2010-2015 North American methane emissions, sectoral contributions, and trends: a high-resolution inversion of GOSAT satellite observations of atmospheric methane

Joannes D. Maasakkers[1,2], Daniel J. Jacob[1], Melissa P. Sulprizio[1], Tia R. Scarpelli[1], Hannah Nesser[1], Jianxiong Sheng[1,3], Yuzhong Zhang[1,4,5,6], Xiao Lu[1], A. Anthony Bloom[7], Kevin W. Bowman[7,8], John R. Worden[7], and Robert J. Parker[9,10]

[1]Harvard University, Cambridge, Massachusetts 02138, United States
[2]SRON Netherlands Institute for Space Research, Utrecht, The Netherlands
[3]Massachusetts Institute of Technology, Cambridge, MA, United States
[4]Environmental Defense Fund, Washington, DC, USA
[5]School of Engineering, Westlake University, Hangzhou, Zhejiang Province, China
[6]Institute of Advanced Technology, Westlake Institute for Advanced Study, Hangzhou, Zhejiang Province, China
[7]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA
[8]Joint Institute for Regional Earth System Science and Engineering, University of California, Los Angeles, CA, USA
[9]Earth Observation Science, School of Physics and Astronomy, University of Leicester, Leicester, UK
[10]NERC National Centre for Earth Observation, Leicester, UK

Correspondence: J.D. Maasakkers (j.d.maasakkers@sron.nl)

Abstract. We use 2010-2015 GOSAT satellite observations of atmospheric methane columns over North America in a high-resolution inversion of methane emissions, including contributions from different sectors and long-term trends. The inversion involves analytical solution to the Bayesian optimization problem for a Gaussian mixture model (GMM) of the emission field with up to $0.5° \times 0.625°$ resolution in concentrated source regions. Analytical solution provides a closed-form characteriza-

5 tion of the information content from the inversion and facilitates the construction of a large ensemble of solutions exploring the effect of different uncertainties and assumptions. Prior estimates for the inversion include a gridded version of the EPA Inventory of U.S. Greenhouse Gas Emissions and Sinks (GHGI) and the WetCHARTs model ensemble for wetlands. Our best estimate for mean 2010-2015 US anthropogenic emissions is 30.6 (range: 29.4-31.3) Tg a^{-1}, slightly higher than the gridded EPA inventory (28.7 (26.4-36.2) Tg a^{-1}). The main discrepancy is for the oil and gas production sectors where we find higher

10 emissions than the GHGI by 35% and 22% respectively. The most recent version of the EPA GHGI revises downward its estimate of emissions from oil production and we find that these are a factor 2 lower than our estimate. Our best estimate of US wetland emissions is 10.2 (5.6-11.1) Tg a^{-1}, on the low end of the prior WetCHARTs inventory uncertainty range (14.2 (3.3-32.4) Tg a^{-1}) and calling for better understanding of these emissions. We find an increasing trend in US anthropogenic emissions over 2010-2015 of 0.4% a^{-1}, lower than previous GOSAT-based estimates but opposite to the decrease reported

15 by the EPA GHGI. Most of this increase appears driven by unconventional oil/gas production in the eastern US. We also find that oil/gas production emissions in Mexico are higher than in the nationally reported inventory, though there is evidence for a

2010-2015 decrease in emissions from offshore oil production.

1 Introduction

Methane is the second most important greenhouse gas in terms of radiative forcing (Stocker et al., 2013). Major emission source
5   sectors include wetlands (the main natural source), livestock, the fossil fuel industry, and waste management (Kirschke et al.,
2013). Individual countries report their anthropogenic emissions to the United Nations Framework Convention on Climate
Change (UNFCCC) using methods prescribed by The Intergovernmental Panel on Climate Change (IPCC) (United Nations,
1992; IPCC, 2006). The reports use "bottom-up" methods, where activity data (e.g., heads of cattle) are combined with emis-
sion factors (e.g., emission per head of cattle) to estimate total emissions. US emissions are calculated and reported in this
10   manner by the Environmental Protection Agency (EPA) in its annual Inventory of U.S. Greenhouse Gas Emissions and Sinks
(GHGI) (EPA, 2020). Measurements of atmospheric methane, including from satellites, can be used through inverse modeling
to provide an independent comparison of these emission estimates (Streets et al., 2013; Jacob et al., 2016). Here, we evaluate
2010-2015 North American emissions by inversion of data from the Greenhouse Gases Observing Satellite (GOSAT), which
measures methane concentrations at high precision by solar backscatter in the shortwave infrared (SWIR) (Butz et al., 2011;
15   Buchwitz et al., 2015; Kuze et al., 2016). We take the gridded version of the EPA GHGI (Maasakkers et al., 2016) as prior
estimate for the inversion, enabling us to use the inversion results to evaluate the GHGI and guide improvements in its repre-
sentation of emission processes.

Several inverse studies using observations of atmospheric methane have found higher US emissions than reported in bottom-
20   up inventories. Miller et al. (2013) estimated methane emissions 50% higher than the EPA GHGI based on 2007-2008 surface
and aircraft observations. They attributed this difference to fossil fuel extraction. Wecht et al. (2014) estimated 40% higher live-
stock emissions using 2004 data from the SCIAMACHY satellite instrument (Frankenberg et al., 2006). Turner et al. (2015)
found anthropogenic emissions to be 50% higher than the EPA GHGI by inversion of 2009-2011 GOSAT data, attributing
the difference to a combination of oil/gas and livestock emissions. Janardanan et al. (2017) found 28% higher anthropogenic
25   emissions over North America based on 2009-2012 GOSAT data. All these studies used the global gridded EDGAR inventory
(European Commission, 2011; Olivier and Janssens-Maenhout, 2012; Janssens-Maenhout et al., 2019) as prior estimate, but
the EDGAR spatial distributions have large errors that affect inversion results and their interpretation (Maasakkers et al., 2016).
Alvarez et al. (2018) used surface and aircraft data in oil/gas fields to find emissions from oil/gas production 60% higher than
in the EPA GHGI.

30

There has also been substantial debate as to the contribution of North American emissions to the rising global methane trend
since 2007. Hausmann et al. (2016) proposed an increase in US gas production as explanation for the 2007-2014 methane
and ethane concentration trends at the Zugspitze mountain site in southern Germany. Turner et al. (2016) found a 2.5% a^{-1}

increase in US emissions for 2002-2014 on the basis of GOSAT and surface methane data. By contrast, Bruhwiler et al. (2017) found from an ensemble of inversions that North American emissions had been flat for 2000-2012. Sheng et al. (2018a) analyzed 2010-2016 GOSAT enhancements over North America and found a 2.5 ± 1.4 % a^{-1} increase over the US driven by oil/gas production and livestock emissions in the Midwest. They found no significant trend over Canada (but large year-to-year

5   variation driven by wetlands) and a 0.8-1.7 % a^{-1} decrease in Mexican emissions driven by a decrease in livestock. Using 2006-2015 surface and aircraft observations over the US and Canada, Lan et al. (2019) found a 0.7 ± 0.3 % a^{-1} increase in total US emissions.

Our inverse analysis of the 2010-2015 GOSAT data over North America optimizes both mean emissions and their long-term

10 trends at high resolution (up to 50 km). The inversion uses dynamic boundary conditions from a consistent global inversion of the 2010-2015 GOSAT data previously reported by Maasakkers et al. (2019). We use an analytical solution to the Bayesian inverse optimization problem (Jacob et al., 2016), which provides closed-form characterization of the information content of the solution, and also enables us to perform a range of sensitivity inversions (inversion ensemble) at no added computational cost. We relate the results from our inversion ensemble to the EPA GHGI emissions in order to inform knowledge of US emissions,

15 their trends, and the contributions from different sectors.

2   Data and methods

We perform a continental-scale inversion of 2010-2015 GOSAT methane data from the University of Leicester proxy CH$_4$ retrieval (Parker et al., 2015; ESA CCI GHG project team , 2018). We use the individual GOSAT observations to optimize a state

20 vector of mean methane emissions and linear emission trends trends at spatial resolution of up to $0.5° \times 0.625°$ ($\approx 50 \times 50$ km^2). The forward model for the inversion is the GEOS-Chem chemical transport model (www.geos-chem.org) applied in nested mode over North America with a spatial resolution of $0.5° \times 0.625°$. The main results presented here will be from a base inversion with specifications given below. In addition to this base inversion, we conducted an ensemble of 9 inversions in which we explored the sensitivity of the results to different assumptions. Specifications for these sensitivity inversions are

25 given below and summarized in Section 2.5. We use the range of results from the inversion ensemble as a range of uncertainty in the results from the base inversion.

2.1   GOSAT observations

The GOSAT satellite has been observing dry column methane mixing ratios in the SWIR using the TANSO-FTS instrument

30 since April 2009 (Butz et al., 2011). GOSAT in its default mode observes circular pixels of 10 km diameter at 13:00 local time, separated by $\sim$250 km along-track and cross-track, repeating observation on the same track every 3 days. Additional locations can be observed using the target mode. GOSAT methane retrievals have a 24% success rate, limited mainly by cloud

cover. Observations have a precision of 13 ppb and relative bias of 2 ppb compared to the Total Carbon Column Observing Network (Buchwitz et al., 2015). There is no significant spectral degradation over time (Kuze et al., 2016). Figure 1 shows the 156,110 retrievals over land used to optimize emissions in this study (Parker et al., 2015). Each retrieval comes with an esti-mated retrieval error (11 ppb on average). We use observations over land from January 2010 to December 2015, excluding data
5 above 60N° for which model errors are large (Maasakkers et al., 2019). Most observations (95,365) are over the Contiguous United States (CONUS). The data are spatially sparse but this reflects the observing strategy of repeated measurements at the same locations in the default mode. Thus most observation locations in Figure 1 have a large number of data points to inform temporal variability and trends (Sheng et al., 2018a).

10 **2.2   Prior estimates**

Table 1 summarizes our prior emissions estimates and Figure 2 shows their spatial distributions for the major sectors. For all US anthropogenic emissions including offshore as reported to the UNFCCC, we use the spatially disaggregated (gridded) ver-sion of the EPA GHGI (EPA, 2016) for 2012 from Maasakkers et al. (2016), with improved spatial allocation of emissions and detailed separation of source sectors compared to EDGAR. For oil/gas emissions in Mexico and Canada including offshore, we
15 use gridded versions of the Instituto Mexicano del Petróleo (IMP) inventory for 2010 (IMP, 2012) and the ICF International inventory for 2013 (ICF, 2015), respectively, as reported by Sheng et al. (2017). IMP (2012) oil/gas emissions for Mexico match the values reported by Mexico to UNFCCC, while ICF (2015) oil/gas emissions for Canada are 43% higher than the value reported by Canada (UNFCCC, 2019). For other anthropogenic emissions in Canada/Mexico and other countries in the domain, we use the EDGAR v4.3.2 global emission inventory for 2012 (edgar.jrc.ec.europa.eu [2017]). We separate the general
20 'fuel exploitation' sector reported by EDGAR v4.3.2 into oil/gas and coal components by using additional information em-bedded in the inventory (Greet Maenhout, personal communication). This allows us to use EDGAR coal emissions for Canada and Mexico. The emissions as given by EDGAR v4.3.2 are aseasonal. For manure management and rice cultivation we apply seasonality as prescribed by Maasakkers et al. (2016) and Zhang et al. (2016) respectively. Other anthropogenic emissions remain aseasonal.

Natural emissions are dominated by wetlands for which we use mean monthly emissions from the WetCHARTS v1.0 ex-tended ensemble with $0.5° \times 0.5°$ spatial resolution (Bloom et al., 2017). The ensemble parameters consist of: three global scaling factors (global emissions of 124.5, 166, or 207.5 Tg a^{-1}); three temperature q10 dependencies (1,2, or 3); and two landcover databases that are combined with precipitation data to estimate wetland extent (GLWD from Lehner and Döll (2004)
30 or GlobCover from Bontemps et al. (2011)). Total wetland emissions vary month to month and interannually as driven by temperature and inundation extent (Bloom et al., 2017). Because the WetCHARTs ensemble exhibits considerable but uncer-tain year-to-year variability, we also perform a sensitivity inversion without prior interannual variability in wetland and other emissions. Daily open fire emissions are from the Quick Fire Emissions Dataset (QFED) (Darmenov and da Silva, 2013) and termite emissions are from Fung et al. (1991). We use geological seepage emissions compiled from literature on both point

sources (Etiope, 2015; Kvenvolden and Rogers, 2005) and areal seepage (Kvenvolden and Rogers, 2005; Etiope and Klusman, 2010) as described in Maasakkers et al. (2019).

2.3 Forward model

5 We use the nested version of the GEOS-Chem chemical transport model v11-01 at $0.5° \times 0.625°$ resolution over North America as forward model for the inversion. Earlier versions of this model for methane were described by Wecht et al. (2014) and Turner et al. (2015). The model is driven with MERRA-2 meteorological fields (Bosilovich et al., 2016) from the NASA Global Modeling and Assimilation Office (GMAO). Methane loss from reaction with OH and Cl radicals, soil uptake, and stratospheric oxidation are described in Maasakkers et al. (2019). The simulation is initialized in January 2009 with concen-

10 tration fields from Turner et al. (2015). Three-hourly boundary conditions at the edges of the nested domain are from the $4° \times 5°$ posterior model simulation of Maasakkers et al. (2019), which provides an unbiased fit to the global GOSAT data. Methane chemical and soil sinks are not optimized in our inversion because they are very slow compared to the time scale for ventilation of the North American domain.

15 Following Maasakkers et al. (2019), we correct the GEOS-Chem simulation of GOSAT columns for a latitudinally and seasonally variable background bias likely caused by the extratropical stratosphere (Bader et al., 2016; Saad et al., 2016; Stanevich, 2018). The latitudinal correction term ξ (ppb) follows a quadratic form as in Turner et al. (2015):

$$\xi = \left(4.0\theta^2 - 1.3\theta\right) \times 10^{-3} - 5 \tag{1}$$

with θ the latitude in degrees. The seasonal bias is corrected over rolling $8°$ latitudinal bands. A sensitivity inversion without

20 the seasonal bias correction is performed as part of the inversion ensemble.

2.4 State vector for the inversion and error covariances

Although we could technically carry out the inversion of the GOSAT data at the $0.5° \times 0.625°$ resolution of the GEOS-Chem simulation, the data do not have sufficient information to constrain emissions on that grid and doing so would incur large

25 smoothing error (Wecht et al., 2014). We use instead a 600-element Gaussian mixture model (GMM) as described by Turner and Jacob (2015) to optimally define the emission patterns that can be usefully constrained by the inversion. Each of the 600 Gaussian functions in the GMM is defined by an emission amplitude, mean location, and spread (standard deviation). These parameters are optimized using a similarity vector on the $0.5° \times 0.625°$ grid that takes into account latitude, longitude, and the prior patterns of different source sectors. The state vector $\mathbf{x}$ for the inversion with dimension $n = 2 \times 600$ consists of scal-

30 ing factors adjusting the amplitudes of the Gaussians in the GMM and their 2010-2015 linear trends. This approach allows for effective aggregation of regions with weak or homogeneous emissions while preserving high resolution for concentrated emissions. Each $0.5° \times 0.625°$ grid cell is represented by a unique combination of the Gaussians, so that the optimization of $\mathbf{x}$

Table 1. Methane emissions used as prior 2010-2015 estimates[a].

Source (Tg a^{-1})	CONUS	Canada	Mexico	Other[b]
Natural	15.7	15.3	1.4	3.8
Wetlands	14.2	14.4	1.0	3.4
Open fires	0.5	0.3	0.2	0.1
Termites	0.6	0.3	0.1	0.2
Geological seeps	0.5	0.3	0.1	0
Anthropogenic	28.7	4.5	5.3	5.1
Livestock	9.2	1.0	2.5	1.9
Enteric Fermentation	6.7	0.8	2.2	1.9
Manure Management	2.5	0.2	0.3	0.1
Oil and Natural Gas	9.1	2.4	1.5	1.2
Gas production	4.4	1.2	0.1	
Oil production	2.3	0.5	1.2	
Gas Transmission	1.1	0.3	0	
Gas Processing	0.9	0.3	0.1	
Gas Distribution	0.5	0	0	
Landfills	5.8	0.7	0.4	0.6
Coal Mining	2.9	0.1	0	0
Wastewater	0.7	0.2	0.7	0.7
Rice Cultivation	0.5	0	0	0.2
Other Anthropogenic[c]	0.5	0.1	0.2	0.4
Total Source	44.5	19.8	6.7	8.9

[a] CONUS anthropogenic emissions are from the EPA GHGI for 2012 as spatially disaggregated by Maasakkers et al. (2016). Oil/gas emissions from Canada (2013) and Mexico (2010) are from ICF (2015) and IMP (2012), respectively, spatially disaggregated by Sheng et al. (2017). Other anthropogenic emissions are from EDGAR v4.3.2 for 2012 (edgar.jrc.ec.europa.eu [2017]). Wetlands and open fire emissions are mean values for 2010-2015 from the WetCHARTS ensemble (Bloom et al., 2017) and QFED (Darmenov and da Silva, 2013); Termite emissions are from Fung et al. (1991). Seepage emissions are as described in Maasakkers et al. (2019).

[b] within the inversion domain shown in Figure 1 (10-70N°, 140-40W°) containing parts of Central and South America.

[c] including fossil fuel combustion, industrial processes, agricultural field burning, and composting.

can be mapped to the 0.5° × 0.625° grid. For more details see Turner and Jacob (2015).

Prior emission error variances are defined for each Gaussian on the basis of its spatial distribution and the contributions from different sectors. Emission errors for individual anthropogenic sectors are estimated using the error curves from Maasakkers et al. (2016). The error standard deviation σ for a given source sector and Gaussian is given by:

$$\sigma = (\alpha_0 \exp(-k_\alpha(L - L_0)) + \alpha_N) E \qquad (2)$$

5   Where α_0, k_α, and α_N are source sector specific error coefficients from Maasakkers et al. (2016), L is the effective spatial resolution (length scale) of the Gaussian defined by the number of $0.5° \times 0.625°$ grid cells it represents, $L_0 = 0.1°$ is the native resolution of the prior inventory, and E is the sum of emissions from the source sector within the Gaussian (sum of emissions from $0.5° \times 0.625°$ grid cells weighted by their contributions to the Gaussian). Maasakkers et al. (2016) also include a displacement error related to uncertainty in source location but this error is negligible at our resolution. For wetland emissions,
10   we use the standard deviation in monthly estimates of the 18 WetCHARTs v1.0 extended ensemble members averaged over the Gaussian; the resulting error standard deviation is 78% on average. For the other natural emissions we assume 100% error at the $0.5° \times 0.625°$ model resolution.

     The error variances for all sectors contributing to a given Gaussian are added in quadrature to obtain the corresponding
15   diagonal element of the prior error covariance matrix $\mathbf{S_A}$. Error variances for a given Gaussian are capped at 50% in the base inversion and we also perform a sensitivity inversion without this cap. The mean relative error standard deviation is 37% in the base inversion. The 50% cap mainly affects Gaussians dominated by wetland emissions. For the 2010-2015 emission trends associated with each Gaussian, the prior estimate is set to zero and the prior error standard deviation is a 5% change per year, in line with uncertainties in trend estimates for North America (Turner et al., 2016; Bruhwiler et al., 2017; Sheng et al., 2018a;
20   Lan et al., 2019). We also perform sensitivity inversions with changes of 2.5% and 10% per year as prior error standard deviation. Off-diagonal elements of $\mathbf{S_A}$ are assumed to be zero because Maasakkers et al. (2016) found no spatial error correlation for the gridded EPA inventory; this may be an underestimate for wetland emissions (Bloom et al., 2017).

     Our calculation of $\mathbf{S_A}$ leads to different error variances for each grid cell. To assess the impact of that choice, we also
25   perform a sensitivity inversion using the mean error variance for all the Gaussians. The base inversion assumes normal errors, but we also perform a sensitivity inversion assuming log-normal emission errors following the Levenberg-Marquardt method as described in Maasakkers et al. (2019).

     We use the residual error method (Heald et al., 2004) to construct the diagonal of the observational error covariance matrix
30    $\mathbf{S_O}$. The mean 2010-2015 difference between GOSAT and the prior model (before seasonal correction) for each $0.5° \times 0.625°$ grid cell is assumed to be due to errors in emissions, to be corrected by the inversion. After subtracting this mean difference, the residual standard deviation is taken as estimate of the observational error standard deviation including contributions from instrument, representation, and forward model errors. If this estimate is less than the reported instrument error standard devia-

tion (Parker et al., 2015), we use the latter instead (17% of observations). If it is less than 10 ppb we reset it to 10 ppb (6% of observations). The resulting average observational error standard deviation is 14 ppb. Off-diagonal terms of $\mathbf{S_O}$ are assumed to be zero for lack of better information, but in fact some transport error correlation would be expected in the forward model. We account for this error correlation with a regularization term γ in the inversion (Section 2.5).

2.5 Inversion procedure

We perform an analytical inversion minimizing the Bayesian cost function $J(\mathbf{x})$ assuming normal errors (Rodgers, 2000):

$$J(\mathbf{x}) = (\mathbf{x} - \mathbf{x_A})^T \mathbf{S_A}^{-1} (\mathbf{x} - \mathbf{x_A}) + \gamma (\mathbf{y} - F(\mathbf{x}))^T \mathbf{S_O}^{-1} (\mathbf{y} - F(\mathbf{x})) \tag{3}$$

where $\mathbf{x}$ is the state vector to be optimized, consisting of 600 Gaussians for which we optimize both scaling factors for

10 mean emissions and absolute linear emission trends, for a total of 1200 state vector elements; $\mathbf{x_A}$ is the prior state vector; $\mathbf{S_A}$ is the prior error covariance matrix (Section 2.4); $\mathbf{S_O}$ is the observational error covariance matrix (Section 2.4); and γ is a regularization factor to account for the lack of non-diagonal terms in $\mathbf{S_O}$ and hence prevent overfitting. γ plays a similar role as the regularization parameter in Tikhonov methods (Brasseur and Jacob, 2017) and reflects our inability to precisely quantify error statistics in the Bayesian method. Here we find that $\gamma = 0.5$ provides the best balance of fitting the prior and observational

15 terms in the cost function, following the L-curve approach of Hansen (1999). The value is higher than $\gamma = 0.05$ used in the global inversion of Maasakkers et al. (2019) at $4° \times 5°$ resolution because here we have a smaller number of observations per state vector element. We also conduct sensitivity inversions with $\gamma = 0.1$ and $\gamma = 1$.

The GEOS-Chem forward model ($\mathbf{y} = F(\mathbf{x})$) as implemented here is strictly linear in its relationship between methane

20 column concentrations ($\mathbf{y}$) and the state vector of emissions ($\mathbf{x}$). It can be expressed as $F(\mathbf{x}) = \mathbf{Kx} + \mathbf{c}$ where $\mathbf{K} = \partial \mathbf{y}/\partial \mathbf{x}$ is the Jacobian matrix and $\mathbf{c}$ an initialization constant. This allows the optimal posterior solution $\widehat{\mathbf{x}}$ which minimizes the cost function $J(\mathbf{x})$ to be obtained analytically as:

$$\widehat{\mathbf{x}} = \mathbf{x_A} + \mathbf{S_A} \mathbf{K}^T \left(\mathbf{KS_A}\mathbf{K}^T + \frac{\mathbf{S_O}}{\gamma} \right)^{-1} (\mathbf{y} - \mathbf{Kx_A}) \tag{4}$$

with posterior error correlation matrix $\widehat{\mathbf{S}}$:

25 $$\widehat{\mathbf{S}} = \left(\gamma \mathbf{K}^T \mathbf{S_O}^{-1} \mathbf{K} + \mathbf{S_A}^{-1} \right)^{-1} \tag{5}$$

The information content from the inversion can then be obtained from the averaging kernel matrix ($\mathbf{A} = \partial \widehat{\mathbf{x}}/\partial \mathbf{x}$) which gives the sensitivity of the solution to the true state:

$$\mathbf{A} = \mathbf{I} - \widehat{\mathbf{S}}\mathbf{S_A}^{-1} \tag{6}$$

The trace of **A** gives the degrees of freedom for signal (DOFS), which measures the number of independent pieces of information on the state vector that can be obtained from the inversion. The diagonal elements of **A** (averaging kernel sensitivities) measure the degree to which the inversion can constrain the true values of the corresponding state vector elements (1 = perfectly, 0 = not at all). We will use these measures of information in our presentation of results.

Analytical solution to the inverse problem requires explicit construction of the Jacobian matrix. We perform this construction column by column by perturbing individually the 1200 elements of the state vector and conducting the corresponding GEOS-Chem simulations for the 2010-2015 observation record. This is readily done as a massively parallel calculation. Aside from enabling closed-form characterization of the information content from the inversion, a major advantage of the analytical

10 solution once the Jacobian matrix has been constructed is that the sensitivity of the solution to various assumptions and choices made in the inversion approach can be immediately obtained. In addition to our base inversion, we generate in this manner an ensemble of nine sensitivity inversions introduced in the text above, and for which the ensemble of solutions gives a better measure of posterior error than can be obtained from $\widehat{\mathbf{S}}$ (Heald et al., 2004). To summarize, these sensitivity inversions include: (1) Using a prior estimate with no interannual variability (2012 values) for wetland and biomass burning emissions, and for

15 the soil sink; (2) Not using a seasonal correction to the GOSAT - model mismatch; (3) Using emission error variances without the 50% cap; (4) Using the average emission and absolute trend error variances (37% and 2.3 Mg a^{-2} km^{-2}, respectively) for each Gaussian; (5,6) Assuming error prior standard deviations for the 2010-2015 trend of 2.5% and 10% annual change of emissions (instead of 5%); (7) Assuming log-normal prior emission errors; (8,9) Using regularization factors for the cost function $\gamma = 0.1$ and $\gamma = 1$ (instead of $\gamma = 0.5$).

3 Results and Discussion

Figure 3 shows mean prior and posterior emissions for 2010-2015, the ratio between the two, and the inversion's averaging kernel sensitivities (Equation 6). The averaging kernel sensitivities identify regions where the GOSAT observations provide significant information on emissions. These are regions with a high density of observations and/or high absolute uncertainties

25 on the prior emissions. For example, we achieve good constraints on emissions in central Canada, much of the eastern and central US, California, and southeastern Mexico. Other regions receive little information from the observations, which explains a lack of departure from the prior estimate.

The posterior emissions when implemented in GEOS-Chem reduce the mean squared difference with GOSAT by 3.5%.

30 Overall changes are small because random errors in individual observations are large and because the background is already captured well in the prior simulation through the optimized boundary conditions. We independently evaluated the posterior estimate by comparison to in situ methane concentrations from surface sites reported in the GLOBALVIEWplus CH4 ObsPack v1.0 data product compiled by NOAA Global Monitoring Laboratory (Cooperative Global Atmospheric Data Integra-

tion Project, 2019). The spatial coefficient of determination between the time-averaged GEOS-Chem and NOAA data increases from $r^2 = 0.58$ with the prior emissions to $r^2 = 0.81$ with the posterior emissions, representing an improvement in our ability to fit observed patterns.

3.1 Mean 2010-2015 emissions

Although the inversion yields little change in total emissions for the continental domain, there are large regional changes as shown in Figure 3. We find higher emissions over the south-central and eastern US, and lower emissions in California compared to the gridded EPA inventory. The WetCHARTS inventory overestimates wetland emissions including along the Gulf and east coasts of the US, the upper Midwest, and Canada. Emissions in eastern Mexico are higher than inferred from the IMP (2012) inventory. The inversion also shows large relative increases from oil production off the Louisiana coast and from wetlands/livestock in western Montana but the associated emissions are low. The large-scale correction patterns revealed by the inversion are similar to those of the coarse ($4° \times 5°$) global inversion reported by Maasakkers et al. (2019), which used the same prior estimates, but we have much more detail here.

Figure 4 shows the attribution of the inversion results to individual source sectors for CONUS, Canada, and Mexico. This attribution was made by applying the correction factors to the sectoral emissions in each grid cell, assuming that the relative contributions of individual sectors to emissions in that grid cell is correct in the prior emission inventory (this does not assume that the total prior distribution of sectoral emissions is correct). Vertical bars show the range of results from the inversion ensemble. Narrow ranges may not necessarily reflect confidence in the inversion results but the inability of the inversion to inform emissions from a particular sector when it is small.

The largest decrease is for US wetland emissions, mostly contributed by the Gulf and east coasts (Figure 2). Such an overestimate in the mean of the WetCHARTS wetland inventory ensemble was previously identified in an inversion of aircraft observations over the Southeast US (Sheng et al., 2018b). It may be related to the low organic carbon content of the soil, the difficulty of distinguishing freshwater and saltwater wetlands, uncertainties in anaerobic CH_4:CO_2 respiration rates, and the accounting of partial wetland land-cover areas (Holmquist et al., 2018; Lehner and Döll, 2004; Bloom et al., 2017). We also find an overestimate of wetland emissions in (mainly eastern) Canada. The large uncertainty range is driven by the inversion ensemble member without seasonal correction. Based on the root mean square error and spatial correlation, our inversion results are most consistent with the WetCHARTs ensemble members that use GlobCover wetland extent, a $q_{10} = 2$ value for the factor increase in the CH_4 to CO_2 emission ratio per 10 K temperature increase, and global scaling at the low end or middle of the range (global wetlands emission range 125-166 Tg a^{-1}). Sheng et al. (2018b) also found their inversion results to be most consistent with GlobCover but also favored no CH_4:CO_2 temperature dependence ($q_{10} = 1$). Their observations were much more limited in space and time (August-September 2013).

Figure 4 also shows some significant sectoral corrections for anthropogenic emissions in the US and Mexico. Over Mexico higher livestock (+13 (5-24)%) and oil/gas emissions (+22 (-24-42)%). Uncertainty ranges in correcting individual sectors are large for Mexico because of the extensive spatial overlap between sectors (Figure 2). Most of the oil/gas correction is for coastal/offshore oil production (Figure 3). We also find 56 (31-120) % higher emissions over Mexico City, which is attributed

5   to wastewater based on the EDGAR spatial patterns. Compared to EDGAR v4.3.2 the Mexico City Secretariat of Environment (SEDEMA) air quality emission inventory (SEDEMA, 2018) predicts lower emissions from wastewater (68 versus 259 Gg a^{-1}) but much higher landfill emissions (222 versus 1 Gg a^{-1}) indicating that our higher emission estimate may be related to landfill emissions being misallocated in EDGAR v4.3.2.

10   Inversion results for the US mapped on to the detailed source sectors and subsectors from the gridded EPA inventory (Maasakkers et al., 2016) are given in Table 2. The table also includes the averaging kernel sensitivity a_i from the inversion for each sector and subsector i, which we estimate by summing emissions from sector/subsector i for all $0.5° \times 0.625°$ grid cells over the CONUS into one state vector element using a summation matrix (Calisesi et al., 2005; Maasakkers et al., 2019). The one-dimensional summation matrix ($\mathbf{W_i} = (w_{i,k})$) weighs the relative contribution $w_{i,k}$ of sector/subsector i to the

15   total emission in Gaussian k in the prior inventory:

$$a_i = \mathbf{W_i} \mathbf{A} \mathbf{W_i^*} \tag{7}$$

Here $\mathbf{W_i^*} = \mathbf{W_i}^T \left(\mathbf{W_i} \mathbf{W_i}^T \right)^{-1}$ is the generalized pseudo-inverse of $\mathbf{W_i}$, and $\mathbf{A}$ is the averaging kernel matrix (Equation 6). $a_i = 1$ means that the inversion can fully constrain the national total for that emission category, independent of the prior estimate, while $a_i = 0$ means that the inversion provides no information and the estimate cannot depart from the prior. The

20   averaging kernel sensitivities for individual sectors/subsectors in Table 2 vary based on the uncertainty of the prior emission estimates and the GOSAT observation density in the regions of emissions. Posterior wetland emissions ($a_i = 0.68$) are 68% informed by the observations (32% by the prior) because the prior uncertainty is large. Total US anthropogenic emissions are 60% informed by the observations, with less information for individual sectors/subsectors. Emissions from oil/gas production are particularly well informed (45-54%) because they are large and have relatively little spatial overlap with other sectors.

Our posterior estimate for the mean 2010-2015 CONUS anthropogenic source is 30.6 (29.4-31.3) Tg a^{-1}, where the best estimate is from the base inversion and the range is from the inversion ensemble. The 2012 emission total from the EPA GHGI (EPA, 2016) used as prior estimate in our inversion is 28.7 Tg a^{-1}, with an uncertainty range 26.4-36.2. Examining the contributions from different sectors, our best posterior estimates for landfills and livestock are within 5% of the GHGI, and

30   coal emissions are 6% higher. Oil/gas emissions total 11.1 Tg a^{-1} in our base inversion, 22 (12-32)% higher than the GHGI, and driven by oil/gas production as seen for example in Texas, Oklahoma, and offshore in the Gulf of Mexico. Our national estimates for the emissions from oil and gas production are 3.1 (2.7-3.6) and 5.4 (4.9-5.9) Tg a^{-1}, respectively, as compared

to 2.3 and 4.4 Tg a^{-1} in the GHGI.

Our scaling factors to the EPA GHGI are for the 2012 emissions as reported by EPA (2016) and used in the inversion as prior estimates. More recently, EPA (2020) updated its methodology for estimating emissions and applied it to a reanalysis of emissions from previous years including 2012. Changes are important for some oil/gas subsectors, as shown in Figure 5. Gas production emissions are lower by 19% because of a downward correction to gathering and boosting emissions. Oil production emissions are 30% lower than previously reported, caused by a correction for wells that were previously double-counted. Our correction factor from the inversion increases oil production emissions by a factor 1.9 (1.7-2.3) and natural gas production emissions by a factor 1.5 (1.4-1.6) relative to the updated 2012 GHGI from EPA (2020). Similarly, emissions from gas processing in the updated GHGI are 55% lower than previously reported, but our inversion finds them to be higher. Our correction factor from the inversion increases gas processing emissions by a factor of 2.9 (2.6-3.1) relative to the updated GHGI.

Table 2. Mean 2010-2015 methane emissions in the contiguous US (CONUS)

Source (Tg a^{-1})	Prior estimate[a]	Posterior estimate[b]	Sensitivity[c]
Natural	15.7	11.8 (7.1-12.7)	0.62
Wetlands	14.2	10.2 (5.6-11.1)	0.68
Open fires	0.5	0.4 (0.4-0.5)	0.15
Termites	0.6	0.6 (0.6-0.6)	0.18
Geological seeps	0.5	0.5 (0.5-0.5)	0.16
Anthropogenic	28.7	30.6 (29.4-31.3)	0.60
Livestock			
Enteric Fermentation	6.7	6.9 (6.3-7.0)	0.32
Manure Management	2.5	2.5 (2.1-2.5)	0.31
Oil and Natural Gas			
Gas production	4.4	5.4 (4.9-5.9)	0.45
Oil production	2.3	3.1 (2.7-3.6)	0.54
Gas Transmission	1.1	1.1 (1.1-1.2)	0.17
Gas Processing	0.9	1.1 (1.0-1.2)	0.49
Gas Distribution	0.5	0.4 (0.4-0.4)	0.41
Landfills			
Municipal	5.2	5.0 (4.7-5.0)	0.39
Industrial	0.6	0.5 (0.5-0.5)	0.24
Coal Mining			
Underground	2.2	2.4 (2.3-2.5)	0.31
Surface	0.5	0.5 (0.4-0.5)	0.32
Abandoned	0.2	0.3 (0.3-0.3)	0.22
Wastewater			
Municipal	0.5	0.4 (0.4-0.4)	0.22
Industrial	0.2	0.2 (0.1-0.2)	0.19
Rice Cultivation	0.5	0.4 (0.3-0.5)	0.30
Other Anthropogenic[d]	0.5	0.4 (0.4-0.5)	0.34
Total Source	44.5	42.4 (37.0-42.9)	0.63

[a] The prior estimates include the 2012 EPA GHGI emissions (EPA, 2016) and the 2010-2015 mean of the WetCharts inventory ensemble for wetlands (Bloom et al., 2017).

[b] Posterior estimates from our base inversion, with range from the inversion ensemble in parentheses.

[c] Sensitivity of the posterior estimate to the observations as diagnosed from the averaging kernel matrix, ranging from 0 (no sensitivity, posterior equal to prior) to 1 (full sensitivity, posterior solely determined by the observations). For example, a sensitivity of 0.62 means that 62% of the posterior estimate is constrained by the observations and 38% is constrained by the prior.

[d] Including fossil fuel combustion, industrial processes, agricultural field burning, and composting.

3.2 2010-2015 Emission trends

Figure 6 shows linear emission trends for 2010-2015 optimized by the base inversion (top left panel) and the sensitivity inversion including no interannual variability in prior estimates for wetlands and open fires (top right panel). The base inversion shows a trend of increasing emissions from US wetlands, but this is relative to the prior WetCHARTS estimate of interannual

5   variability of wetland emissions which vary from 31 Tg in 2015 to 36 Tg in 2010. Due to the sparsity of independent constraints, the WetCHARTs inter-annual variations have not been extensively evaluated (Bloom et al., 2017), therefore we have little confidence in these year-to-year emission changes. The sensitivity inversion including no prior interannual variability for wetlands shows no large trends in US wetland emissions. Both inversions show similar results for the emission trends in anthropogenic source regions. We find an increase in total anthropogenic CONUS emissions of 0.14 Tg a^{-1} a^{-1} (0.4 % a^{-1}) over the

10   2010-2015 period. This anthropogenic trend is much lower than the 2.8 ± 0.3% a^{-1} increase reported for 2010-2014 by Turner et al. (2016) and more in line with the 2006-2015 trend of 0.7 ± 0.3 % a^{-1} in total US emissions estimated by Lan et al. (2019).

The GHGI (EPA, 2020) reports a 0.35 Tg $a^{-1}a^{-1}$ decrease in anthropogenic US emissions from 2010 to 2015, at odds with our result. The decrease in the GHGI is mainly driven by decreasing emissions from landfills (-0.10 Tg $a^{-1}a^{-1}$) and coal min-

15   ing (-0.17 Tg $a^{-1}a^{-1}$). We find small decreases in the western US that may be related to decreases in emissions from landfills or coal mines (Wyoming). On the national scale, however, we find no evidence for decreasing emissions from landfills or coal mining.

For total oil/gas emissions we find a US trend of 0.4 (0-1) % a^{-1} for 2010-2015, smaller than the 3.4 ± 1.4 % a^{-1} increase

20   reported by Lan et al. (2019) for 2006-2015. The discrepancy may be explained by the different time periods and the fact that the Lan et al. (2019) oil/gas trend is mainly determined by stations in Oklahoma, North Dakota, and Texas. Most of our increase is driven by the Marcellus Shale area in the Northeast US, amounting to 130 (20-190) Gg a^{-1} a^{-1}. This area covering Pennsylvania, Ohio, and West Virginia has seen a large increase in natural gas production driven by unconventional drilling. Natural gas production in the area increased by a factor 7.8 between 2010 and 2015, contributing 22% of US natural gas production in

25   2015 (EIA, 2020a). The latest GHGI shows a national 60 Gg a^{-1} a^{-1} decrease in natural gas production emissions over the 2010-2015 period (EPA, 2020), mainly due to decreasing onshore production and exploration emissions and partly offset by increasing gathering and boosting emissions. For onshore production emissions, the GHGI primarily estimates emissions on the basis of the number of wells, rather than by production rate, and this may underestimate the trend in the Marcellus Shale as the number of wells only increased by 13% over 2010-2015 (EIA, 2020b) despite the large increase in production. The

30   inversion suggests additional increases over production regions in Texas (Permian Basin) and Oklahoma. The Permian Basin has seen a large increase in production after 2015 and is currently the largest oil-producing Basin in the US (Zhang et al., 2020).

GOSAT provides little information over Canada and Mexico when it comes to trends. There are signs that oil/gas production emissions in both Alberta (Canada) and offshore in the Gulf of Mexico are decreasing, and the latter may be driven by

decreasing oil production (Zhang et al., 2019).

3.3 Comparison with other evaluations of the EPA inventory

A number of studies using atmospheric measurements over the US have previously compared their findings to the gridded
5 version of the EPA GHGI as reported by Maasakkers et al. (2016) and used as prior estimate in our inversion. Based on an
upscaling of facility-level measurements and aircraft data, Alvarez et al. (2018) estimated 2015 US oil/gas emissions of 13 (11-
15) Tg a^{-1}, consistent with our posterior estimate of 11.4 (10.3-12.2) Tg a^{-1} for 2015 (posterior mean 2010-2015 emissions
plus trend) and much higher than the 7.3 Tg a^{-1} national total from the latest GHGI (EPA, 2020). Similar to our subsector
attribution, Alvarez et al. (2018) find the largest difference with the GHGI for the production subsector (factor 2), which they
10 attribute to the GHGI emissions not accounting for emissions from abnormal operating conditions. While Alvarez et al. (2018)
did not distinguish between oil and gas production, our results point at a much larger relative discrepancy with the GHGI for
oil production emissions than natural gas production emissions.

In an inversion of data from two tower networks and one aircraft campaign, Cui et al. (2019) found 2014-2016 California
15 methane emissions to be 2.05 ± 0.26 Tg a^{-1}. Our posterior estimate of 2015 California emissions is 1.6 (0.8-1.7) Tg a^{-1},
representing a significant decrease from the prior estimate of 2.3 Tg a^{-1}. This is due to our large reduction of mean 2010-2015
emissions in the Los Angeles Basin. This reduction may be overestimated because of the coarseness of model CO_2 used in the
proxy retrieval, underestimating CO_2 over Los Angeles (Turner et al., 2015). Using aircraft measurements, Ren et al. (2019)
found 70% higher oil/gas production emissions in the Marcellus Shale (Pennsylvania/West Virginia) in 2015 compared to the
20 2012 gridded EPA inventory, which they attribute to an increase in production. We find a 2010-2015 increase in emissions for
that region, as discussed above, but 2015 emissions are still only 22% higher than the GHGI. Based on surface observations
in the Uintah Basin in Utah, Foster et al. (2017) found good agreement with basin-wide emissions from Karion et al. (2013),
and found the gridded EPA inventory to be 45% lower after adjusting emissions based on 2015 production data. Most of the
emissions in the Uintah Basin are concentrated in one of our grid cells and that cell is optimized individually in our inversion
25 with good constraints (Fig 3), finding 2015 posterior emissions that are 37(12-66)% higher than the prior and no significant
linear trend.

Based on aircraft data, Plant et al. (2019) found a factor 2 higher anthropogenic urban methane emissions compared to the
gridded EPA inventory over five cities on the US East Coast. We find no such difference but evaluating urban emissions along
30 the East Coast is difficult because of overlap with large wetlands emissions that are themselves highly uncertain. It should be
possible in principle to separate urban and wetland emissions on the basis of seasonality but we have little confidence in doing
so with the GOSAT data because of the need for a seasonal correction to the model - GOSAT mismatch (Section 2.3) and the
uncertainty in the seasonality of wetland emissions (Melton et al., 2013; Poulter et al., 2017).

4 Conclusions

We have used 2010-2015 methane column data from the GOSAT satellite instrument in a high-resolution inversion of methane emissions and their trends over North America during that period. The inversion for the contiguous US (CONUS) uses as prior estimate a gridded version of the EPA Inventory of U.S. Greenhouse Gas Emissions and Sinks (GHGI), so that results from the

5   inversion are directly relevant for evaluating the GHGI including the contributions from different sectors/subsectors to national methane emissions. We use a 600-member Gaussian mixture model (GMM) as state vector for the inversion that enables us to achieve high resolution ($0.5° \times 0.625°$) in concentrated source regions, and an analytic solution to the Bayesian inverse problem that includes full characterization of information content and facilitates the computation of an ensemble of sensitivity inversions to estimate uncertainty.

We find a best estimate for mean US anthropogenic emissions in 2010-2015 of 30.6 Tg a^{-1} (range of 29.4-31.3 Tg a^{-1} from the inversion ensemble), slightly higher than the EPA GHGI estimate of 28.7 (26.4-36.2) Tg a^{-1}. The difference is mainly from oil and gas production, which we find to be higher by 35% (19-59%) and 22% (11-33%) respectively compared to the GHGI. The most recent version of the GHGI EPA (2020) revises emissions from oil/gas production and gas processing emissions

15   downward, opposite from our results. Thus we find that the estimate of emissions from oil production by EPA (2020) is lower than our result by a factor of 2.

Our best estimate of CONUS wetland emissions is 10.2 (5.6-11.1) Tg a^{-1}, representing 24% of total CONUS methane emissions. This is lower than the ensemble mean from the WetCHARTS inventory (14.2 Tg a^{-1}) used as prior estimate, and

20   is consistent with previous studies pointing to overestimates in US wetland emissions. More work is needed to understand the underlying processes. We find a similar overestimate in wetland emissions over Eastern Canada. We estimate mean 2010-2015 anthropogenic emissions of 4.5 (4.4-4.7)Tg a^{-1} for Canada and 6.1 (5.5-6.3) Tg a^{-1} for Mexico. We find that oil/gas emissions in the IMP (2012) inventory reported to the UNFCCC are too low by 20%, mainly driven by oil production.

25   We find from the inversion a 2010-2015 increase in US anthropogenic emissions of 0.14 Tg a^{-1} (0.4 % a^{-1}), much lower than previous GOSAT-based estimates but at odds with the latest EPA GHGI that reports a 0.35 Tg a^{-1} decrease in emissions over that period. Our increase appears to be largely driven by the rapid growth of unconventional oil/gas production in the eastern US.

30   *Author contributions.* JDM and DJJ designed the study. JDM performed the analysis. JDM and MPS performed the simulations. JDM, DJJ, MPS, TRS, HN, JXS, YZ, XL, AAB, KWB, and JRW discussed the results. AAB provided the WetCHARTS emissions and supporting data. RJP provided the GOSAT data and supporting guidance. JDM and DJJ wrote the paper, and all authors provided input on the paper for revision before submission.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This research was funded by the NASA Carbon Monitoring System (CMS) program. RJP is funded via the UK National Centre for Earth Observation (NCEO grant numbers: nceo020005 and NE/N018079/1). We thank the Japanese Aerospace Exploration Agency, National Institute for Environmental Studies, and the Ministry of Environment for the GOSAT data and their continuous support as

5   part of the Joint Research Agreement. This research used the ALICE High Performance Computing Facility at the University of Leicester for the GOSAT retrievals. Part of this research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

References

Alvarez, R. A., Zavala-Araiza, D., Lyon, D. R., Allen, D. T., Barkley, Z. R., Brandt, A. R., Davis, K. J., Herndon, S. C., Jacob, D. J., Karion, A., Kort, E. A., Lamb, B. K., Lauvaux, T., Maasakkers, J. D., Marchese, A. J., Omara, M., Pacala, S. W., Peischl, J., Robinson, A. L., Shepson, P. B., Sweeney, C., Townsend-Small, A., Wofsy, S. C., and Hamburg, S. P.: Assessment of methane emissions from the U.S. oil and gas supply chain, Science, 361, 186–188, https://doi.org/10.1126/science.aar7204, https://science.sciencemag.org/content/361/6398/186, 2018.

Bader, W., Bovy, B., Conway, S., Strong, K., Smale, D., Turner, A., Blumenstock, T., Boone, C., Coulon, A., Garcia, O., et al.: Ten years of atmospheric methane from ground-based NDACC FTIR observations, Atmospheric Chemistry & Physics Discussions, 2016, 1–31, 2016.

Bloom, A. A., Bowman, K. W., Lee, M., Turner, A. J., Schroeder, R., Worden, J. R., Weidner, R., McDonald, K. C., and Jacob, D. J.: A global wetland methane emissions and uncertainty dataset for atmospheric chemical transport models (WetCHARTs version 1.0), Geoscientific Model Development, 10, 2141, 2017.

Bontemps, S., Defourny, P., Van Bogaert, E., Arino, O., Kalogirou, V., and Perez, J. R.: GLOBCOVER 2009 Products description and validation report, URL: http://ionia1. esrin. esa. int/docs/GLOBCOVER2009_Validation_Report_2, 2, 2011.

Bosilovich, M. G., Lucchesi, R., and Suarez, M.: File Specification for MERRA-2. GMAO Office Note No. 9 (Version 1.1),73pp, Available at gmao.gsfc.nasa.gov/pubs/docs/Bosilovich785.pdf, 2016.

Brasseur, G. and Jacob, D.: Mathematical Modeling of Atmospheric Chemistry, 2017.

Bruhwiler, L. M., Basu, S., Bergamaschi, P., Bousquet, P., Dlugokencky, E., Houweling, S., Ishizawa, M., Kim, H.-S., Locatelli, R., Maksyutov, S., Montzka, S., Pandey, S., Patra, P. K., Petron, G., Saunois, M., Sweeney, C., Schwietzke, S., Tans, P., and Weatherhead, E. C.: U.S. CH4 emissions from oil and gas production: Have recent large increases been detected?, Journal of Geophysical Research: Atmospheres, 122, 4070–4083, https://doi.org/10.1002/2016JD026157, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016JD026157, 2017.

Buchwitz, M., Reuter, M., Schneising, O., Boesch, H., Guerlet, S., Dils, B., Aben, I., Armante, R., Bergamaschi, P., Blumenstock, T., et al.: Comparison and quality assessment of near-surface-sensitive satellite-derived CO2 and CH4 global data sets, Remote Sens. Environ, 162, 344–362, 2015.

Butz, A., Guerlet, S., Hasekamp, O., Schepers, D., Galli, A., Aben, I., Frankenberg, C., Hartmann, J.-M., Tran, H., and Kuze, A.: Toward accurate CO2 and CH4 observations from GOSAT, Geophysical Research Letters, 38, 2011.

Calisesi, Y., T., S. V., and van Oss Roeland: Regridding of remote soundings: Formulation and application to ozone profile comparison, Journal of Geophysical Research: Atmospheres, 110, https://doi.org/10.1029/2005JD006122, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2005JD006122, 2005.

Cooperative Global Atmospheric Data Integration Project: Multi-laboratory compilation of atmospheric methane data for the period 1957-2017; obspack_ch4_1_GLOBALVIEWplus_v1.0_2019_01_08, https://doi.org/10.25925/20190108, https://www.esrl.noaa.gov/gmd/ccgg/obspack/data.php?id=obspack_ch4_1_GLOBALVIEWplus_v1.0_2019-01-08, 2019.

Cui, Y. Y., Vijayan, A., Falk, M., Hsu, Y.-K., Yin, D., Chen, X. M., Zhao, Z., Avise, J., Chen, Y., Verhulst, K., Duren, R., Yadav, V., Miller, C., Weiss, R., Keeling, R., Kim, J., Iraci, L. T., Tanaka, T., Johnson, M. S., Kort, E. A., Bianco, L., Fischer, M. L., Stroud, K., Herner, J., and Croes, B.: A Multiplatform Inversion Estimation of Statewide and Regional Methane Emissions in California during 2014–2016, Environmental Science & Technology, 53, 9636–9645, https://doi.org/10.1021/acs.est.9b01769, https://doi.org/10.1021/acs.est.9b01769, pMID: 31347357, 2019.

Darmenov, A. and da Silva, A.: The quick fire emissions dataset (QFED)–documentation of versions 2.1, 2.2 and 2.4, NASA Technical Report Series on Global Modeling and Data Assimilation, NASA TM-2013-104606, 32, 183, 2013.

EIA: Natural Gas Gross Withdrawals and Production, Available at: eia.gov/dnav/ng/ng_prod_sum_dc_NUS_mmcf_m.htm, 2020a.

EIA: Number of Producing Gas Wells, Available at: eia.gov/dnav/ng/ng_prod_wells_s1_a.htm, 2020b.

5   EPA: Inventory of US Greenhouse Gas Emissions and Sinks: 1990-2014, Available at: epa.gov/ghgemissions/inventory-us-greenhouse-gas-emissions-and-sinks-1990-2014, 2016.

EPA: Inventory of US Greenhouse Gas Emissions and Sinks: 1990-2018, Available at: epa.gov/ghgemissions/inventory-us-greenhouse-gas-emissions-and-sinks-1990-2018, 2020.

ESA CCI GHG project team : ESA Greenhouse Gases Climate Change Initiative (GHG_cci): Column-averaged CH4 from GOSAT generated

10   with the OCPR (UoL-PR) Proxy algorithm (CH4_GOS_OCPR), v7.0. Centre for Environmental Data Analysis, Available at: https://catalogue.ceda.ac.uk/uuid/f9154243fd8744bdaf2a59c39033e659, 2018.

Etiope, G.: Natural Gas Seepage: The Earth's Hydrocarbon Degassing, Springer, 2015.

Etiope, G. and Klusman, R. W.: Microseepage in drylands: flux and implications in the global atmospheric source/sink budget of methane, Global and Planetary Change, 72, 265–274, 2010.

15   European Commission: Emission Database for Global AtmosphericResearch (EDGAR), release version 4.2, 2011.

Foster, C. S., Crosman, E. T., Holland, L., Mallia, D. V., Fasoli, B., Bares, R., Horel, J., and Lin, J. C.: Confirmation of Elevated Methane Emissions in Utah's Uintah Basin With Ground-Based Observations and a High-Resolution Transport Model, Journal of Geophysical Research: Atmospheres, 122, 13,026–13,044, https://doi.org/10.1002/2017JD027480, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017JD027480, 2017.

20   Frankenberg, C., Meirink, J. F., Bergamaschi, P., Goede, A. P. H., Heimann, M., Körner, S., Platt, U., van Weele, M., and Wagner, T.: Satellite chartography of atmospheric methane from SCIAMACHY on board ENVISAT: Analysis of the years 2003 and 2004, Journal of Geophysical Research: Atmospheres, 111, https://doi.org/10.1029/2005JD006235, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2005JD006235, 2006.

Fung, I., John, J., Lerner, J., Matthews, E., Prather, M., Steele, L., and Fraser, P.: Three-dimensional model synthesis of the global methane

25   cycle, Journal of Geophysical Research: Atmospheres, 96, 13 033–13 065, 1991.

Hansen, P. C.: The L-curve and its use in the numerical treatment of inverse problems, 1999.

Hausmann, P., Sussmann, R., and Smale, D.: Contribution of oil and natural gas production to renewed increase in atmospheric methane (2007–2014): top–down estimate from ethane and methane column observations, Atmospheric Chemistry and Physics, 16, 3227–3244, https://doi.org/10.5194/acp-16-3227-2016, https://www.atmos-chem-phys.net/16/3227/2016/, 2016.

30   Heald, C. L., Jacob, D. J., Jones, D., Palmer, P. I., Logan, J. A., Streets, D., Sachse, G. W., Gille, J. C., Hoffman, R. N., and Nehrkorn, T.: Comparative inverse analysis of satellite (MOPITT) and aircraft (TRACE-P) observations to estimate Asian sources of carbon monoxide, Journal of Geophysical Research: Atmospheres, 109, 2004.

Holmquist, J. R., Windham-Myers, L., Bliss, N., Crooks, S., Morris, J. T., Megonigal, J. P., Troxler, T., Weller, D., Callaway, J., Drexler, J., et al.: Accuracy and Precision of Tidal Wetland Soil Carbon Mapping in the Conterminous United States, Scientific reports, 8, 9478, 2018.

35   ICF: Economic Analysis of Methane Emission Reduction Opportunities in the Canadian Oil and Natural Gas Industries, https://www.edf.org/sites/default/files/content/canada_methane_cost_curve_report.pdf, 2015.

IMP: Determinacion de factores de emision para emisiones fugitivas de la industria petrolera en Mexico, https://www.gob.mx/cms/uploads/attachment/file/167851/emisiones_fugitivas.pdf, 2012.

IPCC: Guidelines for National Greenhouse Gas Inventories. The National Greenhouse Gas Inventories Programme, Hayama, Kanagawa, Japan, Eggleston, H.S. and Buendia, L. and Miwa, K. and Ngara, T. and Tanabe, K. (eds.), 2006.

Jacob, D. J., Turner, A. J., Maasakkers, J. D., Sheng, J., Sun, K., Liu, X., Chance, K., Aben, I., McKeever, J., and Frankenberg, C.: Satellite observations of atmospheric methane and their value for quantifying methane emissions, Atmospheric Chemistry and Physics, 16, 14 371–14 396, 2016.

Janardanan, R., Maksyutov, S., Ito, A., Yukio, Y., and Matsunaga, T.: Assessment of anthropogenic methane emissions over large regions based on GOSAT observations and high resolution transport modeling, Remote Sensing, 9, 941, 2017.

Janssens-Maenhout, G., Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Dentener, F., Bergamaschi, P., Pagliari, V., Olivier, J., Peters, J., van Aardenne, J., Monni, S., Doering, U., Petrescu, R., Solazzo, E., and Oreggioni, G.: EDGAR v4.3.2 Global Atlas of the three major Greenhouse Gas Emissions for the period 1970–2012, Earth System Science Data Discussions, 2019, 1–52, https://doi.org/10.5194/essd-2018-164, https://www.earth-syst-sci-data-discuss.net/essd-2018-164/, 2019.

Karion, A., Sweeney, C., Pétron, G., Frost, G., Michael Hardesty, R., Kofler, J., Miller, B. R., Newberger, T., Wolter, S., Banta, R., Brewer, A., Dlugokencky, E., Lang, P., Montzka, S. A., Schnell, R., Tans, P., Trainer, M., Zamora, R., and Conley, S.: Methane emissions estimate from airborne measurements over a western United States natural gas field, Geophysical Research Letters, 40, 4393–4397, https://doi.org/10.1002/grl.50811, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/grl.50811, 2013.

Kirschke, S., Bousquet, P., Ciais, P., Saunois, M., Canadell, J. G., Dlugokencky, E. J., Bergamaschi, P., Bergmann, D., Blake, D. R., and Bruhwiler, L.: Three decades of global methane sources and sinks, Nature Geoscience, 6, 813–823, 2013.

Kuze, A., Shiomi, K., Tanaka, M., Ueda, Y., Deguchi, A., Yoshida, J., Yamamoto, Y., Kataoka, F., Taylor, T., Buijs, H., et al.: Update on GOSAT TANSO-FTS performance, operations, and data products after more than six years in space, Atmospheric Measurement Techniques, 2016, 1–1, 2016.

Kvenvolden, K. A. and Rogers, B. W.: Gaia's breath—global methane exhalations, Marine and Petroleum Geology, 22, 579–590, 2005.

Lan, X., Tans, P., Sweeney, C., Andrews, A., Dlugokencky, E., Schwietzke, S., Kofler, J., McKain, K., Thoning, K., Crotwell, M., Montzka, S., Miller, B. R., and Biraud, S. C.: Long-Term Measurements Show Little Evidence for Large Increases in Total U.S. Methane Emissions Over the Past Decade, Geophysical Research Letters, 46, 4991–4999, https://doi.org/10.1029/2018GL081731, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018GL081731, 2019.

Lehner, B. and Döll, P.: Development and validation of a global database of lakes, reservoirs and wetlands, Journal of Hydrology, 296, 1–22, 2004.

Maasakkers, J. D., Jacob, D. J., Sulprizio, M. P., Turner, A. J., Weitz, M., Wirth, T., Hight, C., DeFigueiredo, M., Desai, M., and Schmeltz, R.: Gridded national inventory of US methane emissions, Environmental science & technology, 50, 13 123–13 133, 2016.

Maasakkers, J. D., Jacob, D. J., Sulprizio, M. P., Scarpelli, T. R., Nesser, H., Sheng, J.-X., Zhang, Y., Hersher, M., Bloom, A. A., Bowman, K. W., Worden, J. R., Janssens-Maenhout, G., and Parker, R. J.: Global distribution of methane emissions, emission trends, and OH concentrations and trends inferred from an inversion of GOSAT satellite data for 2010–2015, Atmospheric Chemistry and Physics, 19, 7859–7881, https://doi.org/10.5194/acp-19-7859-2019, https://www.atmos-chem-phys.net/19/7859/2019/, 2019.

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

Miller, S. M., Wofsy, S. C., Michalak, A. M., Kort, E. A., Andrews, A. E., Biraud, S. C., Dlugokencky, E. J., Eluszkiewicz, J., Fischer, M. L., Janssens-Maenhout, G., Miller, B. R., Miller, J. B., Montzka, S. A., Nehrkorn, T., and Sweeney, C.: Anthropogenic emissions of methane in the United States, Proceedings of the National Academy of Sciences, 110, 20 018–20 022, https://doi.org/10.1073/pnas.1314392110, http://www.pnas.org/content/110/50/20018, 2013.

5 Olivier, J. and Janssens-Maenhout, G.: CO2 Emissions From Fuel Combustion (2012 Edition): Part III: Greenhouse-Gas Emissions, 2012.

Parker, R. J., Boesch, H., Byckling, K., Webb, A. J., Palmer, P. I., Feng, L., Bergamaschi, P., Chevallier, F., Notholt, J., Deutscher, N., Warneke, T., Hase, F., Sussmann, R., Kawakami, S., Kivi, R., Griffith, D. W. T., and Velazco, V.: Assessing 5 years of GOSAT Proxy XCH$_4$ data and associated uncertainties, Atmospheric Measurement Techniques, 8, 4785–4801, https://doi.org/10.5194/amt-8-4785-2015, https://www.atmos-meas-tech.net/8/4785/2015/, 2015.

10 Plant, G., Kort, E. A., Floerchinger, C., Gvakharia, A., Vimont, I., and Sweeney, C.: Large Fugitive Methane Emissions From Urban Centers Along the U.S. East Coast, Geophysical Research Letters, 46, 8500–8507, https://doi.org/10.1029/2019GL082635, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019GL082635, 2019.

Poulter, B., Bousquet, P., Canadell, J. G., Ciais, P., Peregon, A., Saunois, M., Arora, V. K., Beerling, D. J., Brovkin, V., Jones, C. D., Joos, F., Gedney, N., Ito, A., Kleinen, T., Koven, C. D., McDonald, K., Melton, J. R., Peng, C., Peng, S., Prigent, C., Schroeder, R.,
15 Riley, W. J., Saito, M., Spahni, R., Tian, H., Taylor, L., Viovy, N., Wilton, D., Wiltshire, A., Xu, X., Zhang, B., Zhang, Z., and Zhu, Q.: Global wetland contribution to 2000–2012 atmospheric methane growth rate dynamics, Environmental Research Letters, 12, 094 013, https://doi.org/10.1088/1748-9326/aa8391, https://doi.org/10.1088%2F1748-9326%2Faa8391, 2017.

Ren, X., Hall, D. L., Vinciguerra, T., Benish, S. E., Stratton, P. R., Ahn, D., Hansford, J. R., Cohen, M. D., Sahu, S., He, H., Grimes, C., Fuentes, J. D., Shepson, P. B., Salawitch, R. J., Ehrman, S. H., and Dickerson, R. R.: Methane Emissions from the Marcel-
20 lus Shale in Southwestern Pennsylvania and Northern West Virginia Based on Airborne Measurements, Journal of Geophysical Research: Atmospheres, 124, 1862–1878, https://doi.org/10.1029/2018JD029690, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018JD029690, 2019.

Rodgers, C. D.: Inverse methods for atmospheric sounding: theory and practice, vol. 2, World scientific, 2000.

Saad, K. M., Wunch, D., Deutscher, N. M., Griffith, D. W., Hase, F., De Mazière, M., Notholt, J., Pollard, D. F., Roehl, C. M., and Schnei-
25 der, M.: Seasonal variability of stratospheric methane: implications for constraining tropospheric methane budgets using total column observations, Atmospheric Chemistry and Physics, 16, 14 003, 2016.

SEDEMA: Emission Inventory of the Mexico City Metropolitan Area 2016, 2018.

Sheng, J.-X., Jacob, D. J., Maasakkers, J. D., Sulprizio, M. P., Zavala-Araiza, D., and Hamburg, S. P.: A high-resolution (0.1°× 0.1°) inventory of methane emissions from Canadian and Mexican oil and gas systems, Atmospheric Environment, 158, 211–215, 2017.

30 Sheng, J.-X., Jacob, D. J., Turner, A. J., Maasakkers, J. D., Benmergui, J., Bloom, A. A., Arndt, C., Gautam, R., Zavala-Araiza, D., Boesch, H., and Parker, R. J.: 2010–2016 methane trends over Canada, the United States, and Mexico observed by the GOSAT satellite: contributions from different source sectors, Atmospheric Chemistry and Physics, 18, 12 257–12 267, https://doi.org/10.5194/acp-18-12257-2018, https://www.atmos-chem-phys.net/18/12257/2018/, 2018a.

Sheng, J.-X., Jacob, D. J., Turner, A. J., Maasakkers, J. D., Sulprizio, M. P., Bloom, A. A., Andrews, A. E., and Wunch, D.: High-
35 resolution inversion of methane emissions in the Southeast US using SEAC[4]RS aircraft observations of atmospheric methane: anthropogenic and wetland sources, Atmospheric Chemistry and Physics, 18, 6483–6491, https://doi.org/10.5194/acp-18-6483-2018, https://www.atmos-chem-phys.net/18/6483/2018/, 2018b.

Stanevich, I.: Variational data assimilation of satellite remote sensing observations for improving methane simulations in chemical transport models, Ph.D. thesis, University of Toronto, 2018.

Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M.: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, 1535 pp, 2013.

Streets, D. G., Canty, T., Carmichael, G. R., de Foy, B., Dickerson, R. R., Duncan, B. N., Edwards, D. P., Haynes, J. A., Henze, D. K., Houyoux, M. R., Jacob, D. J., Krotkov, N. A., Lamsal, L. N., Liu, Y., Lu, Z., Martin, R. V., Pfister, G. G., Pinder, R. W., Salawitch, R. J., and Wecht, K. J.: Emissions estimation from satellite retrievals: A review of current capability, Atmospheric Environment, 77, 1011 – 1042, https://doi.org/https://doi.org/10.1016/j.atmosenv.2013.05.051, http://www.sciencedirect.com/science/article/pii/S1352231013004007, 2013.

Turner, A. J. and Jacob, D. J.: Balancing aggregation and smoothing errors in inverse models, Atmospheric Chemistry and Physics, 15, 7039–7048, https://doi.org/10.5194/acp-15-7039-2015, https://www.atmos-chem-phys.net/15/7039/2015/, 2015.

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

Turner, A. J., Jacob, D. J., Benmergui, J., Wofsy, S. C., Maasakkers, J. D., Butz, A., Hasekamp, O., and Biraud, S. C.: A large increase in U.S. methane emissions over the past decade inferred from satellite data and surface observations, Geophysical Research Letters, 43, 2218–2224, https://doi.org/10.1002/2016GL067987, http://dx.doi.org/10.1002/2016GL067987, 2016GL067987, 2016.

UNFCCC: United Nations Framework Convention on Climate Change: Greenhouse Gas Inventory Data, Described at: https://unfccc.int/process/transparency-and-reporting/greenhouse-gas-data/ghg-data-unfccc, obtained from http://di.unfccc.int/detailed_data_by_party, 2019.

United Nations: United Nations Framework Convention on Climate Change, Article 4(1)(a), Available at unfccc.int, 1992.

Wecht, K. J., Jacob, D. J., Frankenberg, C., Jiang, Z., and Blake, D. R.: Mapping of North American methane emissions with high spatial resolution by inversion of SCIAMACHY satellite data, Journal of Geophysical Research: Atmospheres, 119, 7741–7756, 2014.

Zhang, B., Tian, H., Ren, W., Tao, B., Lu, C., Yang, J., Banger, K., and Pan, S.: Methane emissions from global rice fields: Magnitude, spatiotemporal patterns, and environmental controls, Global Biogeochemical Cycles, 30, 1246–1263, 2016.

Zhang, Y., Gautam, R., Zavala-Araiza, D., Jacob, D. J., Zhang, R., Zhu, L., Sheng, J.-X., and Scarpelli, T.: Satellite-Observed Changes in Mexico's Offshore Gas Flaring Activity Linked to Oil/Gas Regulations, Geophysical Research Letters, 46, 1879–1888, https://doi.org/10.1029/2018GL081145, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018GL081145, 2019.

Zhang, Y., Gautam, R., Pandey, S., Omara, M., Maasakkers, J. D., Sadavarte, P., Lyon, D., Nesser, H., Sulprizio, M. P., Varon, D. J., Zhang, R., Houweling, S., Zavala-Araiza, D., Alvarez, R. A., Lorente, A., Hamburg, S. P., Aben, I., and Jacob, D. J.: Quantifying methane emissions from the largest oil-producing basin in the United States from space, Science Advances, 6, https://doi.org/10.1126/sciadv.aaz5120, https://advances.sciencemag.org/content/6/17/eaaz5120, 2020.
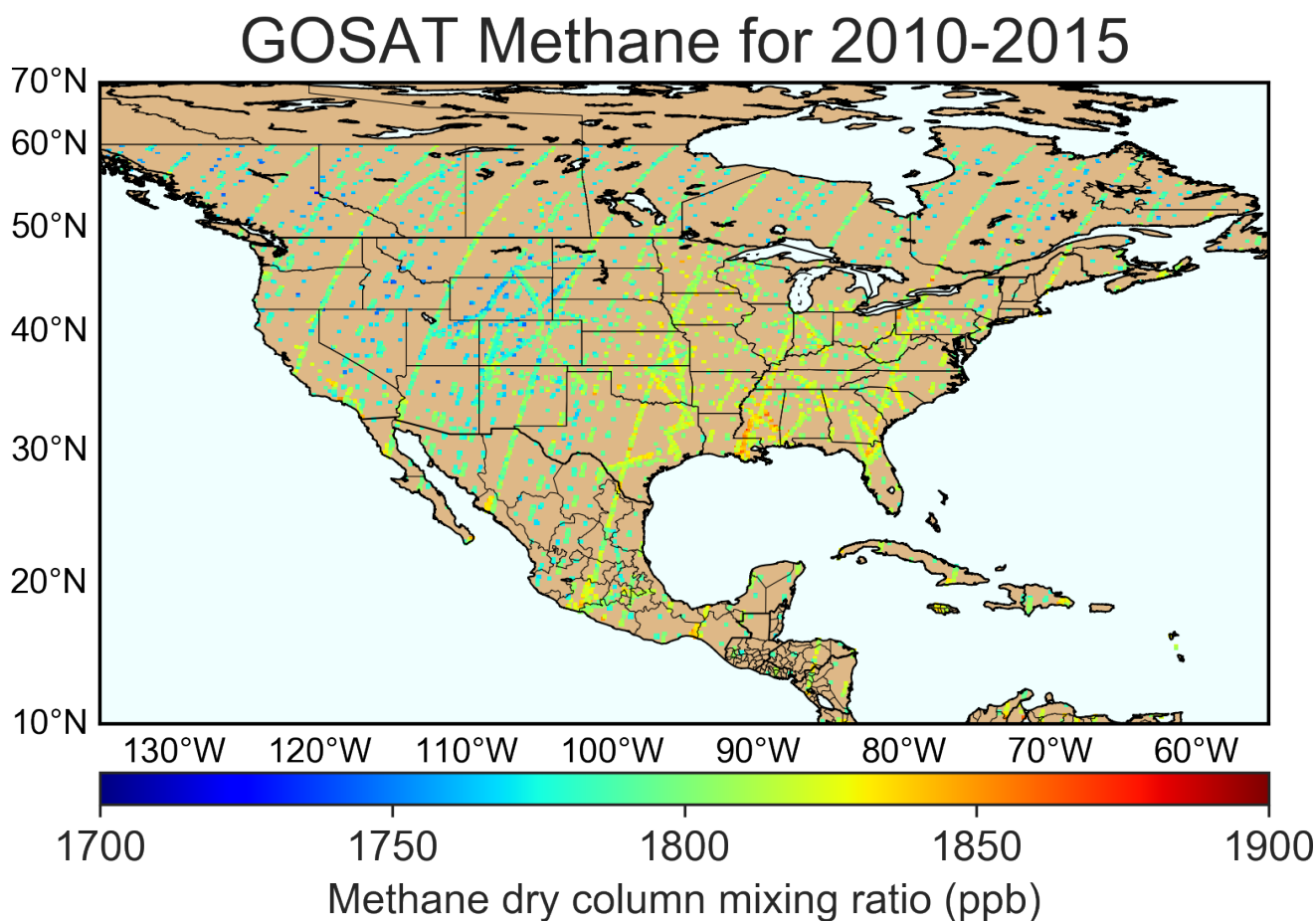
Figure 1. Average 2010-2015 methane dry column mixing ratios over North America observed by GOSAT. There are 156110 individual observations over land used in the inversion. The GOSAT data have 10km pixel resolution but we inflate them here to $0.3° \times 0.3°$ for visibility. GOSAT generally takes repeated observations of the same pixels so that most pixels shown here average a number of observations. GOSAT observations north of $60N°$ are excluded because of their seasonal limitation and uncertainty about the stratospheric correction.
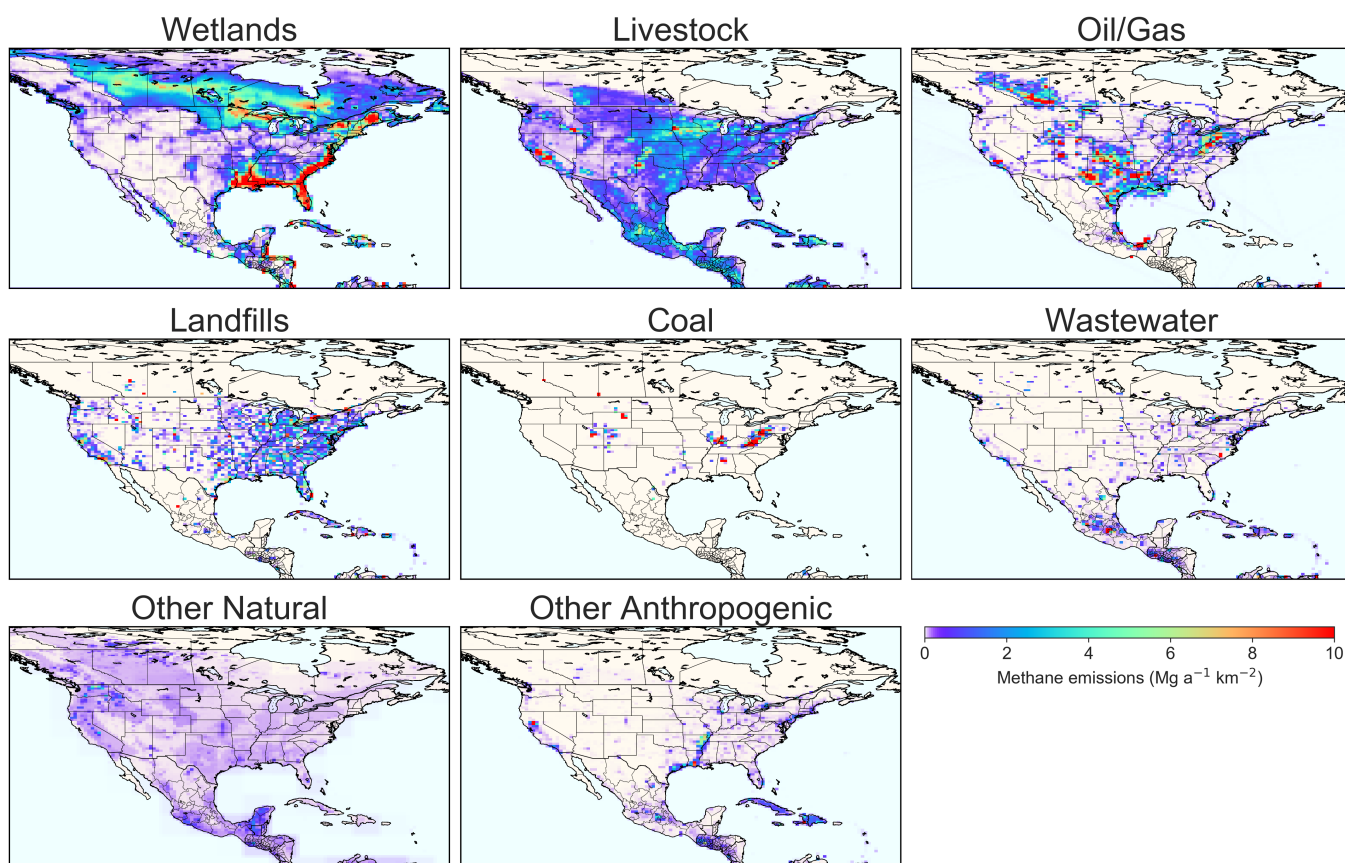
Figure 2. Methane emissions taken as prior estimate for 2010-2015. National totals, subsector breakdowns, and references are in Table 1.
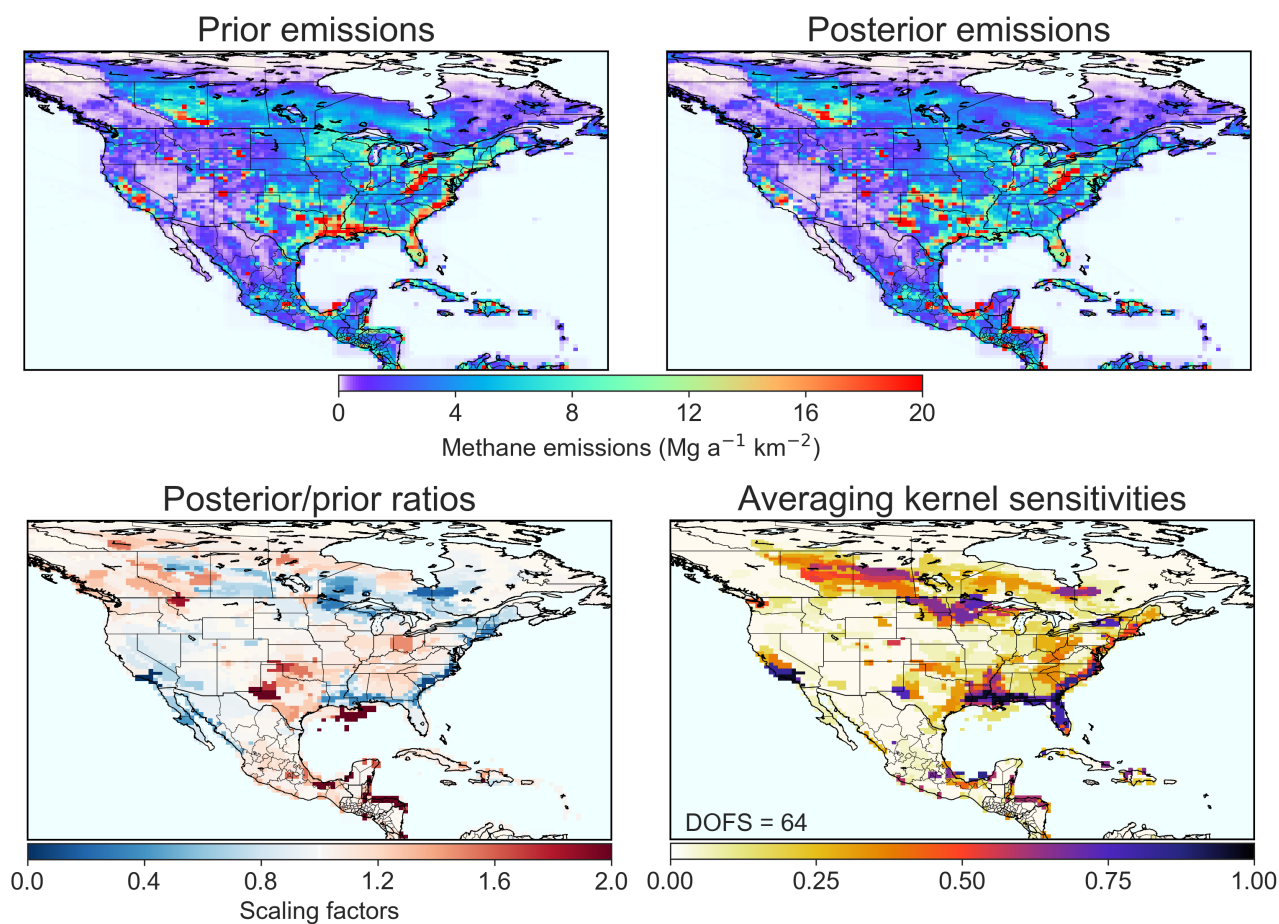
Figure 3. Mean 2010-2015 posterior methane emissions and comparison to the prior estimate. Results are from the base inversion. The bottom right panel shows the averaging kernel sensitivities (diagonal elements of the averaging kernel matrix). The trace of the averaging kernel matrix, i.e., the degrees of freedom for signal (DOFS), is given inset. It represents the number of independent pieces of information that can be constrained by the inversion.
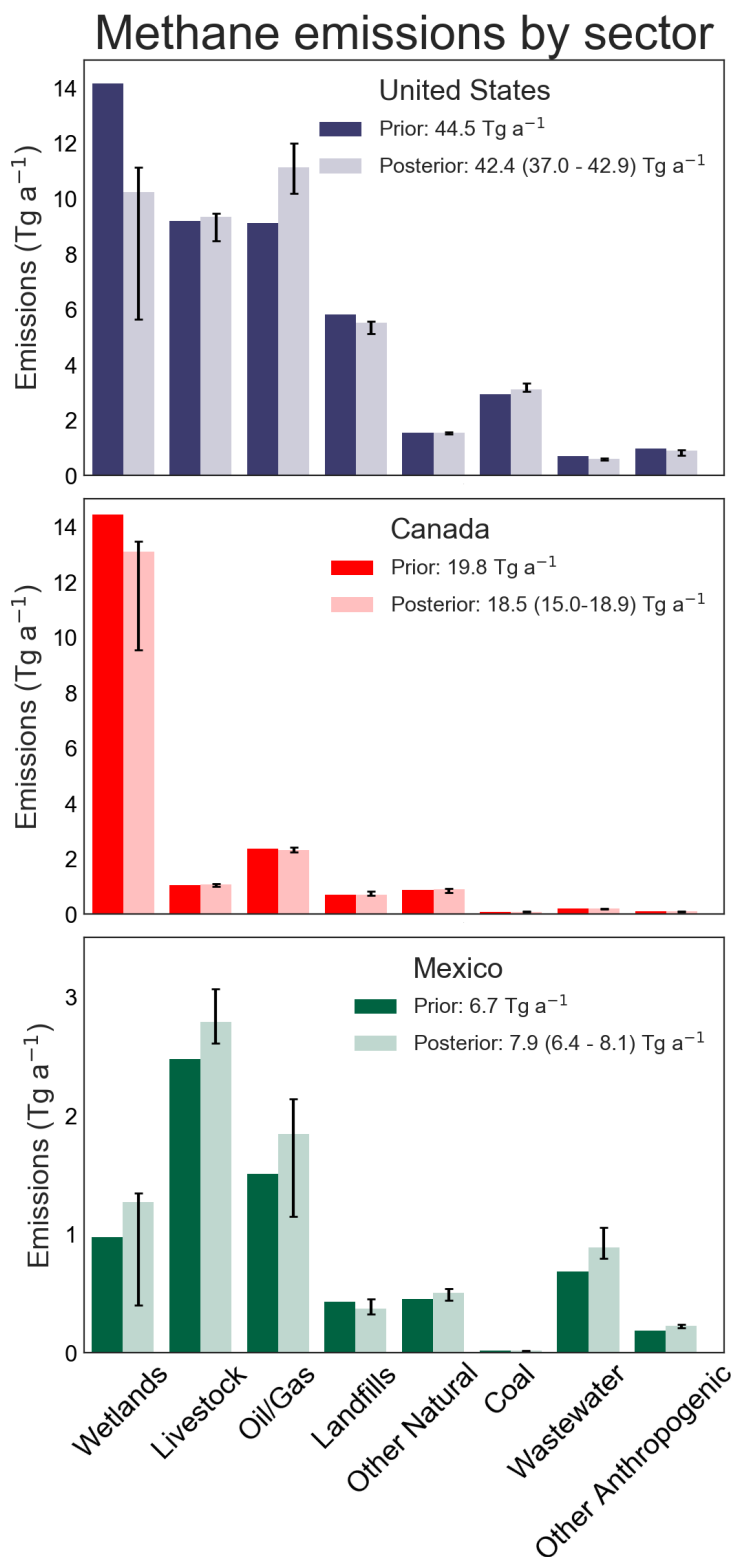
Figure 4. Mean 2010-2015 methane emissions per source sector for the contiguous US (CONUS), Canada, and Mexico. Values are shown for prior estimates (Table 1) and posterior estimates from the inversion of GOSAT data. Vertical bars show the ranges of results from the inversion ensemble.
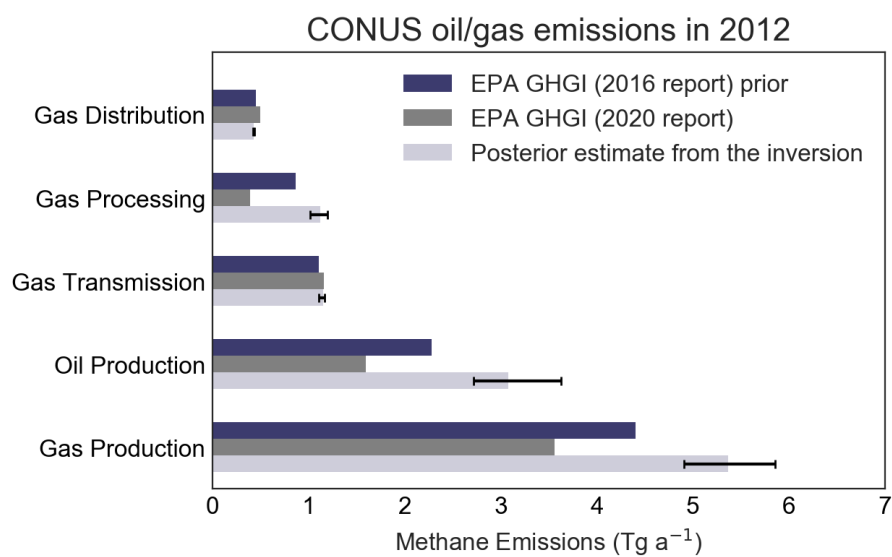
Figure 5. Methane emissions from the oil/gas sector in the contiguous US (CONUS) in 20212. The figure shows the original EPA GHGI estimates for 2012 used as prior in the inversion (EPA, 2016), the updated estimates for 2012 based on revised methodology (EPA, 2020), and the posterior results from the inversion. Horizontal bars give the ranges of the inversion ensemble.
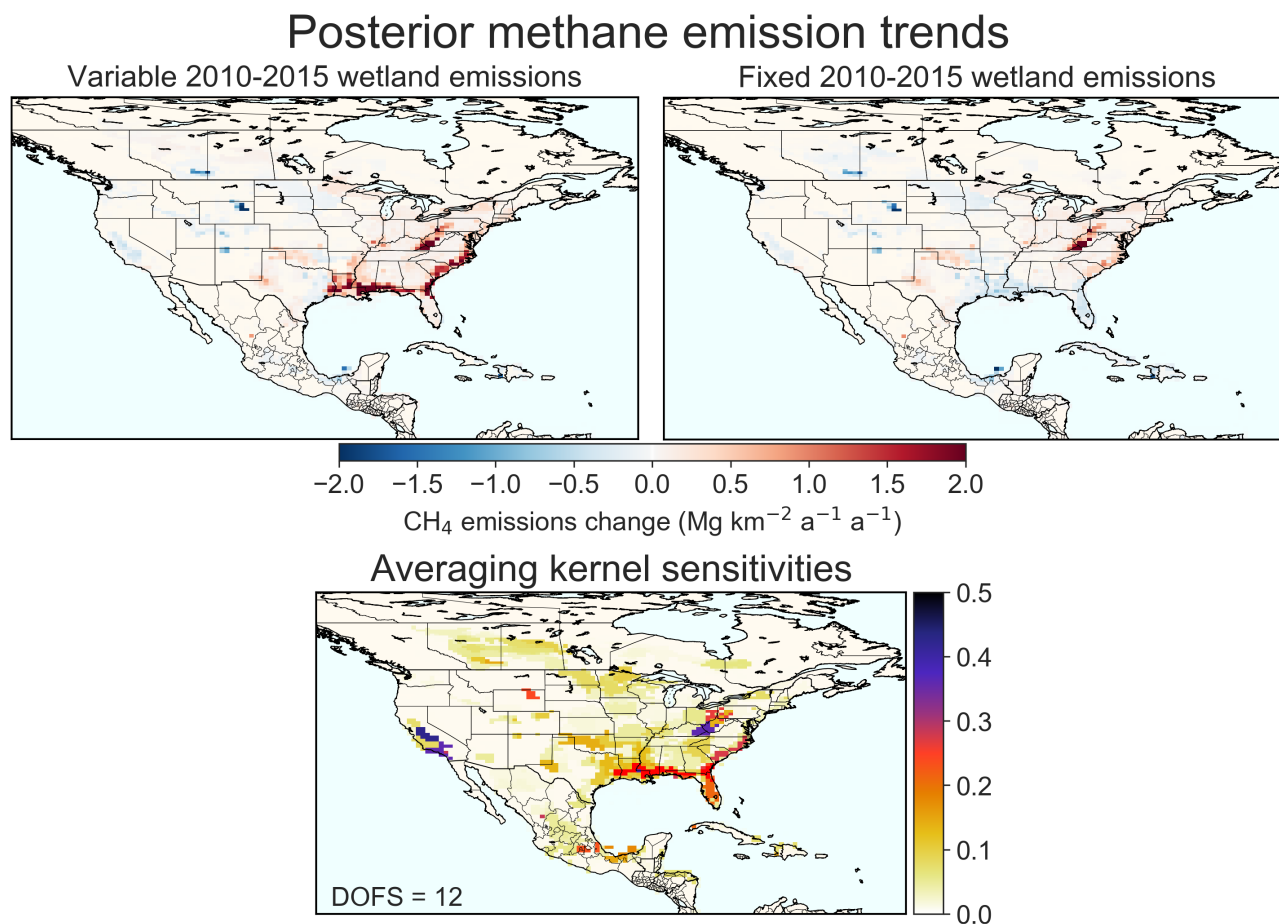
Figure 6. 2010-2015 methane emission trends in North America. The top panels show the posterior estimates from the inversion of GOSAT satellite data, allowing for 2010-2015 interannual variability in the prior estimates for wetlands and open fires (base inversion, left) and assuming constant prior estimates for 2010-2015 (right). The bottom figure shows the averaging kernel sensitivities, with the degrees of freedom for signal (DOFS) indicated inset.