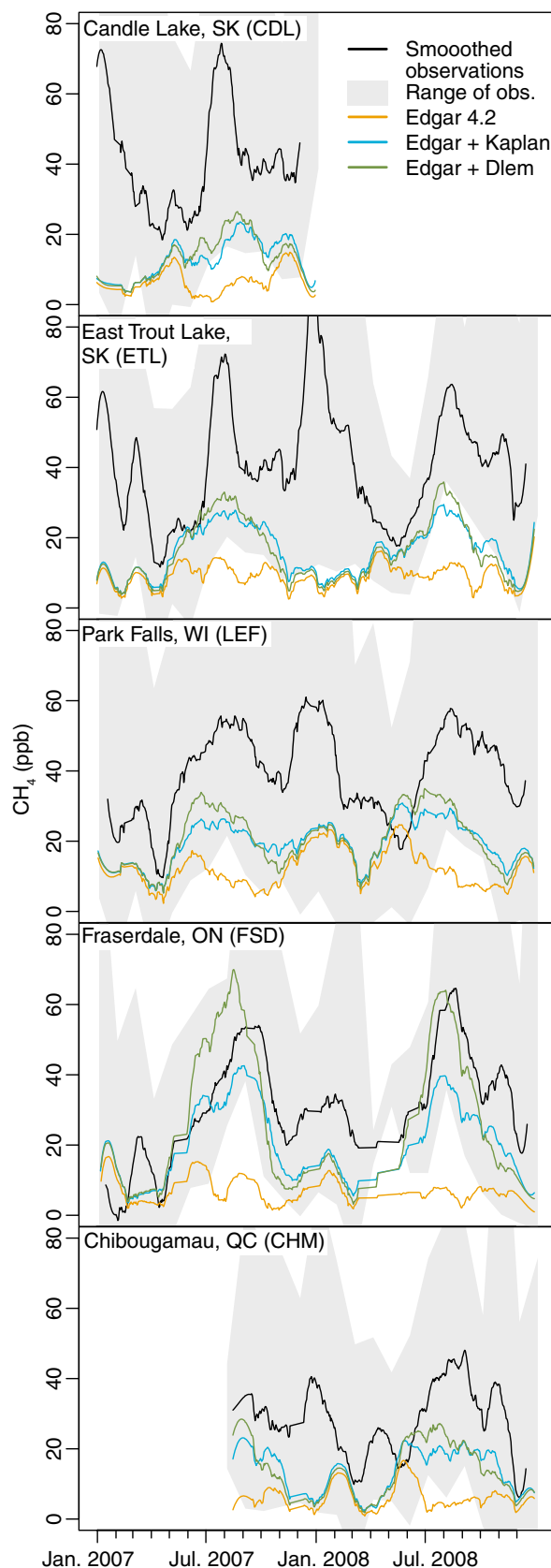


Figure 2

LPJ and surface water data from the Global Inundation Extent from Multi-Satellites (GIEMS) database [Prigent *et al.*, 2007; Papa *et al.*, 2010]. Refer to the supporting information for maps of these auxiliary data sets.

In addition to wetland-related data sets, we also consider multiple data sets or predictors for the distribution of anthropogenic emissions. Specifically, we consider including the EDGAR v4.2 anthropogenic inventory in the deterministic model as well as the individual sector-by-sector emissions estimates from EDGAR. A companion study found that EDGAR v4.2 did not match the estimated distribution of anthropogenic emissions in the United States [Miller *et al.*, 2013]. Hence, we consider additional proxies other than EDGAR v4.2 for the spatial distribution of anthropogenic emissions. For example, we construct smooth tricube functions centered over known anthropogenic source regions (e.g., Alberta, Oklahoma, California, and the U.S. East Coast; refer to the supporting information). The subsequent section discusses how to choose among this array of auxiliary data sets when constructing the deterministic model.

3.2.2. Selection of Auxiliary Data

It would be ill advised to use all auxiliary data sets from Table 1 in the deterministic model; the resulting model would be an overfit with problematic colinearity [e.g., Zucchini, 2000]. We instead use a statistical selection method to choose an optimal set of auxiliary data sets for the deterministic model. These methods select as many data sets for X as can explain variability in the methane fluxes but will prevent an overfit or unreliable coefficient estimates. We implement one of the most common methods, the Bayesian information criterion (BIC) (as in Gourdji *et al.* [2012]). The BIC numerically scores all possible combinations of auxiliary data based on how well they reduce the model-measurement residuals and applies an increasing penalty for model complexity (refer to the supporting information). Specifically, this penalty increases with the number of columns

Table 2. BIC Scores for a Selection of Candidate Deterministic Models^a

Candidate Model	BIC
$\beta_0 + \beta_1$ [smooth functions over anthropogenic source regions] + $\beta_2[W][M]f_{\text{Kaplan}}(T)$	16,725
$\beta_0 + \beta_1$ [smooth functions ...] + $\beta_2[W][M_{\text{Tot}}]f_{\text{Kaplan}}(T)$	16,728
$\beta_0 + \beta_1$ [smooth functions ...] + $\beta_2[W][M]f_{\text{Kaplan}}(T)f_{\text{Kaplan}}(C)$	16,729
$\beta_0 + \beta_1$ [smooth functions ...] + β_2 [full Kaplan model]	16,735
$\beta_0 + \beta_1$ [smooth functions ...] + $\beta_2[W][M]f_{\text{Kaplan}}(T)$ using NARR surface soil layer	16,744
$\beta_0 + \beta_1$ [smooth functions ...] + $\beta_2[W]f_{\text{DLEM}}(M)f_{\text{DLEM}}(T)$	16,750
$\beta_0 + \beta_1$ [EDGAR v4.2] + $\beta_2[W][M]f_{\text{Kaplan}}(T)$	16,885

^aWe test all possible combinations and interactions of the auxiliary variables in Table 1 and display only a sample here. The table is intended to show the range of BIC scores for the best scoring models and a few other notable models. The drift coefficients (β) scale the magnitude of the auxiliary data to match the methane observations. All models above use inputs from NARR (10 cm depth) and LPJ, unless otherwise noted. $f_{\text{Kaplan}}(\cdot \cdot \cdot)$ refers to the functional form used in the Kaplan model and $f_{\text{DLEM}}(\cdot \cdot \cdot)$ the functional form in DLEM.

in \mathbf{X} and with the log of the number of observations. Unlike frequentist statistics, these scores do not support p values or traditional hypothesis testing. The best model is simply the one with the lowest score. *Kass and Raftery* [1995] provide a qualitative assessment of model strength based on the difference in BIC scores. A score difference greater than 2 is “worth mentioning” and greater than 10 is “very strong.”

In many cases, one might expect that the product of two or more different environmental variables may be a better predictor than an additive model, so we test multiplicative interactions among the wetland-related auxiliary data sets. Additionally, several of the auxiliary data sets are colinear (e.g., total soil moisture and unfrozen soil moisture), and we are careful not to include similar or colinear predictors in the same candidate model for \mathbf{X} . For consistency, we do not mix WRF and NARR data sets in the same candidate model.

4. Results and Discussion

4.1. Model-Data Comparison Using Existing Flux Estimates

Methane concentrations modeled with existing flux estimates exhibit a variable fit against the atmospheric data (see Figure 2). For example, both the Kaplan and DLEM models match the general shape of the seasonal cycle at eastern tower sites (LEF, FSD, and CHM) but underestimate the magnitude of the measurements. Among these sites, models match observations most closely at Fraserdale, Ontario (FSD), possibly because *Pickett-Heaps et al.* [2011] validated the Kaplan model at Fraserdale. Existing methane flux estimates, however, perform far worse at the western sites (CDL and ETL). For example, the models underestimate both observed summer and winter maxima at these sites. The observed summer maxima are likely caused by peak summer wetland fluxes, while the winter maxima likely reflect a combination of advected anthropogenic emissions and limited vertical mixing within the troposphere. This result implies that existing inventories underestimate both wetland and anthropogenic fluxes in western Canada.

The model-data comparison in Figure 2 also reveals important conclusions about the interdependence of wetland and anthropogenic emissions estimates. Gaps in anthropogenic emissions inventories can affect the perceived amplitude or seasonality of the wetland flux model. Even in remote regions like the HBL, the estimation of wetland fluxes hinges on a reliable anthropogenic emissions estimate. For example, the Kaplan/EDGAR v4.2 modeled concentrations are consistently too low at the Fraserdale site, but the amplitude of the summer maximum is similar to the amplitude of the data. This discrepancy could reflect one of two problems: either the wetland inventory has the incorrect magnitude and seasonal structure or the

Figure 2. A comparison of modeled mixing ratios against measurements at the observation sites. The estimated boundary condition values have been subtracted from the observations; the difference indicates the effect of North American methane sources on the measurement sites. EDGAR v4.2 is an anthropogenic emissions inventory, while Kaplan and DLEM model wetlands. The model and observations are smoothed using a third-order Savitzky-Golay filter with a 61-point window.

Table 3. Canada Methane Budgets From the Deterministic Model (South of 65°N) and Several Inventory Estimates

Flux Model	Canada Budget (Tg C yr ⁻¹)
Deterministic model	
$\hat{\beta}_0$	5.4 ± 1.5
$\hat{\beta}_1$ [smooth functions]	7.9 ± 0.9
$\hat{\beta}_2[W][M]f_{\text{Kaplan}}(T)$	3.2 ± 0.6
Existing wetland models	
Kaplan model	4.4
DLEM model	5.6
Existing anthropogenic inventories	
Environment Canada	3.3
EDGAR v4.2	3.9

anthropogenic inventory (EDGAR v4.2) is simply too low. The time series at Park Falls (LEF), Wisconsin, further illustrates the importance of the anthropogenic emissions estimate. It appears that the wetland flux models begin producing methane too early in the spring of 2008 at LEF. A closer examination of Figure 2, however, reveals large (~25 ppb) modeled concentrations from anthropogenic sources during this period. This model-data discrepancy could stem from misspecified anthropogenic emissions, not problems in the seasonal structure of the wetland model. These examples highlight

the difficulty of disentangling anthropogenic and wetland methane fluxes.

Subsequent sections discuss the deterministic model and geostatistical inversion results in greater detail.

4.2. Environmental Predictors of Wetland Fluxes

This section explores the results of the deterministic model ($X\beta$, sections 3.1 and 3.2). As discussed in the methods sections, the deterministic model is analogous to a weighted multivariate regression. Model selection methods (like the BIC, section 3.2.2) play a crucial role in constructing this deterministic model; they select auxiliary data sets (Table 1) for the deterministic model that can best explain the atmospheric methane data. In this way, model selection provides a means to objectively understand and assess biogeochemical methane models at continental scale.

The BIC selection chooses the following deterministic model for methane fluxes in Canada (Table 2):

$$\beta_0 + \beta_1[\text{smooth functions}] + \beta_2[W][M]f_{\text{Kaplan}}(T) \tag{10}$$

This selected model for methane fluxes is relatively simple. The first term (β_0) is a constant component, equivalent to the intercept in a regression. It describes the average magnitude of all sources not explicitly included in other components of the deterministic model. For example, this component might include agriculture, landfills, and wastewater treatment sources (among other possibilities).

The second term (β_1 [smooth functions ...]) parameterizes anthropogenic sources (section 3.2.1). This term places smooth geometric functions over known source regions, including Alberta, California, Oklahoma, and the U.S. east coast. The BIC does not choose the EDGAR v4.2 anthropogenic inventory for the deterministic model because it fits the atmospheric data less well than the smooth geometric functions (Table 2). Hence, we do not utilize EDGAR within the atmospheric inversion.

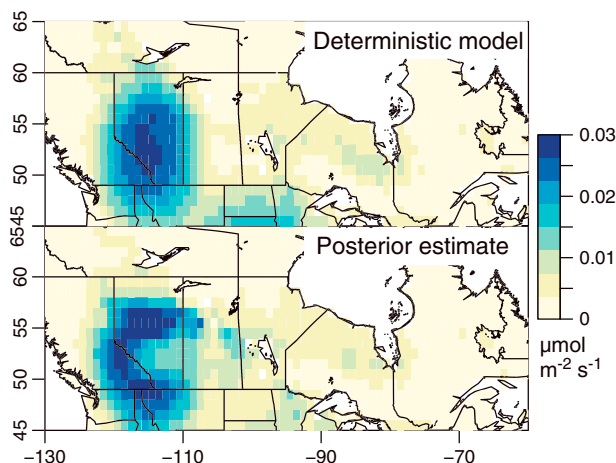
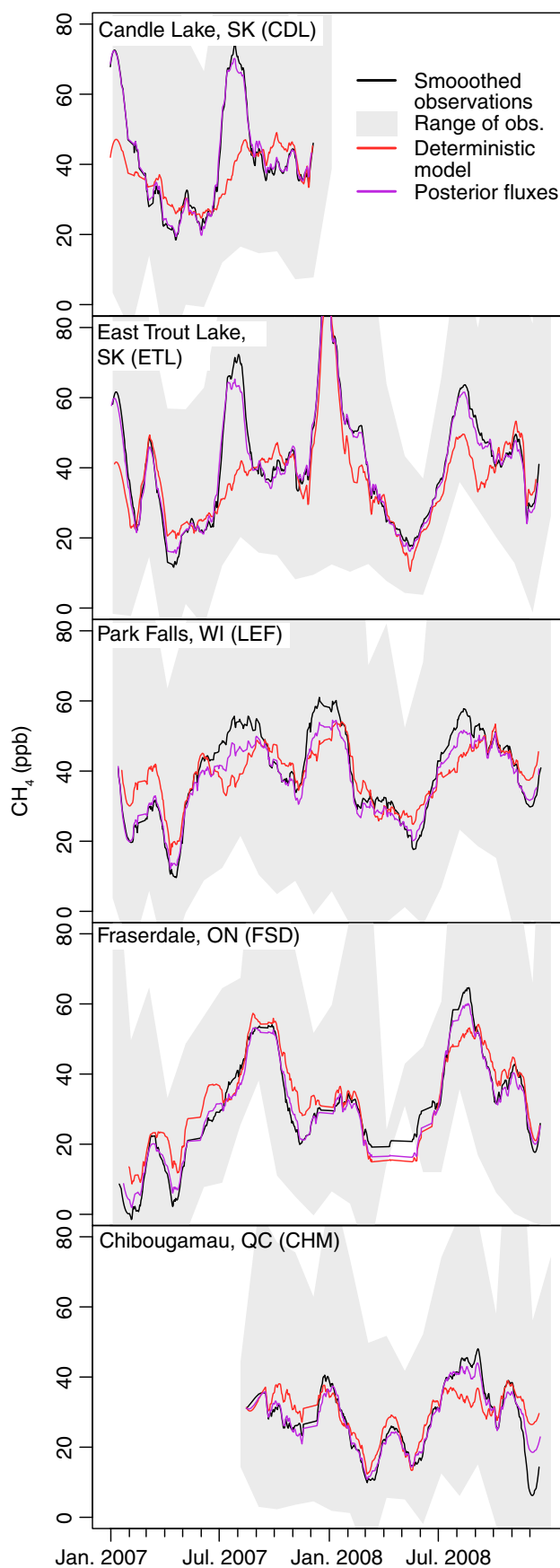


Figure 3. The 24 month (2 year) mean estimated methane flux from (top) the deterministic model and (bottom) the final posterior estimate.

The final component of the deterministic model ($\beta_2[W][M]f_{\text{Kaplan}}(T)$) parameterizes wetland fluxes. This term in the deterministic model includes three auxiliary data sets: the distribution of wetlands (W), a map of unfrozen soil moisture (M), and an Arrhenius equation based upon soil temperature ($f_{\text{Kaplan}}(T)$). The optimal deterministic model uses the wetland map from the LPJ model and soil variables from NARR (at 10 cm soil depth). All other possible combinations and interactions of the auxiliary variables in Table 1 produce higher BIC scores (Table 2). For example, we test wetland models that include soil carbon, environmental variables at different depths in the



soil profile, and different estimates for wetland distribution. Furthermore, we test a deterministic model that uses the functional form of temperature and/or soil moisture from the DLEM model.

This selected wetland model is similar to the Kaplan flux model but with soil carbon removed. Section 4.4 synthesizes the wetland flux results from this study and highlights what this parameterized wetland model might indicate about biogeochemical methane modeling.

Table 3 lists the Canadian methane budget associated with each component of the deterministic model and compares these estimates against existing inventories. It is important to remember that the methane budgets from the deterministic model are estimated using the atmospheric data—via the unknown coefficients, β . The smooth functions represent the largest component of the deterministic model, followed by the constant component, and finally the wetland component. When interpreting these budgets, however, it is important to note that the constant component $\hat{\beta}_0$ could represent either anthropogenic emissions or wetland fluxes.

Figures 3 and 4.2 visualize the deterministic model, both spatially and in relation to the atmospheric methane data. Figure 3 displays the annual average of the deterministic model. The smooth geometric functions to parameterize anthropogenic emissions are evident over the province of Alberta and over the Dakotas. The wetland model is more difficult to distinguish in this annual mean plot but is largest south of Hudson Bay in eastern Canada and near Great Slave Lake in the Northwest Territories. The deterministic model is nonzero everywhere across Canada, and this reflects

Figure 4. A comparison of modeled mixing ratios against measurements at the observation sites. This figure is similar to Figure 2 but compares the deterministic model and posterior emissions estimate instead of existing flux models.

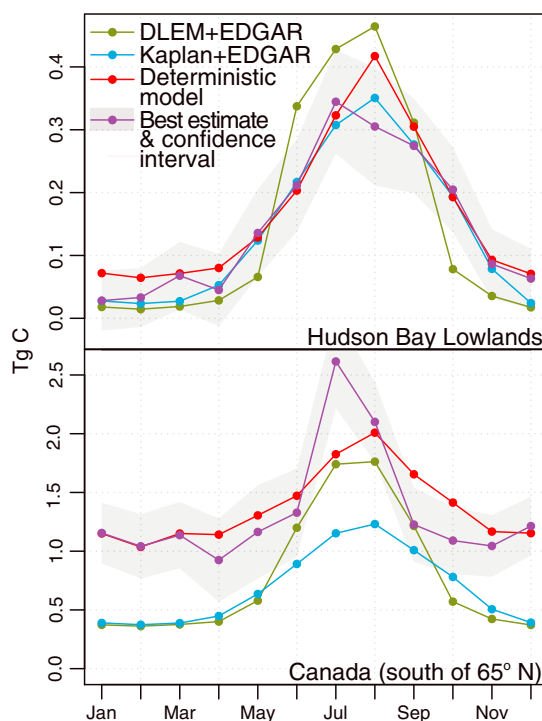


Figure 5. The monthly average methane budget estimated for the HBL and Canada in 2007–2008. Existing models underestimate wetland fluxes in western Canada. This regional shortfall explains much of the summertime discrepancy between the flux models and the posterior estimate in the lower panel.

the constant term β_0 of the deterministic model. This term has an estimated magnitude of $2 \pm 0.5 \times 10^{-3} \mu\text{mol m}^{-2} \text{s}^{-1}$ ($5.4 \pm 1.5 \text{ Tg C yr}^{-1}$ over Canada, Table 3).

Despite the simplicity of the deterministic model, the mixing ratios estimated with model match favorably against atmospheric measurements (Figure 4.2). The deterministic model fits the atmospheric methane observations ($R = 0.72$, root mean squared error (RMSE) = 20.9 ppb) better than either the model setup with Kaplan and EDGAR v4.2 ($R = 0.12$, RMSE = 37.1 ppb) or DLEM and EDGAR v4.2 ($R = 0.08$, RMSE = 37.2 ppb). The formulation of anthropogenic emissions in the deterministic model may account for much of this improved fit against the atmospheric data. Despite the improvement, the deterministic model displays two notable shortfalls. First, the deterministic model does not reproduce the summer maxima observed at western observation sites (CDL and ETL). Second, the deterministic model underestimates the summer maxima at the Wisconsin (LEF) and Quebec (CHM) observation sites. These shortfalls suggest that the spatial distribution of wetland fluxes in the deterministic model may be too restrictive. In other words, wetland fluxes likely extend

further west, east, and south than in the deterministic model, which places the largest wetland fluxes in the HBL.

Subsequent sections discuss the final methane flux estimate from the geostatistical inversion ($\hat{\xi}$). This final, best estimate ($R = 0.89$, RMSE = 12.0 ppb) is henceforth referred to as the “posterior” fluxes.

4.3. The Spatial and Temporal Distribution of Emissions

The posterior flux estimate identifies two major source regions in Canada (Figure 3): over Alberta in western Canada and over the HBL in eastern Canada. This discussion analyzes each geographic region individually.

In western Canada, the inversion identifies a large, seasonally constant methane source region over Alberta. In the deterministic model, this source is represented by a smooth function. But in the posterior estimate, this source region becomes a more well-defined crescent shape over Alberta (Figure 3). These emissions likely originate from anthropogenic activity, and a future study will give an in-depth analysis of anthropogenic emissions in Canada. The posterior flux estimate also includes a large summer source in Alberta and Saskatchewan. As discussed previously, these fluxes are not represented by the auxiliary environmental data sets in the deterministic model. This omission in western Canada dominates the discrepancy in summertime Canadian methane between the deterministic model and posterior fluxes (Figure 5). The omission implies that either the LPJ wetland or the NARR soil moisture map is an underestimate in westerly regions of Canada. Unfortunately, the atmospheric data in this region have limited capability to pinpoint the exact location of these western wetland fluxes; atmospheric observations are sparse in western Canada, and wetland fluxes are colocated with large anthropogenic sources. In sum, this study identifies Alberta as a region with poorly known wetland fluxes and as a possible hot spot of anthropogenic emissions. We recommend that future methane measurement efforts focus on Alberta because this province is a key uncertainty in current understanding of Canadian methane sources.

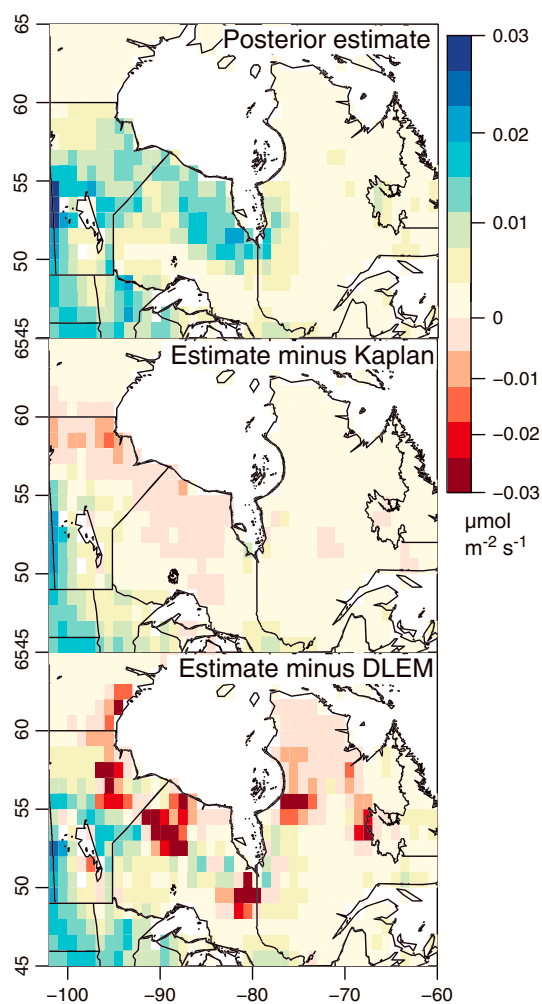


Figure 6. (top) The posterior flux estimate averaged over all summer months (July to September for 2007–2008). (Bottom) The difference between the posterior estimate and the DLEM and Kaplan methane models.

existing estimates. Anthropogenic emissions in western Canada may explain much of this discrepancy. In contrast, our annual HBL budget is consistent with that of DLEM and *Pickett-Heaps et al.* [2011] who use the Kaplan wetland model, but our estimate diverges from a site-based study by *Roulet et al.* [1992] and a box model study by *Worthy et al.* [2000] (see the supporting information). Furthermore, the HBL budget estimated here is low compared to the array of biogeochemical models listed in *Melton et al.* [2013]. The HBL budgets in those models range from 1.7 to 8.5 Tg C yr⁻¹. This range in wetland methane estimates is likely greater than the interannual variability in wetland fluxes. For example, *Tian et al.* [2010] estimate an 11% standard deviation in annual North American methane fluxes.

4.4. A Synthesis Perspective on Biogeochemical Methane Models

This section explores the study's implications for biogeochemical methane modeling. The inversion results (e.g., section 4.2) raise the question of why a simple flux model fits the atmospheric methane data as well as sophisticated process models. The deterministic model developed here excludes a number of factors that can affect methane fluxes: soil carbon, plant-mediated transport, and heterogeneities in microbial communities, among many other processes. This question could be answered in two ways.

First, simple parameterizations may be sufficient when regional-scale flux patterns are the primary goal. For example, a synthesis study of existing chamber measurement sites found that methane fluxes across all sites are influenced most strongly by only a few environmental variables: water table height, soil

Eastern Canada, in contrast, is dominated by seasonal methane fluxes that presumably emanate from wetlands. Figure 5 compares the seasonal cycle of DLEM, Kaplan, the deterministic model, and the posterior flux estimate over the HBL. The seasonal cycle of the deterministic model and posterior flux estimate compare similarly to the Kaplan model but have a broader seasonality than DLEM. The posterior flux estimate matches the Kaplan model more closely than the deterministic model over the HBL (though the deterministic model is a better match than Kaplan/EDGAR v4.2 in other regions of Canada and the northern U.S.).

Seasonal structure aside, the flux models also diverge in spatial distribution. Figure 6 displays the mean summer (July, August, and September) methane flux estimated by the inversion for eastern Canada. It also displays the difference between this estimate and the DLEM and Kaplan models. Our flux estimate is more spatially dispersive than DLEM across the Hudson Bay region. The differences between the posterior estimate and Kaplan are more subtle. The posterior estimate indicates methane fluxes across a broader region than Kaplan: into Minnesota, Wisconsin, Manitoba, and further west.

Figure 7 summarizes the findings of this study as an annual methane budget estimate for the HBL and for all of Canada (south of 65° latitude). Our methane estimate for Canada is a factor of 1.5 to 2.2 times

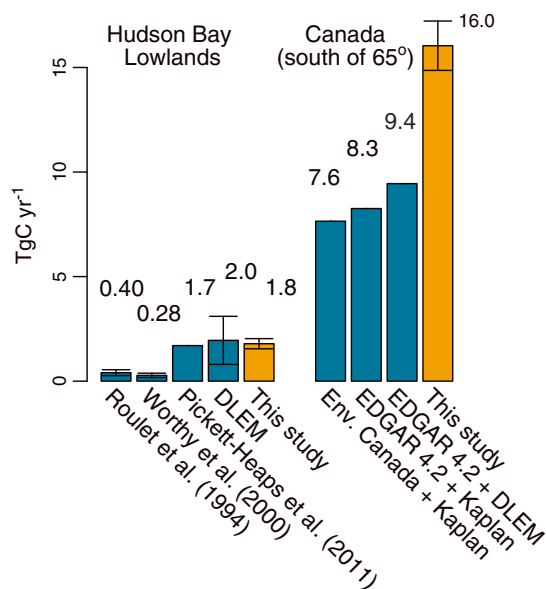


Figure 7. Total methane budgets from this study and others for the HBL and for Canada. The HBL estimates listed here are from DLEM and from observational studies. *Melton et al.* [2013] list numerous additional model-based HBL methane budgets, which range from 1.7 to 8.5 Tg C yr⁻¹. “Env. Canada” refers to *Environment Canada’s* [2013] National Inventory Report. The confidence intervals for this study do not encompass uncertainties in model selection and therefore may underestimate the total budget uncertainty.

temperature, and vegetation type [*Olefeldt et al.*, 2013]. Furthermore, *Bubier et al.* [1993] and *Waddington and Roulet* [1996] argue that most centimeter-scale flux variability ultimately depends on two primary parameters: temperature and water table position. These studies imply that a simple model may adequately parameterize regional-scale flux variability. A second reason may account for the simplicity of the deterministic model. Complex methane flux processes can be challenging to upscale, meaning that the most complete methane model is not always the most accurate at regional scales. The spatial distribution of many flux-related processes is highly uncertain [e.g., *Melton et al.*, 2013] due to a paucity of both land surface and methane flux data. This uncertainty means that models with many processes and parameters could run the risk of overfitting limited, available data (A paper by *Zucchini* [2000] illustrates the hazards of overfitting.). For example, a number of physical processes like ebullition, plant-mediated transport, and microbial community dynamics are all thought to play a role in methane emissions

[*Bridgham et al.*, 2013, and references therein], but how these processes or features vary on regional spatial scales is often poorly understood.

To that end, model selection methods, like the BIC used here, provide a means to diagnose weaknesses in flux model upscaling from plot level to regional or continental scale [e.g., *Olefeldt et al.*, 2013]. Model selection methods choose the set of predictors that can best explain variability in any available methane data. If model selection does not choose a given predictor, that outcome implies one of several conclusions. Either the distribution of the predictor does not match against the distribution implied by the methane data, or the available methane data are insufficient to constrain the effect of that predictor. In either case, any conclusions based upon the predictor would likely overfit the available data at the expense of describing the large-scale flux process of interest.

5. Conclusions

This study uses atmospheric methane observations and geostatistical inverse modeling to understand North American boreal methane fluxes and associated biogeochemical models. The conclusions of this study fall under three general themes. First, we find that a simple wetland flux model, when combined with WRF-STILT, provides as good agreement with atmospheric methane observations as more complex flux process models. This result may have several possible causes: either simple models adequately parameterize regional-scale flux patterns, or the spatiotemporal distribution of important but complex flux processes is difficult to accurately model with available data at this geographic scale.

Second, we estimate both the spatial and seasonal distributions of methane fluxes over much of boreal North America. We find wetland fluxes that are more broadly distributed than in existing inventories, even extending into Minnesota, Wisconsin, Manitoba, and western Canada. This result implies that existing maps may under-represent the extent of soil moisture and/or the distribution of wetlands.

Finally, we calculate regional and Canadian methane budgets. Our HBL budget is on the upper range of observational studies but on the lower range of biogeochemical model estimates (Figure 7 and Melton *et al.* [2013]). In addition, we estimate total Canadian emissions that exceed existing inventories, largely due to sources in or near Alberta. Available atmospheric data are limited near Alberta during the study period, and this work highlights a need for more intensive methane measurements over that region.

Acknowledgments

This work was supported by the American Meteorological Society Graduate Student Fellowship/DOE Atmospheric Radiation Measurement Program, the DOE Computational Science Graduate Fellowship, and the National Science Foundation Graduate Research Fellowship Program. We thank Thomas Nehrkorn and Janusz Eluszkiewicz for their help with the WRF meteorology and the NASA Advanced Supercomputing Division for their computing support. We further thank Catherine Prigent for providing GIEMS data. Support for this research was provided by NASA grants NNX08AR47G and NNX11AG47G, NOAA grants NA09OAR4310122 and NA11OAR4310158, NSF grant ATM-0628575, and Environmental Defense Fund grant 0146-10100. NOAA measurements were funded in part by the Atmospheric Composition and Climate Program and the Carbon Cycle Program of NOAA's Climate Program Office.

References

- Avis, C. A., A. J. Weaver, and K. J. Meissner (2011), Reduction in areal extent of high-latitude wetlands in response to permafrost thaw, *Nat. Geosci.*, *4*(7), 444–448, doi:10.1038/NNGEO1160.
- Bergamaschi, P., et al. (2013), Atmospheric CH₄ in the first decade of the 21st century: Inverse modeling analysis using SCIAMACHY satellite retrievals and NOAA surface measurements, *J. Geophys. Res. Atmos.*, *118*, 7350–7369, doi:10.1002/jgrd.50480.
- Bergamaschi, P., et al. (2007), Satellite cartography of atmospheric methane from SCIAMACHY on board ENVISAT: 2. Evaluation based on inverse model simulations, *J. Geophys. Res.*, *112*, D02304, doi:10.1029/2006JD007268.
- Bergamaschi, P., et al. (2010), Inverse modeling of European CH₄ emissions 2001–2006, *J. Geophys. Res.*, *115*, D22309, doi:10.1029/2010JD014180.
- Bridgman, S. D., H. Cadillo-Quiroz, J. K. Keller, and Q. Zhuang (2013), Methane emissions from wetlands: Biogeochemical, microbial, and modeling perspectives from local to global scales, *Global Change Biol.*, *19*(5), 1325–1346, doi:10.1111/gcb.12131.
- Bubier, J., A. Costello, T. R. Moore, N. T. Roulet, and K. Savage (1993), Microtopography and methane flux in boreal peatlands, northern Ontario, Canada, *Can. J. Botany*, *71*(8), 1056–1063, doi:10.1139/b93-122.
- Butler, J. (2012), The NOAA annual greenhouse gas index (AGGI). [Available at <http://www.esrl.noaa.gov/gmd/aggi/aggi.html>.]
- Chen, Y.-H., and R. G. Prinn (2006), Estimation of atmospheric methane emissions between 1996 and 2001 using a three-dimensional global chemical transport model, *J. Geophys. Res.*, *111*, D10307, doi:10.1029/2005JD006058.
- Ciais, P., C. Sabine, G. Bala, L. Bopp, V. Brovkin, and J. Canadell (2013), *Carbon and Other Biogeochemical Cycles—Final Draft Underlying Scientific Technical Assessment*, chap. 6, IPCC Secretariat, Geneva.
- Comas, X., L. Slater, and A. Reeve (2005), Geophysical and hydrological evaluation of two bog complexes in a northern peatland: Implications for the distribution of biogenic gases at the basin scale, *Global Biogeochem. Cycles*, *19*, GB4023, doi:10.1029/2005GB002582.
- Dlugokencky, E. J., E. G. Nisbet, R. Fisher, and D. Lowry (2011), Global atmospheric methane: Budget, changes and dangers, *Philos. Trans. R. Soc. London, Ser. A*, *369*(1943), 2058–2072, doi:10.1098/rsta.2010.0341.
- Environment Canada (2013), National inventory report 1990–2011: Greenhouse gas sources and sinks in Canada - executive summary, *Tech. Rep. ISSN: 1910-7064*, Environment Canada.
- Fraser, A., et al. (2013), Estimating regional methane surface fluxes: The relative importance of surface and GOSAT mole fraction measurements, *Atmos. Chem. Phys.*, *13*(11), 5697–5713, doi:10.5194/acp-13-5697-2013.
- Gedney, N., P. Cox, and C. Huntingford (2004), Climate feedback from wetland methane emissions, *Geophys. Res. Lett.*, *31*, L20503, doi:10.1029/2004GL020919.
- Gourdji, S. M., et al. (2012), North American CO₂ exchange: Inter-comparison of modeled estimates with results from a fine-scale atmospheric inversion, *Biogeosciences*, *9*(1), 457–475, doi:10.5194/bg-9-457-2012.
- Gourdji, S. M., K. L. Mueller, K. Schaefer, and A. M. Michalak (2008), Global monthly averaged CO₂ fluxes recovered using a geostatistical inverse modeling approach: 2. Results including auxiliary environmental data, *J. Geophys. Res.*, *113*, D21114, doi:10.1029/2007JD009733.
- Hegarty, J., R. R. Draxler, A. F. Stein, J. Brioude, M. Mountain, J. Eluszkiewicz, T. Nehrkorn, F. Ngan, and A. Andrews (2013), Evaluation of Lagrangian particle dispersion models with measurements from controlled tracer releases, *J. Appl. Meteorol. Climatol.*, *52*(12), 2623–2637.
- Hendriks, D. M. D., J. van Huissteden, and A. J. Dolman (2010), Multi-technique assessment of spatial and temporal variability of methane fluxes in a peat meadow, *Agric. For. Meteorol.*, *150*(6), 757–774, doi:10.1016/j.agrformet.2009.06.017.
- Hugelius, G., C. Tarnocai, G. Broll, J. G. Canadell, P. Kuhry, and D. K. Swanson (2013), The Northern Circumpolar Soil Carbon Database: Spatially distributed datasets of soil coverage and soil carbon storage in the northern permafrost regions, *Earth Syst. Sci. Data*, *5*(1), 3–13, doi:10.5194/essd-5-3-2013.
- Kaplan, J. (2002), Wetlands at the Last Glacial Maximum: Distribution and methane emissions, *Geophys. Res. Lett.*, *29*(6), 1079, doi:10.1029/2001GL013366.
- Kass, R., and A. Raftery (1995), Bayes factors, *J. Am. Stat. Assoc.*, *90*(430), 773–795, doi:10.2307/2291091.
- Khvorostyanov, D. V., P. Ciais, G. Krinner, and S. A. Zimov (2008), Vulnerability of east Siberia's frozen carbon stores to future warming, *Geophys. Res. Lett.*, *35*, L10703, doi:10.1029/2008GL033639.
- Kim, H.-S., S. Maksyutov, M. V. Glagolev, T. Machida, P. K. Patra, K. Sudo, and G. Inoue (2011), Evaluation of methane emissions from West Siberian wetlands based on inverse modeling, *Environ. Res. Lett.*, *6*(3), 035201, doi:10.1088/1748-9326/6/3/035201.
- Kirschke, S., et al. (2013), Three decades of global methane sources and sinks, *Nat. Geosci.*, *6*(10), 813–823.
- Kitanidis, P. (1995), Quasi-linear geostatistical theory for inverting, *Water Resour. Res.*, *31*(10), 2411–2419, doi:10.1029/95WR01945.
- Kitanidis, P. K., and E. G. Vomvoris (1983), A geostatistical approach to the inverse problem in groundwater modeling (steady state) and one-dimensional simulations, *Water Resour. Res.*, *19*(3), 677–690, doi:10.1029/WR019i003p0677.
- Koven, C. D., B. Ringeval, P. Friedlingstein, P. Ciais, P. Cadule, D. Khvorostyanov, G. Krinner, and C. Tarnocai (2011), Permafrost carbon-climate feedbacks accelerate global warming, *Proc. Natl. Acad. Sci. U.S.A.*, *108*(36), 14,769–14,774, doi:10.1073/pnas.1103910108.
- Lin, J., C. Gerbig, S. Wofsy, A. Andrews, B. Daube, K. Davis, and C. Grainger (2003), A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model, *J. Geophys. Res.*, *108*(D16), 4493, doi:10.1029/2002JD003161.
- Lupascu, M., J. Wadham, E. R. Hornibrook, and R. Pancost (2012), Temperature sensitivity of methane production in the permafrost active layer at Stordalen, Sweden: A comparison with non-permafrost northern wetlands, *Arct. Alp. Res.*, *44*(4), 469–482.
- Melton, J. R., et al. (2013), Present state of global wetland extent and wetland methane modelling: Conclusions from a model inter-comparison project (WETCHIMP), *Biogeosciences*, *10*(2), 753–788, doi:10.5194/bg-10-753-2013.
- Mesinger, F., et al. (2006), North American Regional Reanalysis., *Bull. Am. Meteorol. Soc.*, *87*, 343–360, doi:10.1175/BAMS-87-3-343.

- Michalak, A., L. Bruhwiler, and P. Tans (2004), A geostatistical approach to surface flux estimation of atmospheric trace gases, *J. Geophys. Res.*, *109*, D14109, doi:10.1029/2003JD004422.
- Miller, S. M., et al. (2013), Anthropogenic emissions of methane in the United States, *Proc. Natl. Acad. Sci. U.S.A.*, *110*(50), 20,018–20,022, doi:10.1073/pnas.1314392110.
- Miller, S. M., A. M. Michalak, and P. J. Levi (2014), Atmospheric inverse modeling with known physical bounds: An example from trace gas emissions, *Geosci. Model Dev.*, *7*, 303–315, doi:10.5194/gmd-7-303-2014.
- Mueller, K. L., S. M. Gourdj, and A. M. Michalak (2008), Global monthly averaged CO₂ fluxes recovered using a geostatistical inverse modeling approach: 1. Results using atmospheric measurements, *J. Geophys. Res.*, *113*, D21114, doi:10.1029/2007JD009734.
- Nehrkorn, T., J. Eluszkiewicz, S. C. Wofsy, J. C. Lin, C. Gerbig, M. Longo, and S. Freitas (2010), Coupled Weather Research and Forecasting-Stochastic Time-Inverted Lagrangian Transport (WRF-STILT) model, *Meteorol. Atmos. Phys.*, *107*(1–2), 51–64, doi:10.1007/s00703-010-0068-x.
- O'Connor, F. M., et al. (2010), Possible role of wetlands, permafrost, and methane hydrates in the methane cycle under future climate change: A review, *Rev. Geophys.*, *48*, RG4005, doi:10.1029/2010RG000326.
- Olefeldt, D., M. R. Turetsky, P. M. Crill, and A. D. McGuire (2013), Environmental and physical controls on northern terrestrial methane emissions across permafrost zones, *Global Change Biol.*, *19*(2), 589–603, doi:10.1111/gcb.12071.
- Papa, F., C. Prigent, F. Aires, C. Jimenez, W. B. Rossow, and E. Matthews (2010), Interannual variability of surface water extent at the global scale, 1993–2004, *J. Geophys. Res.*, *115*, D12111, doi:10.1029/2009JD012674.
- Petrescu, A. M. R., L. P. H. van Beek, J. van Huissteden, C. Prigent, T. Sachs, C. A. R. Corradi, F. J. W. Parmentier, and A. J. Dolman (2010), Modeling regional to global CH₄ emissions of boreal and arctic wetlands, *Global Biogeochem. Cycles*, *24*, GB4009, doi:10.1029/2009GB003610.
- Pickett-Heaps, C. A., et al. (2011), Magnitude and seasonality of wetland methane emissions from the Hudson Bay Lowlands (Canada), *Atmos. Chem. Phys.*, *11*(8), 3773–3779, doi:10.5194/acp-11-3773-2011.
- Prigent, C., F. Papa, F. Aires, W. B. Rossow, and E. Matthews (2007), Global inundation dynamics inferred from multiple satellite observations, 1993–2000, *J. Geophys. Res.*, *112*, D12107, doi:10.1029/2006JD007847.
- Roulet, N. T., R. Ash, and T. R. Moore (1992), Low boreal wetlands as a source of atmospheric methane, *J. Geophys. Res.*, *97*(D4), 3739–3749, doi:10.1029/91JD03109.
- Schuur, E., et al. (2013), Expert assessment of vulnerability of permafrost carbon to climate change, *Clim. Change*, *119*(2), 359–374, doi:10.1007/s10584-013-0730-7.
- Sitch, S., et al. (2003), Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model, *Global Change Biol.*, *9*(2), 161–185, doi:10.1046/j.1365-2486.2003.00569.x.
- Spahni, R., et al. (2011), Constraining global methane emissions and uptake by ecosystems, *Biogeosciences*, *8*(6), 1643–1665, doi:10.5194/bg-8-1643-2011.
- Tarnocai, C. (2009), The impact of climate change on Canadian peatlands, *Can. Water Resour. J.*, *34*(4), 453–466, doi:10.4296/cwrj3404453.
- Tarnocai, C., J. G. Canadell, E. A. G. Schuur, P. Kuhry, G. Mazhitova, and S. Zimov (2009), Soil organic carbon pools in the northern circumpolar permafrost region, *Global Biogeochem. Cycles*, *23*, GB2023, doi:10.1029/2008GB003327.
- Tian, H., X. Xu, M. Liu, W. Ren, C. Zhang, G. Chen, and C. Lu (2010), Spatial and temporal patterns of CH₄ and N₂O fluxes in terrestrial ecosystems of North America during 1979–2008: Application of a global biogeochemistry model, *Biogeosciences*, *7*(9), 2673–2694, doi:10.5194/bg-7-2673-2010.
- Tian, H., et al. (2012), Contemporary and projected biogenic fluxes of methane and nitrous oxide in North American terrestrial ecosystems, *Front. Ecol. Environ.*, *10*(10), 528–536.
- van Hulzen, J., R. Segers, P. van Bodegom, and P. Leffelaar (1999), Temperature effects on soil methane production: An explanation for observed variability, *Soil Biol. Biochem.*, *31*(14), 1919–1929, doi:10.1016/S0038-0717(99)00109-1.
- Villani, M. G., P. Bergamaschi, M. Krol, J. F. Meirink, and F. Dentener (2010), Inverse modeling of European CH₄ emissions: Sensitivity to the observational network, *Atmos. Chem. Phys.*, *10*(3), 1249–1267, doi:10.5194/acp-10-1249-2010.
- Waddington, J., and N. Roulet (1996), Atmosphere-wetland carbon exchanges: Scale dependency of CO₂ and CH₄ exchange on the developmental topography of a peatland, *Global Biogeochem. Cycles*, *10*(2), 233–245, doi:10.1029/95GB03871.
- Whalen, S. (2005), Biogeochemistry of methane exchange between natural wetlands and the atmosphere, *Environ. Eng. Sci.*, *22*(1), 73–94, doi:10.1089/ees.2005.22.73.
- Worthy, D., I. Levin, F. Hopper, M. Ernst, and N. Trivett (2000), Evidence for a link between climate and northern wetland methane emissions, *J. Geophys. Res.*, *105*, 4031–4038.
- Zhang, Y., T. Sachs, C. Li, and J. Boike (2012), Upscaling methane fluxes from closed chambers to eddy covariance based on a permafrost biogeochemistry integrated model, *Global Change Biol.*, *18*(4), 1428–1440, doi:10.1111/j.1365-2486.2011.02587.x.
- Zhao, C., A. E. Andrews, L. Bianco, J. Eluszkiewicz, A. Hirsch, C. MacDonald, T. Nehrkorn, and M. L. Fischer (2009), Atmospheric inverse estimates of methane emissions from central California, *J. Geophys. Res.*, *114*, D16302, doi:10.1029/2008JD011671.
- Zhu, X., Q. Zhuang, M. Chen, A. Sirin, J. Melillo, D. Kicklighter, A. Sokolov, and L. Song (2011), Rising methane emissions in response to climate change in Northern Eurasia during the 21st century, *Environ. Res. Lett.*, *6*(4), 045,211, doi:10.1088/1748-9326/6/4/045211.
- Zucchini, W. (2000), An introduction to model selection, *J. Math. Psychol.*, *44*(1), 41–61, doi:10.1006/jmps.1999.1276.