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

Scot M. Miller Doug E.J Worthy Anna M. Michalak Steven C. Wofsy
Eric A. Kort Talya C. Havice Arlyn E. Andrews Edward J. Dlugokencky
Jed O. Kaplan Patricia J. Levi Hanqin Tian Bowen Zhang

S1 Detailed model explanation and validation

S1.1 Meteorology overview

The WRF meteorological simulations used in this study have been constructed explicitly for trace gas modeling with STILT, and the same WRF fields have been used in a number of greenhouse gas modeling studies (e.g., *Kort et al.*, 2010; *Gourdji et al.*, 2010; *Huntzinger et al.*, 2011; *Gourdji et al.*, 2012; *Miller et al.*, 2013, 2014; *Shiga et al.*, 2013). Among other features, these simulations conserve mass by a factor of ten better than other meteorological products like the NCEP global analysis fields (FNL) (*Nehrkorn et al.*, 2010). *Nehrkorn et al.* (2010) and *Hegarty et al.* (2013) validate the WRF fields generated for STILT against meteorological observations and controlled tracer release experiments, respectively. The former study compares horizontal winds in WRF against radiosondes: WRF exhibits a root mean squared error (RMSE) of 2.5 – 4 m/s with no change in error statistics at the top of the planetary boundary layer. The latter study compares the performance of several particle trajectory models, including STILT, coupled with a variety of wind fields. WRF fields produce the best results among the examined wind fields when tested against controlled tracer release experiments (*Hegarty et al.*, 2013).

Section S3 describes the land surface model used in WRF.

S1.2 Nested meteorology setup

The WRF-fields used in this study are nested: 10-km resolution wind fields drive particle trajectories over most of the continental US and much of Canada while 40-km wind fields drive

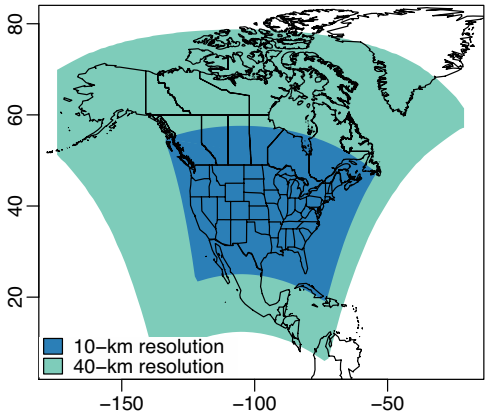

Figure S1: The 10-km and 40-km resolution domains used in the WRF model runs.

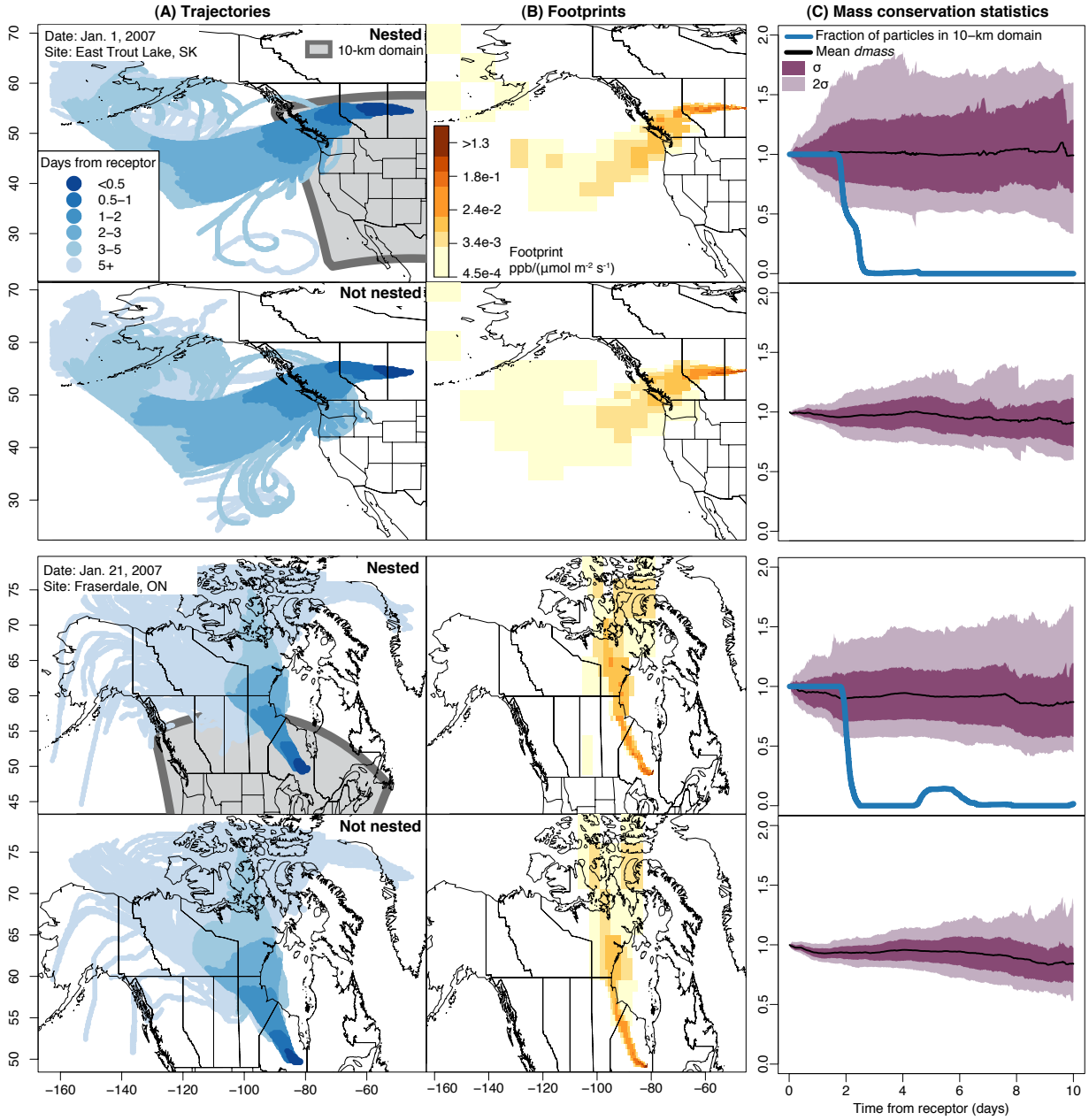

Figure S2: A comparison of STILT trajectories run with nested-resolution WRF fields (odd rows of the figure) and non-nested 40-km fields (even rows). Column (A) shows the trajectories associated with single methane measurements at East Trout Lake, Saskatchewan, on Jan. 1, 2007, and at Fraserdale, Ontario, on Jan. 21, 2007. Each ensemble of trajectories contains 500 particles that run backward in time along the WRF wind fields. Column (B) displays the footprints associated with each set of trajectories (used to construct $\mathbf{H}$). Only particles within the mixed layer have a non-zero footprint, so the footprints have a less dispersed spatial distribution compared to the trajectories. Column (C) plots *dmass*, a cumulative metric of mass conservation (refer to section S1.2). Panels show both the mean and distribution of *dmass* at each time step. *dmass* values reach as high as 50 and as low as zero when STILT is paired with other meteorological products (section S1.2).

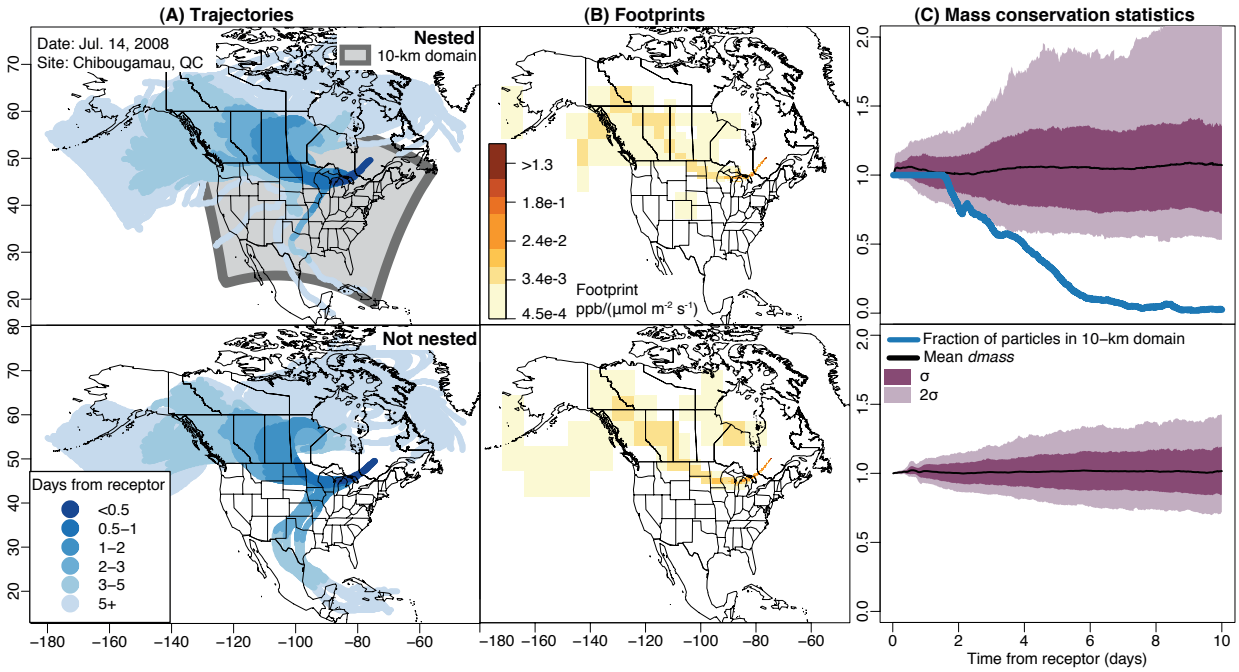

Figure S3: This figure is identical to Fig. S2 but compares trajectories from Chibougamau, Quebec, on July 14, 2008.

trajectories in more distant regions (Fig. S1). This nested setup affords a detailed description of winds within 24–48 hours of the measurement location, and the 40-km resolution at greater distances allows substantial computational savings. This arrangement makes it computationally feasible to run an ensemble of 500 back-trajectories for each of ∼15,000 hourly methane measurements. Nested WRF fields have been employed in a number of STILT modeling studies (e.g. *Zhao et al.*, 2009; *Nehrkorn et al.*, 2010; *Kort et al.*, 2010; *Gourdji et al.*, 2010; *Huntzinger et al.*, 2011; *Pillai et al.*, 2011; *Gourdji et al.*, 2012; *Jeong et al.*, 2012; *McKain et al.*, 2012; *Pillai et al.*, 2012; *Xiang et al.*, 2013; *Miller et al.*, 2013, 2014; *Hegarty et al.*, 2013; *Newman et al.*, 2013; *Shiga et al.*, 2013).

Figures S2 and S3 compare sample STILT trajectories created with the nested WRF fields against trajectories created with the 40-km fields only. The examples in these figures span winter to summer seasons and represent different synoptic transport patterns. Individual panels display the back-trajectories (Figs. S2a and S3a), the influence footprints (used to construct **H**, Figs. S2b and S3b), and a metric of mass conservation (*dmass*, Figs. S2c and S3c). For the nested WRF setup, transport is continuous and smooth at the 10-km/40-km wind field boundaries. Furthermore, synoptic-scale transport patterns are consistent in nested and non-nested model runs.

Both nested and non-nested WRF-STILT setups display favorable mass conservation statistics. The *dmass* parameter (Figs. S2c and S3c) provides an estimate of cumulative mass creation or loss in each back-trajectory. A value of one indicates perfect mass conservation while 0.5 indicates 50% loss, and 1.5 indicates 50% mass gain (*Nehrkorn et al.*, 2010). STILT uses the *dmass* parameter to correct for mass violation in the footprints. In general, the 40-km wind fields produce less mass creation/loss compared to the 10-km wind fields. *Hegarty et al.* (2013) report that the 10-km resolution fields nonetheless reproduce atmospheric transport better than coarser-resolution winds. In the nested runs, mass conservation statistics are also continuous at the wind field boundaries. In comparison, other meteorological products like EDAS-40 (Eta Data Assimilation System) and FNL (Final) produce *dmass* values as high as 50 and as low as

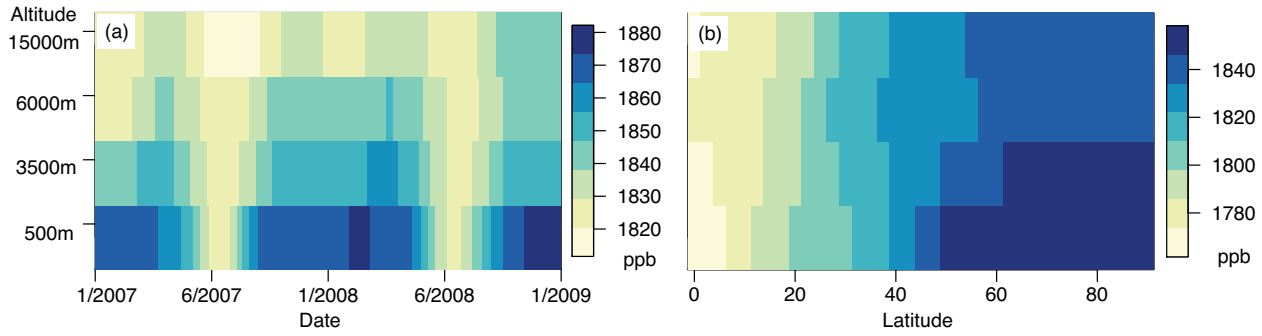

Figure S4: Examples of the methane boundary condition (a) at 50°N and a variety of dates and altitudes and (b) on August 1, 2008, and a variety of latitudes and altitudes. Note the different scales on each panel.

zero (*Miller*, 2007; *Nehrkorn et al.*, 2010).

Note that we use nested wind fields exclusively for computing the back-trajectories. The Kaplan wetland flux model uses many inputs from WRF (e.g., soil moisture, soil temperature), and these inputs are drawn exclusively from the 40-km fields.

S1.3 Boundary condition

WRF-STILT only models methane sources over the North American continent and therefore requires a methane boundary condition (see sections 2.2 in the main article). The boundary condition represents the mixing ratio of methane in air masses before they reach the continent, from both westerly and northerly synoptic air flow.

This study uses an empirical boundary that interpolates a variety of trace gas measurements from ground-based sites and aircraft in the NOAA ESRL Global Monitoring Division's Cooperative Global Air Sampling Network. The interpolated boundary curtain includes methane measurements from Greenland, Alaska, California, Hawaii, and from ship tracks in the Pacific Ocean, among other sites. The measurements are interpolated latitudinally and vertically using the a curve-fitting procedure (*Thoning et al.*, 1989). Fig. S4 displays example cross-sections of the boundary condition by date (Fig. S4a) and latitude (Fig. S4b).

The estimated boundary value associated with each trajectory run depends on the ending latitudes, altitudes, and days of the trajectories.

For most model-data comparisons in the main article, we have subtracted the estimated boundary value from the measurements to clearly depict the effect of North American methane sources on the measurement sites (e.g., Figs. 2 and 4 in the main article). The full mixing ratio measurements show large seasonality (∼40ppb) unrelated to emissions, due to changes northern hemisphere hydroxyl radical concentrations. A time series plot of full mixing ratio measurements would visually obscure the effect of North American methane fluxes on the measurements.

Existing studies validate this empirical boundary condition against methane data from aircraft in the North American free troposphere. *Miller et al.* (2013) compare the boundary condition against NOAA aircraft data over the continental US (which, in some cases, were also used to construct the boundary condition). They find that the boundary has small regional and/or seasonal biases of 0.75 to 7ppb compared to this aircraft data. *Miller et al.* (2013) further adjust the empirical boundary condition to remove these regional and/or seasonal biases compared to aircraft data. However, regular methane data from aircraft are sparse over Canada, so it is not feasible in this study to fit the methane boundary condition to regional free troposphere values.

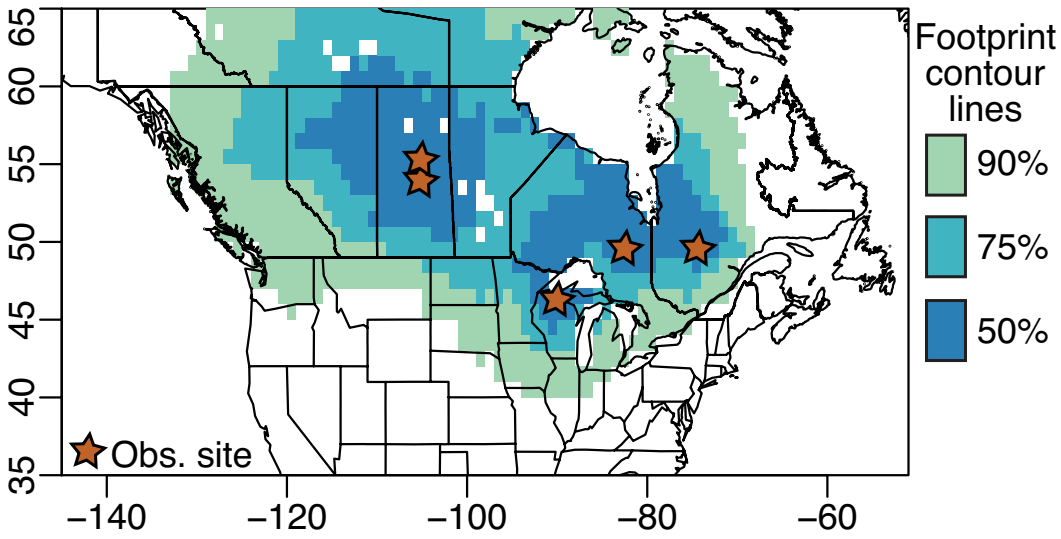

Figure S5: This figure visualizes the sensitivity of the observations to emissions in different regions. The contour lines indicate the regions that encompass 50, 75, and 90% of the total sensitivity or footprint of the observations. In other words, the figure provides a measure of geographic coverage of the inversion.

S1.4 Consistency between WRF-STILT and other atmospheric models

A number of existing STILT studies produce surface flux estimates that are comparable to studies with different transport models and/or meteorologies. These comparisons support model transport as represented by WRF-STILT and suggest a lack of overall bias in the WRF-STILT footprints (**H**). *Miller et al.* (2012) use both WRF and the Regional Atmospheric Modeling System (RAMS) with STILT to estimate US nitrous oxide emissions; the US budgets match within 12 ± 6%. Furthermore, STILT studies of carbon monoxide and the present methane study both produce budgets comparable to top-down emissions estimates with the Geos-Chem model. Constraints on summertime US carbon monoxide emissions with RAMS-STILT and Geos-Chem match to within 10% (*Hudman et al.*, 2008; *Miller et al.*, 2008), and methane budgets for the Hudson Bay Lowlands (HBL) in this study match *Pickett-Heaps et al.* (2011) to within ∼5%.

S1.5 Geographic coverage and data limitations

Fig. S5 displays the summed footprints (**H**, units of mixing ratio per flux) for the two year (2007–2008) study period. The footprints indicate the sensitivity of the measurements to fluxes in different geographic regions. This figure displays the contour lines of the footprints: the regions that encompass 50, 75, and 90% of the total sensitivity. The footprints show broad coverage across Canada. Observations are particularly sensitive to fluxes in the Hudson Bay and Great Lakes regions as well as fluxes in Saskatchewan and eastern Alberta.

Three additional Canadian measurement sites (Lac Labiche, Alberta; Egbert, Ontario; and Sable Island, Nova Scotia) are excluded from the analysis. The Lac Labiche and Egbert sites lie relatively close to the ground (10m and 6m agl, respectively) and are therefore particularly sensitive to near-surface mixing or unresolved turbulent eddies. Sable Island lies far from methane sources and sees mostly marine background air.

Fig. S5 indicates several gaps in the 2007–2008 observation network. The network has limited coverage across Arctic Canada, in eastern Quebec and maritime Canada, and over the southern half of Alberta.

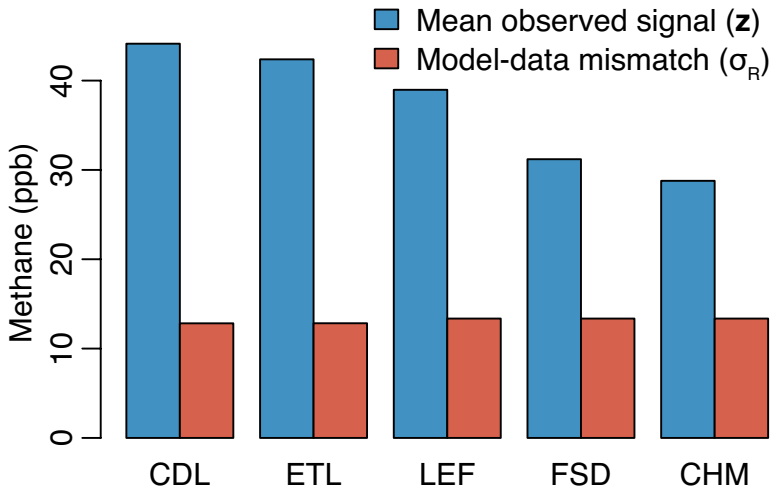

Figure S6: Blue bars illustrate the mean methane signal from North American sources at each tower site ($\boldsymbol{z}$, the measured mixing ratio minus the boundary condition). The red bars depict the estimated model-data mismatch errors (e.g., the square root of the diagonal of $\mathbf{R}$). These errors encompass uncertainties unrelated to the emissions (e.g., transport, methane boundary condition, etc.).

S2 Detailed statistical methodology

S2.1 Covariance matrix structure and estimation

We estimate the parameters ($\boldsymbol{\theta}$) that define the covariance matrices ($\mathbf{R}$ and $\mathbf{Q}$) using a method known as restricted maximum likelihood (REML). Maximum likelihood methods are standard estimation tools for statistical parameters, including variances and covariances (e.g., *Devore*, 2012, ch. 6.2). Restricted Maximum Likelihood makes it possible to estimate the variance and/or covariance when the mean, and any other drift coefficients ($\boldsymbol{\beta}$), are unknown, by first removing the effect of these unknown variables (*Corbeil and Searle*, 1976). REML has been used for covariance parameter estimation in the solution of a variety of inverse problems (e.g., *Kitanidis*, 1995), including the estimation of atmospheric trace gas sources and sinks (e.g., *Michalak et al.*, 2004; *Gourdji et al.*, 2012).

REML will estimate the parameters ($\boldsymbol{\theta}$) that are most likely given both the measurements ($\boldsymbol{z}$) and atmospheric model ($\mathbf{H}$) (e.g., *Michalak et al.*, 2004):

$$\begin{aligned} L_{\boldsymbol{\theta}} &= -\ln \int_{\boldsymbol{\beta}} \int_{\boldsymbol{s}} p(\boldsymbol{s}, \boldsymbol{\beta}, \boldsymbol{\theta} | \boldsymbol{z}, \mathbf{H}) \, d\boldsymbol{s} \, d\boldsymbol{\beta} \\ &= \tfrac{1}{2} \ln |\boldsymbol{\Psi}| + \tfrac{1}{2} \ln |\mathbf{X}^T \mathbf{H}^T \boldsymbol{\Psi}^{-1} \mathbf{H} \mathbf{X}| + \tfrac{1}{2} \boldsymbol{z}^T \boldsymbol{\Xi} \boldsymbol{z} \end{aligned} \tag{S1}$$

where

$$\boldsymbol{\Psi} = \mathbf{H} \mathbf{Q} \mathbf{H}^T + \mathbf{R} \tag{S2}$$

$$\boldsymbol{\Xi} = \boldsymbol{\Psi}^{-1} - \boldsymbol{\Psi}^{-1} \mathbf{H} \mathbf{X} (\mathbf{X}^T \mathbf{H}^T \boldsymbol{\Psi}^{-1} \mathbf{H} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{H}^T \boldsymbol{\Psi}^{-1} \tag{S3}$$

The best estimate of the covariance matrix parameters ($\boldsymbol{\theta}$) is obtained by minimizing the negative log likelihood of the observations ($L_{\boldsymbol{\theta}}$ in Eq. S1). This is equivalent to maximizing the probability $p(\boldsymbol{s}, \boldsymbol{\beta}, \boldsymbol{\theta} | \boldsymbol{z}, \mathbf{H})$ after integrating out the unknown fluxes ($\boldsymbol{s}$) and coefficients ($\boldsymbol{\beta}$).

For the inversion setup here, the model-mismatch covariance matrix ($\mathbf{R}$) is diagonal. This setup is consistent with a number of previous inversion studies that use tower-based trace gas data (e.g., *Zhao et al.*, 2009; *Bergamaschi et al.*, 2010; *Jeong et al.*, 2013; *Gourdji et al.*, 2012).

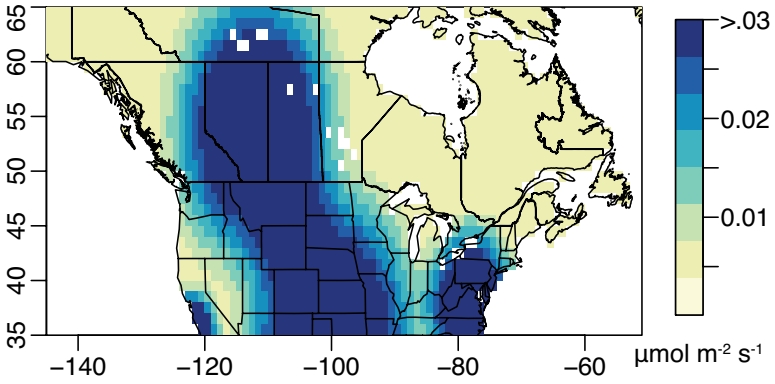

Figure S7: The estimated standard deviation of the difference between the deterministic model and the true unknown fluxes ($\hat{\sigma}_Q$). This is synonymous with the standard deviation of the stochastic component, all fluxes not described by the deterministic model. We estimate these quantities using REML (section S2.1).

We use REML to estimate a different model-data mismatch variance (σ_R^2) for the eastern (LEF, CHM, FSD) and western (CDL, ETL) tower sites. Furthermore, we model this variance to be constant across all seasons. Fig. S6 compares the estimated model-data mismatch values at each measurement site against the average measured signal from North American sources ($\boldsymbol{z}$).

The a priori covariance matrix ($\mathbf{Q}$) describes the differences between the deterministic model and the true fluxes. It requires a more complex structure for several reasons. First, the deterministic model may be a better (or worse) fit over wetland flux regions than in regions dominated by anthropogenic emissions. Second, the deterministic model fits wetland fluxes in eastern Canada better than in western Canada (see Fig. 4 in the main article). Finally, the deterministic model may describe wetland fluxes more adeptly in some seasons than in others (e.g., Fig. 4 in the main article). For these three reasons, the inversion problem requires a covariance function that is both spatially and temporally non-stationary. To this end, we model the diagonal elements of $\mathbf{Q}$, denoted σ_Q^2, as follows:

$$\sigma_Q \;=\; \alpha_1 \;+\; \alpha_2[\text{monthly Kaplan HBL budget}] \;+\; \alpha_3[\text{smooth functions over anthropogenic regions}] \quad \text{(S4)}$$

where α_1, α_2, and α_3 are constants estimated by REML. In other words, the covariance function has a component that is spatially and temporally constant (α_1), that varies seasonally with the magnitude of wetland fluxes (α_2[monthly Kaplan HBL budget]), and that varies spatially (α_3[smooth functions...]). Figure S7 displays the annual mean σ_Q as estimated by REML.

The covariances decay exponentially with distance:

$$Q_{ij}(h_{ij}|\sigma_{Qi}, \sigma_{Qj}, l) = \sigma_{Qi}\sigma_{Qj} \exp\left(-\frac{h_{ij}}{l}\right) \quad \text{(S5)}$$

where h_{ij} is the great circle distance between locations i and j, l is the decorrelation lengthscale, and σ_{Qi} and σ_{Qj} are the standard deviations given in Eq. S4 at locations i and j.

REML estimates a total decorrelation length ($3l$) of 381 ± 36km. REML did not converge on temporal decorrelation length, and we omit temporal correlations in $\mathbf{Q}$.

In summary, we implement REML to estimate six parameters ($\boldsymbol{\theta}$) that help define the covariance matrices ($\mathbf{R}$ and $\mathbf{Q}$): α_1, α_2, α_3, l, and two values of σ_R.

S2.2 The Bayesian Information Criterion

This study uses the Bayesian Information Criterion (BIC) to select predictors for the deterministic model. We calculate BIC scores using the following equations (as presented in *Gourdji et al.*, 2012):

$$BIC \;=\; \ln|\boldsymbol{\Psi}| + \boldsymbol{z}^T\boldsymbol{\Xi}\boldsymbol{z} + p\ln n \tag{S6}$$

The final BIC score for each candidate model is unit-less. This score increases with the magnitude of the combined covariance matrices ($\boldsymbol{\Psi}$), with the weighted sum of squared residuals ($\boldsymbol{z}^T\boldsymbol{\Xi}\boldsymbol{z}$), and with the complexity of the deterministic model ($p\ln n$).

We add a further constraint that none of the environmental datasets in $\mathbf{X}$ can contribute negatively to the methane fluxes. In other words, it would be illogical for either an anthropogenic or wetland emissions model to have a net negative flux. Therefore, we check the estimated coefficients $\hat{\boldsymbol{\beta}}$ on the selected model to ensure they are positive:

$$\hat{\boldsymbol{\beta}} \;=\; (\mathbf{X}^T\mathbf{H}^T\boldsymbol{\Psi}^{-1}\mathbf{H}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{H}^T\boldsymbol{\Psi}^{-1}\boldsymbol{z} \tag{S7}$$

This equation calculates the coefficients that are most likely given the atmospheric data $\boldsymbol{z}$ and $\mathbf{H}$, analogous to a weighted least squares regression.

S2.3 Estimation of the posterior fluxes

A standard atmospheric inversion setup will minimize the cost function but is oblivious to any known physical bounds on the emissions. However, large negative methane fluxes would be unrealistic; methane has a small soil sink (∼4% of the global sink), and only ∼0.5% of the global atmospheric methane loss occurs in boreal soils (*Dutaur and Verchot*, 2007). Hence, we assume that all fluxes estimated by the inversion should be nonnegative. Of further concern, unrealistic negative fluxes can exaggerate source strength in adjacent, high emissions regions (*Miller et al.*, 2014). We implement nonnegativity in the inversion with Lagrange multipliers via an iterative trust region algorithm. The companion paper *Miller et al.* (2014) describes this approach in detail.

S2.4 Uncertainties in the flux estimate

The posterior error covariance matrices provides a measure of uncertainty in both the estimated fluxes ($\hat{\boldsymbol{s}}$) and coefficients ($\hat{\boldsymbol{\beta}}$) (e.g., *Michalak et al.*, 2004):

$$\begin{bmatrix} \mathbf{V}_{\hat{\boldsymbol{s}}} & \mathbf{V}_{\hat{\boldsymbol{s}}\hat{\boldsymbol{\beta}}} \\ \mathbf{V}_{\hat{\boldsymbol{\beta}}\hat{\boldsymbol{s}}} & \mathbf{V}_{\hat{\boldsymbol{\beta}}} \end{bmatrix} \;=\; \begin{bmatrix} \mathbf{Q}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} & \mathbf{Q}^{-1}\mathbf{X} \\ \mathbf{X}^T\mathbf{Q}^{-1} & \mathbf{X}^T\mathbf{Q}^{-1}\mathbf{X} \end{bmatrix}^{-1} \tag{S8}$$

Michalak et al. (2004) and *Miller et al.* (2014) discusses posterior covariances in greater detail, and the former paper includes an alternative formulation of Eq. S8 that does not require calculating the inverse of $\mathbf{Q}$.

In this paper, we use the posterior covariance matrices to calculate all uncertainty estimates. The variances and covariances of the estimated fluxes ($\hat{\boldsymbol{s}}$) account for uncertainties in the drift coefficients and in the stochastic component of the emissions, and they account for uncertainty due to randomly-distributed model-data mismatch errors. Note that the uncertainties listed for the coefficients and posterior fluxes in this study are 2σ, unless otherwise noted.

Of note, existing statistical inversions cannot explicitly account for model-data mismatch errors that produce overall bias (e.g., biased mixed layer heights or biased boundary condition estimates). The consistency between STILT-based flux estimates and other model studies (Section S1.4) implies an absence of large or flagrant systematic errors in WRF-STILT.

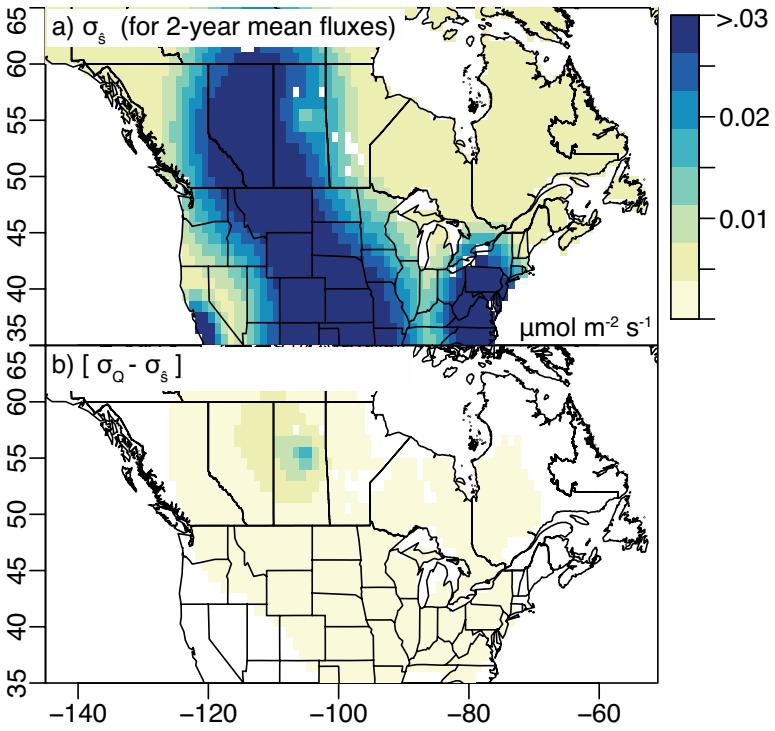

Figure S8: (a) Uncertainty in the posterior estimate ($\sigma_{\hat{\boldsymbol{s}}}$). (b) The reduction in uncertainty relative to the prior estimate ($\sigma_Q - \sigma_{\hat{\boldsymbol{s}}}$). Note that these uncertainties decrease over increasing spatial scales. Mathematically, this decrease is due to negative covariances in the posterior covariance matrix.

Furthermore, the posterior covariances do not account for any uncertainty in the choice of covariates for the deterministic model. If there were many plausible candidate models for $\mathbf{X}$, all with similar BIC scores, this could be an important component of the uncertainty in the flux estimate. In this study, most candidate models have BIC scores that are more than 10 points above the optimal model. This indicates a 'very strong' level of evidence against those models (see section 3.2.2 in the main article and *Kass and Raftery* (1995)). A notable exception is the decision to exclude soil carbon from the deterministic model. The evidence against including soil carbon is 'positive' but not 'strong' or 'very strong.' In other words, the possibility of including soil carbon within $\mathbf{X}$ and its effect on the final flux estimate ($\hat{\boldsymbol{s}}$) is a source of uncertainty not explicitly reflected in the posterior covariance matrices. The choice of covariate over Alberta represents an additional possible uncertainty. We use a set of smooth functions to describe anthropogenic emissions over Alberta (Section S4). We could not find a spatial pattern that fit the atmospheric methane data more adeptly. However, a better spatial predictor of Alberta emissions could exist.

Figure S8a displays a map of the posterior uncertainty, denoted $\sigma_{\hat{\boldsymbol{s}}}$, and Fig. S8b displays the corresponding decline from the prior uncertainties. It is important to note that the posterior uncertainties often decrease exponentially at increasing spatial scales (e.g. *Miller et al.*, 2013). This decrease occurs because of negative posterior covariances among emissions points distributed geographically. In other words, the atmospheric data data may provide a weak constraint on the exact distribution of emissions between any two adjacent grid boxes. However, the data may nonetheless constrain regional and national methane budgets with high relative confidence. For example, the uncertainties in Fig. S8a relative to the flux estimate can be 100%, but the relative uncertainty in the total Canadian methane budget (Fig. 7 of the main article) is far smaller.

The posterior uncertainties indicate key information about methane emissions in Alberta. These uncertainties at grid-scale are comparable to the magnitude of emissions. This result indicates that the precise spatial distribution of the posterior emissions over Alberta is highly uncertain. The inversion places large emissions in Northern Alberta and near the Rocky Mountain Front Range, but this assignment is poorly constrained by the available methane data. As explained above, the atmospheric methane data better constrain regional-scale totals. For example, the inversion estimates a budget for western Canada of 11.3 ± 2.4 (west of -100° lon.). The uncertainties in this regional-scale total are small relative to uncertainties at grid-scale.

The spatial distribution of emissions over Alberta remains an important, open question. Two efforts, in particular, could reduce these uncertainties in future estimates. First, expansion of the Environment Canada measurement network will provide stronger constraints on both the magnitude and spatial distribution of emissions. The 2014 network includes methane measurements at Abbotsford, British Columbia; Esther, Alberta; Fort McKay, Alberta, and Bratt's Lake, Saskatchewan (in addition to the sites used in this study). Second, top-down emissions estimates would benefit from accurate, detailed maps of spatial processes related to methane emissions. These maps could be used to construct a more capable deterministic model. The better the deterministic model describes the atmospheric methane measurements, the smaller the variance of the stochastic component, and the smaller the posterior variances and uncertainties.

S3 Detailed comparison of biogeochemical model drivers

This section provides visual comparison of the wetland-related environmental datasets that are used in Kaplan, DLEM, and the deterministic model.

Figure S9 compares all of the wetland-related variables that are listed in Table 1 of the main article. The top panels display soil moisture and temperature estimates from WRF and NARR. These meteorologies are similar for two reasons; NARR serves as the initial condition for the WRF runs, and both NARR and WRF use the Noah land surface model. The Noah model is described in a multitude of publications, and a list of relevant references is available in *Ek et al.* (2003) and *Niu et al.* (2011). Fig. S10 further illustrates the soil variables by season, in this case for NARR. Total soil moisture peaks in spring and declines into the summer. Temporal patterns in unfrozen soil moisture, by contrast, are dominated by the seasonal freeze-thaw cycle.

It is important to note that the NARR and WRF outputs used here are constrained by meteorological observations. For example, the NARR model-data assimilation framework adjusts land surface model output at each time step to match available meteorological observations, including temperature and precipitation. *Mesinger et al.* (2006) provide a succinct overview of NARR and visually compare precipitation measurements against NARR estimates. The WRF runs, in contrast, are re-initiated at regular intervals with initial conditions from NARR. The WRF runs do not assimilate meteorological observations directly, but this setup ensures that the WRF runs cannot stray far from the available meteorological data.

Subsequent panels of Fig. S9 display the wetland and soil carbon estimates. Wetland estimates include output from the LPJ model and GIEMS. GIEMS is a monthly-resolution satellite product that estimates global surface water (*Prigent et al.*, 2007; *Papa et al.*, 2010). In this study, we use long term, summer-averaged surface water from GIEMS as a wetland proxy (1993–2007, averaged over July–September). As discussed in *Melton et al.* (2013), surface water may not necessarily be an ideal proxy for wetlands. Among other issues, this proxy would underestimate wetlands that hold the majority of water in the subsurface. For this and other possible reasons, GIEMS and the LPJ model have different spatial distributions. For example, GIEMS estimates high surface water over northern Quebec while the LPJ wetland estimate is largely restricted to the HBL and Great Slave Lake regions.

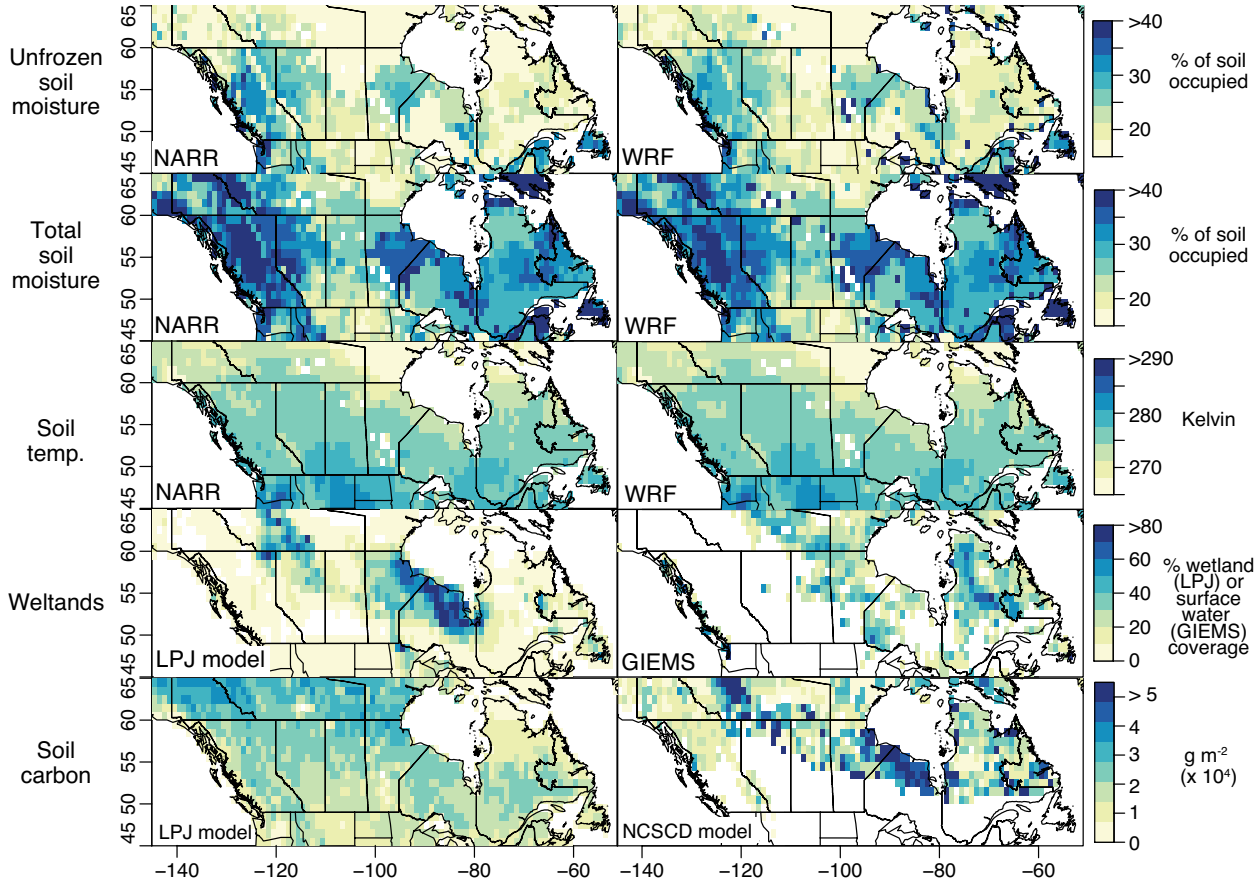

Figure S9: Environmental datasets tested for the deterministic model (Table 1 in the main article). NARR and WRF plots are annual averages.

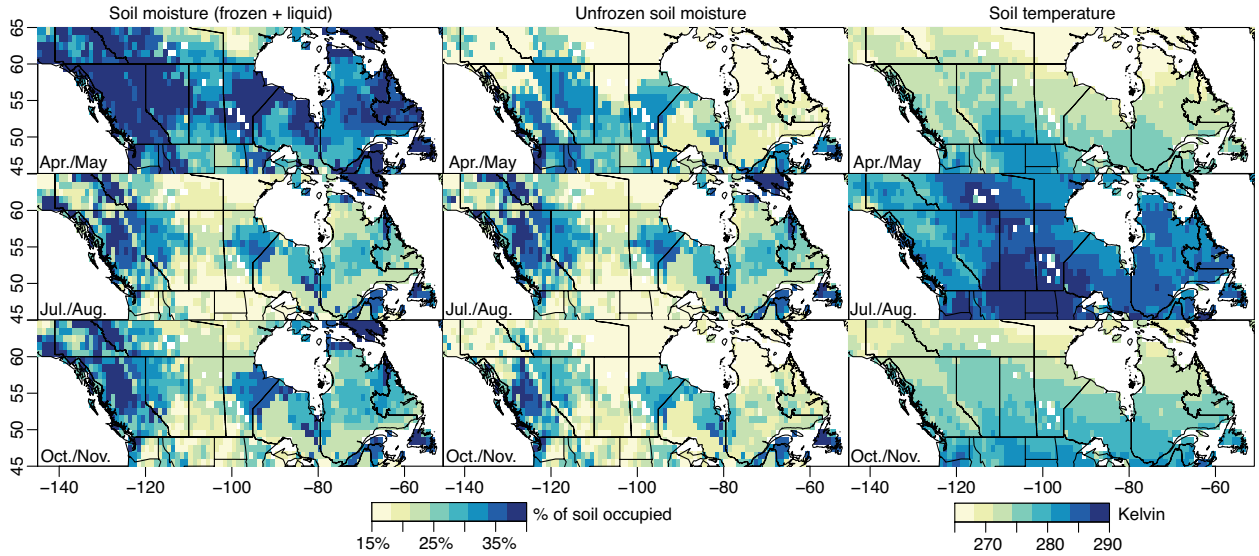

Figure S10: Soil variables from NARR for selected months of the year (averaged over 2007 – 2008).

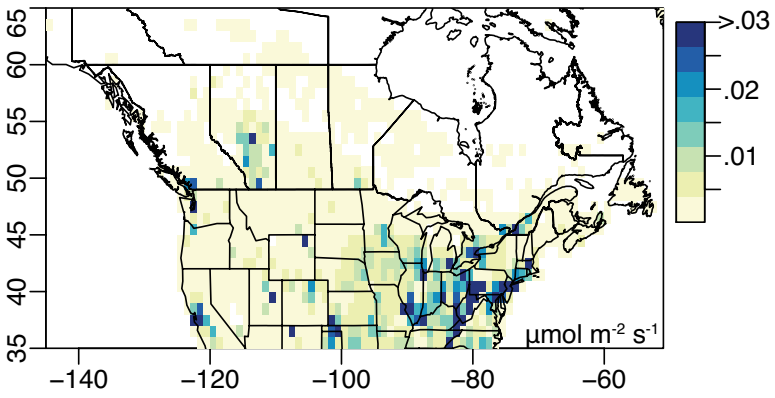

Figure S11: The EDGAR v4.2 anthropogenic methane emissions inventory.

The bottom panels of Fig. S9 visualize the soil carbon estimates, and these estimates exhibit very different spatial distributions. The LPJ soil carbon estimate correlates closely with latitude. Note that this estimate is multiplied by a factor of 4.15 before it is used in the Kaplan model in order to match the LPJ inputs used in *Pickett-Heaps et al.* (2011) (refer to section 2.4.1 of the main article). The NCSCD soil carbon estimate (*Tarnocai et al.*, 2009; *Hugelius et al.*, 2013), in contrast, has a spatial distribution more similar to the LPJ wetland estimate. NCSCD notably only estimates soil carbon in permafrost regions.

The deterministic model, described in the main article (section 4.2), excludes any estimate of soil carbon. This feature may be reconcilable with existing methane models. One might expect wetland location to correlate with soil carbon, so both environmental datasets may not be strictly necessary in a wetland flux model. In the Northern Circumpolar Carbon Database (*Tarnocai et al.*, 2009), for example, wet peatland soils (histels and histosols) have the highest carbon density of any soil type by a factor of $2-20$. Other environmental factors may complicate this correlation between the wetland distribution and soil carbon; both the type of plant carbon and age since deposition can affect the rate of methane production (*Chanton et al.*, 2008; *Bridgham et al.*, 2013). However, existing soil carbon datasets disagree even on the continental-scale distribution of soil carbon, so it is doubtful that an existing dataset would capture more subtle features or effects. Furthermore, several chamber studies argue that methane fluxes correlate more strongly with other environmental factors than with soil carbon (*Prater et al.*, 2007; *Olefeldt et al.*, 2013).

S4 Anthropogenic methane estimates

Figure S11 displays the EDGAR version 4.2 inventory over Canada. In the figure, the inventory is regridded to a 1° lat. by 1° lon. resolution for easier visual comparison with the inversion results in this study.

We use an alternative formulation of anthropogenic sources in the deterministic model. Specifically, we place smooth tricube functions over anthropogenic source regions (Fig S12):

$$f(d_i, r) = \left(1 - \left(\frac{d_i}{r}\right)^3\right)^3 \tag{S9}$$

$$d_i = \sqrt{(\text{latitude}_i - \text{latitude}_c)^2 + (\text{longitude}_i - \text{longitude}_c)^2} \tag{S10}$$

where $f()$ is the tricube function. d_i is the latitude-longitude distance between location i and the center c of the function. Both the function center (c) and radius (r) are defined by the modeler.

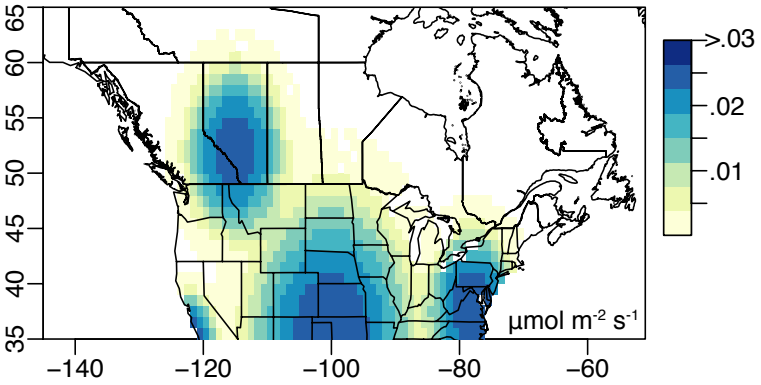

Figure S12: The set of smooth, tri-cube functions used as a component of the deterministic model ($\mathbf{X}\boldsymbol{\beta}$). In this plot, the functions are scaled by a coefficient estimated from the atmospheric methane data ($\hat{\beta}_1$, section 4.2 of the main article). Note that the Canadian observation network is not sensitive to methane emissions over most of the continental US (e.g., Fig. S5). Instead, refer to *Kort et al.* (2008, 2010) or *Miller et al.* (2013) for top-down analysis of US methane emissions.

We define the distance d_i in terms of latitude and longitude instead of great circle distance in order to better fit the shape of anthropogenic source regions. Most large anthropogenic source regions in North America span longer north-south distances than east-west (e.g., the US Eastern Seaboard, California Central Valley, and resource extraction regions of Alberta). To this end, Eq. S10 approximates anthropogenic source regions over North America with an ellipse that has been elongated in the north-south direction.

We test multiple tricube functions with different centers (c) and radii (r) and choose the combination of functions that produce the lowest BIC score. The functions chosen for use in the deterministic model are centered over Alberta (52° lat., -115° lon.), the southern Great Plains (35° lat., -100° lon.), the US East Coast (38° lat., -78° lon.), and California (35° lat., -122° lon.) (Fig. S12).

S5 Comparison with previous observational studies

The HBL methane budget estimated in this paper is higher than two previous observational studies: *Roulet et al.* (1992) and *Worthy et al.* (2000). Several factors may explain this difference. *Roulet et al.* (1992) extrapolated a flux estimate using ground-based measurement sites at the north and south ends of the HBL. The authors may have missed larger emissions in the central HBL, leading to a possible underestimate for the region. This point is also discussed in *Pickett-Heaps et al.* (2011). *Worthy et al.* (2000) used Fraserdale (FSD) observations and a box model with spatially-uniform flux to estimate an HBL budget. Their budget estimated using atmospheric data depends on the presumed distribution of the fluxes. For example, a modest source near the tower site or a much larger source in more distant regions could produce similar modeled concentrations at Fraserdale. Our study overcomes this distributional challenge by using multiple tower sites and by leveraging existing environmental datasets relevant to the distribution of wetland fluxes.

References

Bergamaschi, P., et al. (2010), Inverse modeling of European CH_4 emissions 2001-2006, *J. Geophys. Res.: Atmos.*, *115*, doi:10.1029/2010JD014180.

Bridgham, 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 Biology*, *19*(5), 1325–1346, doi:10.1111/gcb.12131.

Chanton, J. P., P. H. Glaser, L. S. Chasar, D. J. Burdige, M. E. Hines, D. I. Siegel, L. B. Tremblay, and W. T. Cooper (2008), Radiocarbon evidence for the importance of surface vegetation on fermentation and methanogenesis in contrasting types of boreal peatlands, *Global Biogeochem. Cycles*, *22*(4), doi:10.1029/2008GB003274.

Corbeil, R. R., and S. R. Searle (1976), Restricted maximum likelihood (REML) estimation of variance components in the mixed model, *Technometrics*, *18*(1), pp. 31–38.

Devore, J. (2012), *Probability and Statistics for Engineering and the Sciences*, Brooks/Cole, Cengage Learning, Boston.

Dutaur, L., and L. V. Verchot (2007), A global inventory of the soil CH_4 sink, *Global Biogeochem. Cycles*, *21*(4), doi:10.1029/2006GB002734.

Ek, M. B., K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley (2003), Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model, *J. Geophys. Res.: Atmos.*, *108*(D22), doi:10.1029/2002JD003296.

Gourdji, S. M., A. I. Hirsch, K. L. Mueller, V. Yadav, A. E. Andrews, and A. M. Michalak (2010), Regional-scale geostatistical inverse modeling of north american CO_2 fluxes: a synthetic data study, *Atmos. Chem. Phys.*, *10*(13), 6151–6167, doi:10.5194/acp-10-6151-2010.

Gourdji, S. M., et al. (2012), North American CO_2 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.

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. Meteor. Climatol.*, *52*(12), 2623–2637.

Hudman, R. C., et al. (2008), Biogenic versus anthropogenic sources of CO in the United States, *Geophys. Res. Lett.*, *35*(4), doi:10.1029/2007GL032393.

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 System Science Data*, *5*(1), 3–13, doi:10.5194/essd-5-3-2013.

Huntzinger, D. N., S. M. Gourdji, K. L. Mueller, and A. M. Michalak (2011), The utility of continuous atmospheric measurements for identifying biospheric CO_2 flux variability, *J. Geophys. Res.*, *116*(D6), doi:10.1029/2010JD015048.

Jeong, S., C. Zhao, A. E. Andrews, E. J. Dlugokencky, C. Sweeney, L. Bianco, J. M. Wilczak, and M. L. Fischer (2012), Seasonal variations in N_2O emissions from central California, *Geophys. Res. Lett.*, *39*(16), doi:10.1029/2012GL052307.

Jeong, S., Y.-K. Hsu, A. E. Andrews, L. Bianco, P. Vaca, J. M. Wilczak, and M. L. Fischer (2013), A multitower measurement network estimate of California's methane emissions, *J. Geophys. Res.: Atmos.*, doi:10.1002/jgrd.50854.

Kass, R., and A. Raftery (1995), Bayes Factors, *J. Am. Stat. Assoc.*, *90*(430), 773–795, doi: 10.2307/2291091.

Kitanidis, P. (1995), Quasi-linear geostatistical theory for inversing, *Water Resour. Res.*, *31*(10), 2411–2419, doi:10.1029/95WR01945.

Kort, E. A., et al. (2008), 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*(18), doi:10.1029/2008gl034031.

Kort, E. A., et al. (2010), Atmospheric constraints on 2004 emissions of methane and nitrous oxide in North America from atmospheric measurements and receptor-oriented modeling framework, *Journal of Integrative Environmental Sciences*, *7*(2), 125–133, doi: 10.1080/19438151003767483.

McKain, K., S. C. Wofsy, T. Nehrkorn, J. Eluszkiewicz, J. R. Ehleringer, and B. B. Stephens (2012), Assessment of ground-based atmospheric observations for verification of greenhouse gas emissions from an urban region, *Proc. Natl. Acad. Sci. U. S. A.*, *109*(22), 8423–8428, doi:10.1073/pnas.1116645109.

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.: Atmos.*, *109*(D14), doi:10.1029/ 2003JD004422.

Miller, S. (2007), STILT-CO: A new Lagrangian air pollution model for carbon monoxide. Undergraduate thesis, Harvard University.

Miller, S. M., et al. (2008), Sources of carbon monoxide and formaldehyde in North America determined from high-resolution atmospheric data, *Atmos. Chem. Phys.*, *8*(24), 7673–7696, doi:10.5194/acp-8-7673-2008.

Miller, S. M., et al. (2012), Regional sources of nitrous oxide over the United States: Seasonal variation and spatial distribution, *J. Geophys. Res.: Atmos.*, *117*(D6), doi:10.1029/ 2011JD016951.

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, *Geoscientific Model Development*, *7*(1), 303–315, doi:10.5194/gmd-7-303-2014.

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, *Meteorology and Atmospheric Physics*, *107*(1-2), 51–64, doi:10.1007/s00703-010-0068-x.

Newman, S., et al. (2013), Diurnal tracking of anthropogenic CO_2 emissions in the Los Angeles basin megacity during spring 2010, *Atmos. Chem. Phys.*, *13*(8), 4359–4372, doi:10.5194/acp-13-4359-2013.

Niu, G.-Y., et al. (2011), The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements, *J. Geophys. Res.: Atmos.*, *116*(D12), doi:10.1029/2010JD015139.

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 Biology*, *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.: Atmos.*, *115*(D12), doi:10.1029/2009JD012674.

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.

Pillai, D., C. Gerbig, R. Ahmadov, C. Rödenbeck, R. Kretschmer, T. Koch, R. Thompson, B. Neininger, and J. V. Lavrié (2011), High-resolution simulations of atmospheric CO_2 over complex terrain representing the Ochsenkopf mountain tall tower, *Atmos. Chem. Phys.*, *11*(15), 7445–7464, doi:10.5194/acp-11-7445-2011.

Pillai, D., C. Gerbig, R. Kretschmer, V. Beck, U. Karstens, B. Neininger, and M. Heimann (2012), Comparing Lagrangian and Eulerian models for CO_2 transport – a step towards Bayesian inverse modeling using WRF/STILT-VPRM, *Atmos. Chem. Phys.*, *12*(19), 8979–8991, doi:10.5194/acp-12-8979-2012.

Prater, J. L., J. P. Chanton, and G. J. Whiting (2007), Variation in methane production pathways associated with permafrost decomposition in collapse scar bogs of Alberta, Canada, *Global Biogeochem. Cycles*, *21*(4), doi:10.1029/2006GB002866.

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.: Atmos.*, *112*(D12), 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.

Shiga, Y. P., A. M. Michalak, S. Randolph Kawa, and R. J. Engelen (2013), In-situ CO_2 monitoring network evaluation and design: A criterion based on atmospheric CO_2 variability, *J. Geophys. Res.: Atmos.*, *118*(4), 2007–2018, doi:10.1002/jgrd.50168.

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*, doi:10.1029/2008GB003327.

Thoning, K., P. Tans, W. Komhyr, et al. (1989), Atmospheric carbon dioxide at Mauna Loa Observatory 2. analysis of the NOAA GMCC data, 1974–1985, *J. Geophys. Res*, *94*(06), 8549–8565.

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.: Atmos.*, *105*, 4031–4038.

Xiang, B., et al. (2013), Nitrous oxide (N_2O) emissions from California based on 2010 CalNex airborne measurements, *J. Geophys. Res.: Atmos.*, *118*(7), 2809–2820, doi:10.1002/jgrd.50189.

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.: Atmos.*, *114*, doi:{10.1029/2008JD011671}.