Supplement of

Estimating 2010–2015 Anthropogenic and Natural Methane Emissions in Canada using ECCC Surface and GOSAT Satellite Observations

Sabour Baray et al.

Correspondence to: Sabour Baray (sabour@yorku.ca)

S1.1 Monthly GOSAT Data in the Canadian Domain

Figure S1 shows the GOSAT data available per month using 2013 as an example year, this corresponds to the data coverage shown in Fig. 1 of the main text but highlights the variability in satellite observational coverage over a single year. GOSAT data shown passes all quality assurance flags and includes our domain filter to land data that is within 50°W to 150°W longitude and 45°N to 60°N latitude. The minimum in December observations (*n=112*) and neighbouring months is due to less solar radiation in the winter resulting in less retrievals. Fewer observations cause the inversion to favour the prior state of emissions. There are less methane emissions from Canadian wetlands in the coldest months of the winter, and the comparison between the prior, the posterior using GOSAT data, and the posterior using ECCC data shows very small differences in emissions estimates for these coldest months.

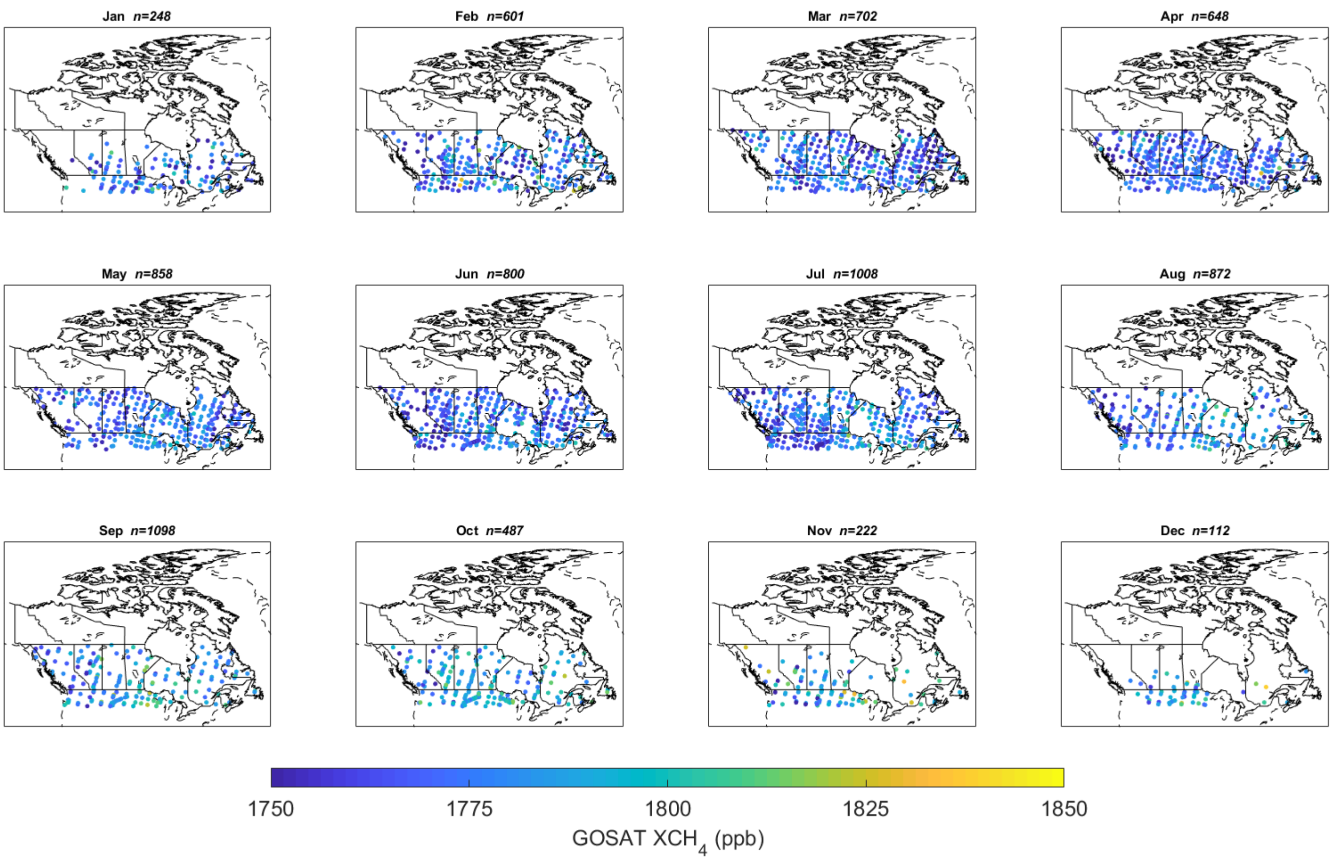

Figure S1: GOSAT observations per month in the year 2013 corresponding to Fig. 1 in the main text (*n=7656* observations for the entire year). Observations are filtered to land data that is within 50°W to 150° W longitude and 45°N to 60° N latitude.

S1.2 Sensitivity of Seasonal Emissions to Climatological Data

We select four climatological stations shown in Table S1 to sample temperature and precipitation data from 2010–2015 in the four provinces where wetlands emissions are concentrated (Alberta, Saskatchewan, Manitoba, and Ontario). These stations are not exhaustive and are chosen for their proximity to the stations shown in Table 1. Station measurements are quality-controlled from the National Climate Data Archive from Environment and Climate Change Canada (Hutchinson et al., 2009).

Table S1: Climatological sites used for air temperature and total precipitation measurements for the seasonality comparison.

Site Name, Province	Latitude	Longitude
Lac La Biche Climate, Alberta	54.8° N	112.0° W
La Ronge, Saskatchewan	55.1° N	105.3° W
Churchill Climate, Manitoba	58.7° N	94.1° W
Moosonee, Ontario	51.3° N	80.6° W

Figure S2 shows the mean 2010–2015 seasonal pattern of natural methane emissions constrained by ECCC and GOSAT data corresponding to Fig. 8 in the main text. These emissions are compared to monthly mean air temperature and precipitation averaged over the four climatological stations in Table S1. We consider air temperature a reasonable proxy for the surface skin temperature that is used in WetCHARTS. Surface skin temperature is itself a proxy for soil temperatures deeper beneath the surface where methane is produced (Miller et al., 2016). Hence both metrics may be lagging indicators for the peak of methane emissions. Both air temperature and precipitation show peaks in July which correspond well with the maxima of methane emissions in the prior from WetCHARTS. Methane emissions in the prior begin to accelerate from March to April, however for both months air temperature is below freezing. It is not likely that soil temperatures and subsurface soil temperatures would be above freezing in these months. Air temperature crosses from below 0° to above freezing one month later from April to May, which corresponds to where the posterior ECCC and GOSAT emissions begin to accelerate. Total precipitation shows the highest acceleration one month later from May to June. As the peak in July is passed, late-summer and autumn air temperatures are higher than the months opposite of the peak (August is warmer than June, September is warmer than May, October is warmer than April). This pattern is corroborated by the precipitation measurements. Air temperatures go below freezing from October to November. As shown by Zona et al. (2016), “zero-curtain” emissions may continue even when the soil is at freezing temperatures. This mechanism may be more likely to occur in the months after the peak if subsurface soils are slower to thaw in the spring and slower to freeze in the autumn. These simple climatological measurements and the described mechanisms suggested in other studies corroborate our posterior results of lower spring methane emissions and lower peak methane emissions in the summer. Our results suggest process models may benefit from better parameterization of possible lagging effects from air temperature and precipitation for Boreal Canada methane emissions.

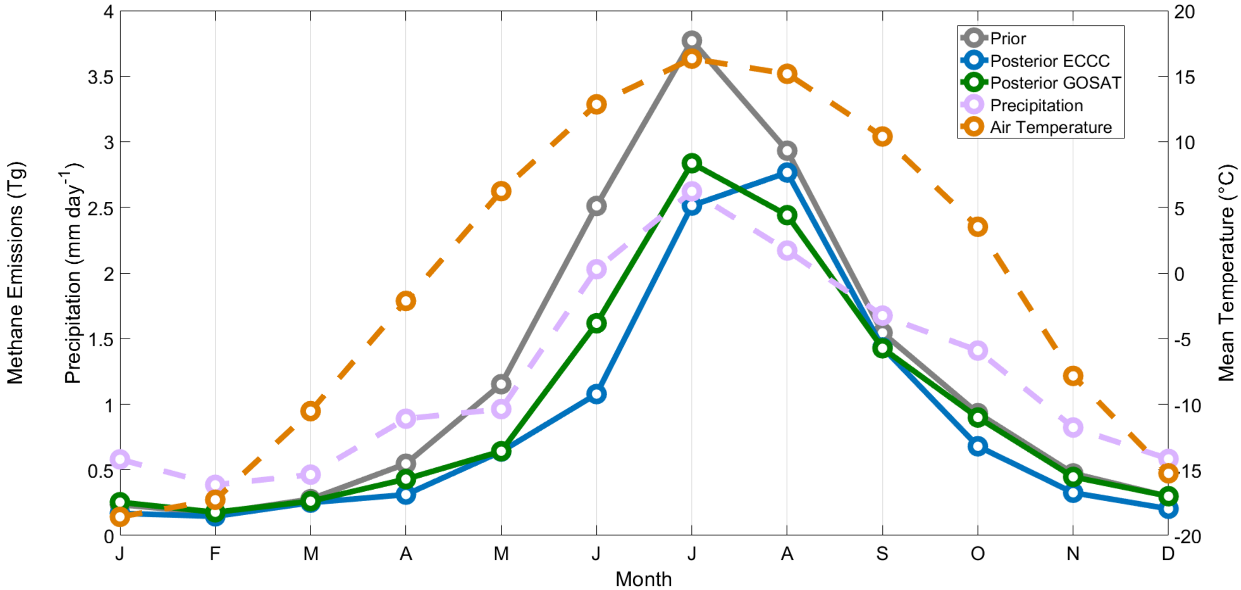

Figure S2: Mean seasonal pattern of 2010-2015 methane emissions from the prior (gray), posterior constrained with ECCC data (blue), posterior constrained with GOSAT data (green). This is compared to the seasonal pattern of monthly mean air temperature (orange, right axis) and precipitation (pink, left axis) from station measurements listed in Table S1. Both air temperature and precipitation show an asymmetry about the July peak, with higher temperature and precipitation in the fall months than the spring.

S1.3 Sensitivity of GOSAT-Constrained Emissions to GEOS-Chem Column Bias Corrections

We test the sensitivity of the posterior GOSAT-constrained methane emissions in our analysis to the use of latitude-dependent and seasonal bias corrections in the GEOS-Chem simulated total column of methane. The latitude-dependent bias correction has a magnitude less than 3.5 ppb for our domain of interest (45 to 60°N). On a global basis the seasonal bias correction has an amplitude of ±4 ppb with a maximum in June and a minimum in December. Figure S3 shows the sensitivity of posterior monthly emissions to these bias corrections using 2013 as an example. We show four versions of the posterior methane emissions using GOSAT data: GOSAT11 (green) is the base case which applies the latitude-dependent bias correction and the seasonal bias correction, GOSAT10 (purple) applies the latitude-dependent bias correction and does not apply the seasonal correction, GOSAT01 (orange) does not apply the latitude-dependent bias correction and applies the seasonal correction, and GOSAT00 (light blue) uses neither bias correction. The range of emissions from all four examples is 9.7 – 10.7 Tg a^{-1}, which are all consistent with the ECCC emissions of 10.0 Tg a^{-1} and lower than the prior emissions of 14.3 Tg a^{-1}. Not applying the latitude-dependent bias correction results in a decrease in the resulting emissions and maintains the same seasonal pattern. Not applying the seasonal bias correction results in a change in the temporal distribution of emissions that better matches the

August peak in the posterior with ECCC data. Emissions are lower than the base case in the spring and higher than the base case in autumn. This change enhances the autumn-shift in emissions that has been described in S1.1. While this may be more consistent with our interpretations, it is not clear whether the difference is due to emissions or transport biases. Stanevich et al. (2019) showed that the latitude dependent bias is most likely due to excessive polar stratospheric transport at high latitudes. If the seasonal bias is indeed due to mischaracterized natural emissions, it is not clear why the bias would be equally large in December (–4 ppb) as June (+4 ppb) on a global basis. The magnitude of natural emissions in December is much lower than June and emissions mischaracterization would not itself produce an equally large bias as the largely overestimated summertime emissions. Our analysis with ECCC data shows most of the adjustments to wetlands are in the peak of summer with some extension into the autumn. These results show that the bias corrections produce minor differences in the magnitude and seasonal pattern of emissions.

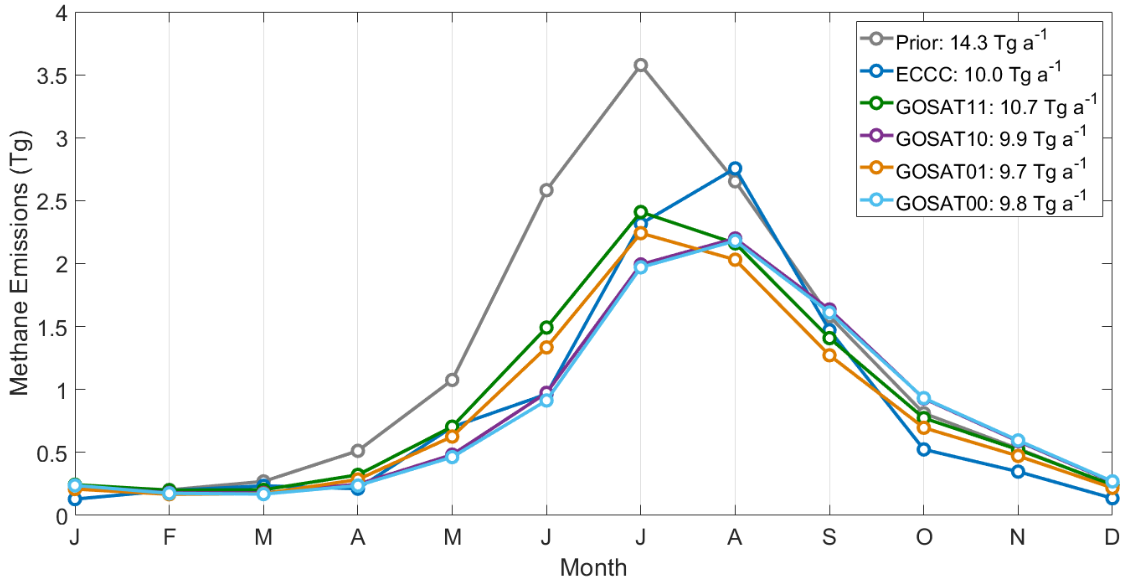

Figure S3: Sensitivity of 2013 posterior GOSAT constrained methane emissions to bias corrections used in the GEOS-Chem simulated total column of methane. For comparison, the prior in 2013 (gray) and the posterior in 2013 constrained by ECCC data (blue) are shown. The digits in the GOSAT label represent the binary use of bias corrections (1 = applied, 0 = not applied). The first digit corresponds to the use of the latitude bias correction, the second digit corresponds to the use of the monthly bias correction, hence GOSAT11 is the base case that applies both bias corrections and GOSAT00 is the case with no bias corrections applied.

S1.4 Diagnostics of Sectoral and Provincial Inversions

In this analysis we first evaluate the correlations and/or independence of the state vector elements from the posterior error covariance matrix $\hat{\mathbf{S}}$ as follows (Heald et al., 2004):

$$r_{ij} = \frac{\hat{s}_{ij}}{\sqrt{\hat{s}_{ii}}\sqrt{\hat{s}_{jj}}} \quad (1s)$$

The error-normalized posterior correlation matrix ***r*** provides information on the independence of the state vector elements. This is corroborated by the averaging kernel matrix **A** which shows which state vector elements contain independent pieces of information, with the trace of **A** providing the total degrees of freedom for signal for the inversion. To further evaluate the signal-to-noise ratio of the observation-constrained state vector elements and their independence from each other we use an eigenanalysis. The Jacobian matrix **K** is normalized about the observational and prior error covariance matrices as follows (Rodgers, 2000):

$$\check{\mathbf{K}} = \mathbf{S}_{\mathrm{o}}^{-1/2}\mathbf{K}\mathbf{S}_{\mathrm{a}}^{1/2} \quad (2s)$$

The singular value decomposition of $\check{\mathbf{K}}$ gives its rank which is the number of singular values greater than one. The singular values also correspond to the signal-to-noise ratio of state vector elements and hence quantify the strength of the observational constraints on individual emissions categories.

Figure S4 shows this series of diagnostics for the sectoral (5 state vector element) inversion and Figure S5 shows the same analysis for the provincial (16 state vector element) inversion. Figure S4 (top) shows the error-normalized correlation matrix for the sectoral inversion. The most important result is that the primary source of natural emissions, wetlands (purple line), is not correlated with the primary source of anthropogenic emissions, energy (blue line). Within the anthropogenic category however, we see that energy is strongly correlated with agriculture, showing that these two elements cannot be distinguished by the observation system. For natural emissions, other natural sources are weakly correlated with wetlands and are not completely independent. Emissions from waste are shown to be independent and can be distinguished from the other sources. The averaging kernel matrix corroborates this result, and shows the three independent pieces of information are energy, wetlands and waste, with partial information content from other natural sources and a lack of information on agriculture. The singular values show strong constraints on wetlands with a signal-to-noise ratio of 37.3, and strong constraints on energy with a signal-to-noise ratio of 5.2. Waste sources are 2.2, other natural are 1.2 and agriculture is below the noise at 0.4. These diagnostics demonstrate that a joint ECCC in situ and GOSAT satellite inversion system can successfully provide constraints on and distinguish the three major categories of methane emissions in Canada: wetlands, energy and waste. Emissions from agricutlure cannot be distinghised in this system and should be aggregated with energy, this is likely because of the strong spatial overlap between these emissions in Western Canada and the lower signal from lower magnitude agriculture emissions. Emissions from other natural sources (biomass burning, seeps, and termites) also are at the noise and should be aggregated with wetlands. This is because minor natural sources are much lower in magnitude (0.8 Tg a^{-1} out of 14.8 Tg a^{-1}) and also show spatial overlap with wetlands.

Figure S5 shows the diagnostics on the provincial (16 state vector element) inversion. This choice of state vector elements challenges the observing system and results in a largely underdetermined solution. These diagnostics allow us to identify where the limitations of the ECCC + GOSAT observing system are. The posterior error correlation matrix $\boldsymbol{r}$ shows the provincial emissions are somewhat correlated a) between anthropogenic/natural emissions of the same province and b) with neighboring provinces in the same category of emissions. For example, AB anthropogenic emissions (solid orange line) show a small inverse correlation with AB natural emissions (dashed-x orange line). The solid orange line also shows a small correlation with the anthropogenic emissions of nearby provinces BC and SK. For the natural emissions, the dashed lines corresponding to natural emissions within a province in most cases extends into the provinces to the east and west. These correlations are not as large as the case of Energy and Agriculture emissions in Fig. S4, and show a more moderate influence of nearby provinces on the optimized emissions. The primary limitation of the provincial inversion is the inability to distinguish provinces with a very small magnitude of emissions. This is shown in the averaging kernel matrix, which has a degrees of freedom for signal of 7.9 out of 16 elements. The 6 regions that are best constrained are AB anthropogenic, ON anthropogenic, AB natural, SK natural, MB natural, and ON natural, with partial constraints on BC anthropogenic, SK anthropogenic, QC anthropogenic, BC natural, QC natural and NOR natural. The singular vectors corroborate this result and show that there are 8 regions that are above the noise and 8 that are at or below the noise. The best constraints on anthropogenic emissions are in Alberta, with a signal to noise ratio as good as 15.1 (solid blue line), followed by Ontario (2.5-2.8).

These diagnostics show that the ECCC+GOSAT observing system for Canada is limited in its ability to characterize agricultural emissions, and somewhat limited in its ability to characterize non-wetlands natural emissions. Hence we present Energy+Agriculture and Wetlands+Other Natural together for our conclusions. More precise and more dense measurements at a finer scale would better disaggregate these sources, although the use of the precise in situ data is primarly limited by the model error (Section 2.3 of the main text). In the provincial inversion, the observing system provides good constraints on anthropogenic emissions from AB and ON and is capable of distinguishing these emissions from natural sources in the same province. However, anthropogenic sources from other provinces with much lower emissions cannot be distinguished. Natural emissions can be characterized from the provinces that are most responsible for wetlands emissions (AB, SK, MB, ON), however the observing system struggles in Atlantic and Northern Canada where the surface and satellite observations we use are limited. The emissions adjustments to state vector elements beneath the noise are due to aliasing with other sources and compensation effects due to interprovincial transport. We limit out conclusions to simple interpretations, we use the limited provincial inversion for spatial attribution to show higher posterior anthropogenic emissions are primarily from the total in Western Canada (BC+AB+SK+MB), and not emissions in Central Canada (ON+QC).

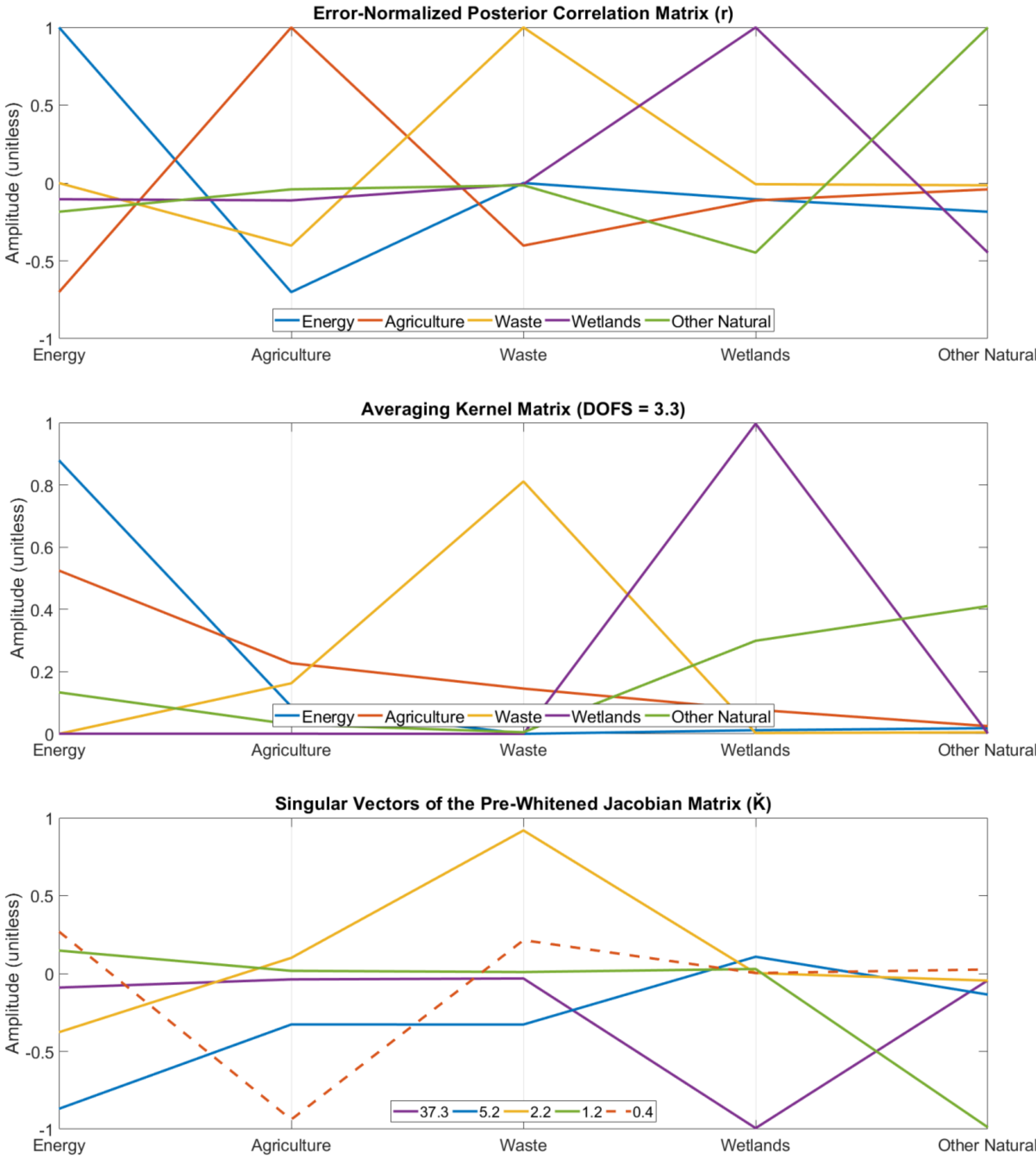

Figure S4: Diagnostics of the sectoral inversion used to evaluate the independence and information content of the 5 state vector elements. The error-normalized posterior correlation matrix (top) shows the correlations between elements. The averaging kernel matrix (middle) shows where the independent pieces of information are (DOFS = 3.3). The singular vector decomposition of the pre-whitened jacobian (bottom) quantifies the signal-to-noise ratio of the significant elements – these are the singular values listed above one (4 in total). The singular vector below noise (agriculture) is shown as a dashed line.

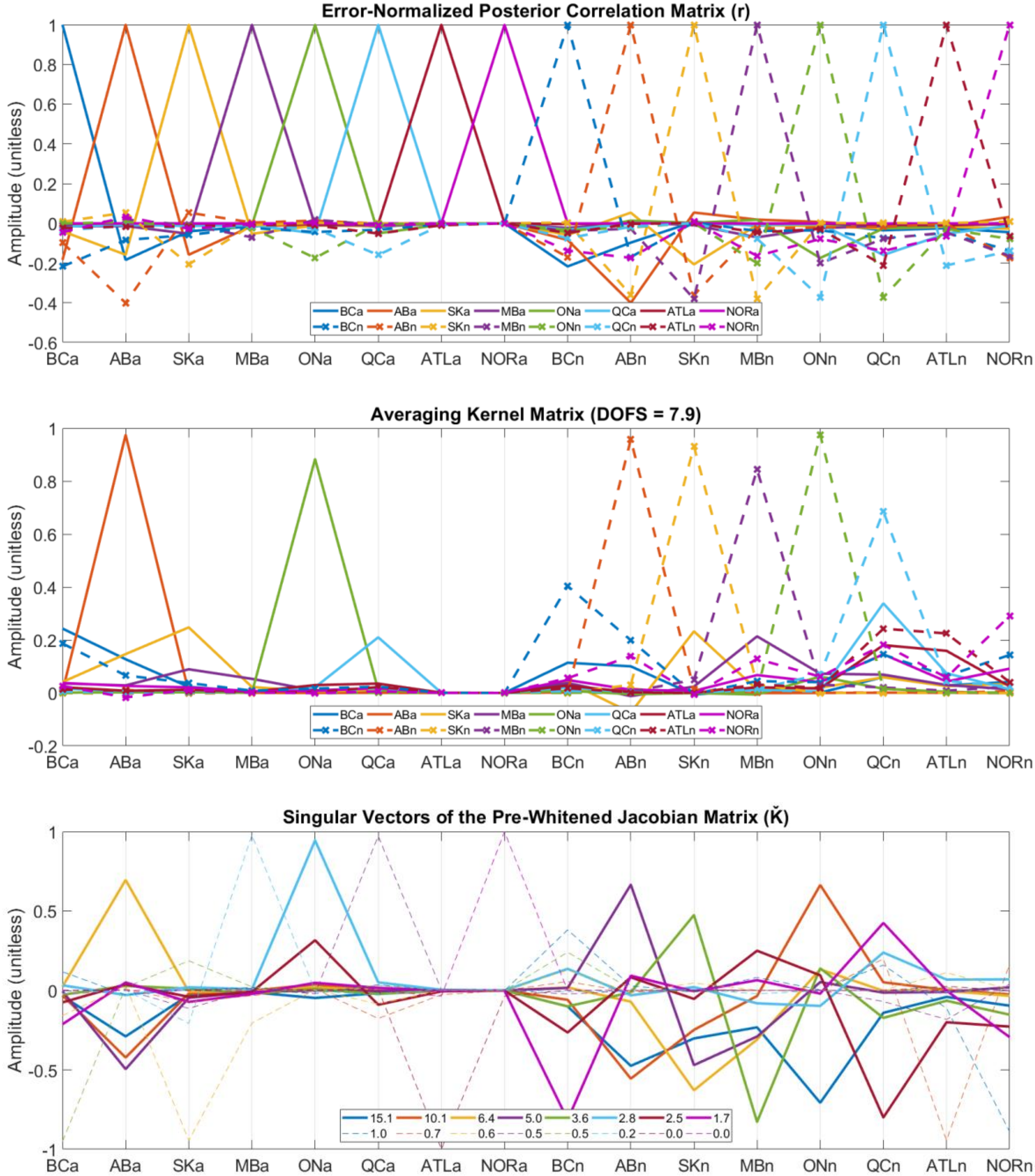

Figure S5: Similar to Fig. S4 for the 16 state vector provincial inversion. The DOFS from the averaging kernel matrix are 7.9, which are consistent with the number of singular values greater than unity in the pre-whitened jacobian matrix (8 in total). Note the difference in meaning of dashed lines between panels: in the top two panels, solid and dashed-x lines of the same colour correspond to anthropogenic and natural emissions of the same province to help visualize the capability for disentangling intra-province emissions. In the bottom panel, the singular vectors below the noise (corresponding to singular values less than one) are shown as light-dashed lines, these show which emissions are not constrained by observations.

A possible solution to improving the resolution of the solution is to combine all six years of data to constrain finer scale emissions for the sectoral and provincial inversions. In the presented approach inversions were completed on a yearly basis for six years to produce an average result for 2010–2015. We used the year to year variance as a representation of noise in the system and real yearly variability in the state (due to emissions and/or transport). In principle using more years of data provides a better signal to noise ratio. However, due to the way our state vector elements are defined in the sectoral and provincial inversions, the inverse approach is sensitive to aggregation error and overfitting the fewer number of well-defined state vector elements. Overfitting can be diagnosed using the reduced chi-squared metric:

$$\chi_v^2 = \frac{\chi^2}{v} \cong \frac{\sum \frac{(y-Kx)^2}{So}}{m} \qquad (3s)$$

Where χ_v^2 is the chi-square per degree of freedom ν. Here, the χ^2 is equal to the ratio of the square of the innovation, $\mathbf{S}_O$ is the diagonal element of the observational error covariance matrix corresponding to the same observation, *m* is the number of rows of the observation vector and *n* is the number of state vector elements. A value of χ_v^2 less than one indicates overfitting. We calculate a value of 0.65 for the total vector containing ECCC and GOSAT data which shows evidence of overfitting. Hence using a larger amount of data for the same number of state vector elements would exasperate the issue.

We further test the improvement from combining 6 years of data against independent measurements. To evaluate the differences between using a repeated 1-year approach and a 6-year approach we use independent observations from NOAA ETL aircraft measurements and ECCC CHA in situ surface measurements. Table S2 lists the metrics of agreement that were in Figure 10 and compares them to the results using all 6 years of data simultaneously. For the sectoral inversion, using 6 years of data provides a small improvement in the slope (0.96 vs. 0.91), no improvement in the R^2 (0.20) and degrades the mean bias (–4.3 ppb vs. –0.4 ppb) when comparing to NOAA ETL. Similarly with ECCC CHA data, using 6 years of data for the sectoral inversion provides an improvement in the slope (1.01 vs. 0.98), a slightly worse R^2 (0.43 vs. 0.44) and largely degrades the mean bias comparison (–10.6 ppb vs. –5.9 ppb). For the provincial inversion evaluation at NOAA ETL, using 6 years of data slightly degrades the slope (0.83 vs. 0.86), gives an improvement in the R^2 (0.27 vs. 0.22), and degrades the mean bias (–3.2 ppb vs. –0.5 ppb). The same comparison at ECCC CHA degrades agreement in the slope (0.87 vs. 0.91), improves the R^2 (0.51 vs. 0.47), and improves the mean bias (–4.1 ppb vs. –4.9 ppb). These results show that using 6 years of data for the subnational inversions does not improve agreement against independent data and in many cases degrades the mean bias. The inversion converges on a solution within our defined prior error matrix $\mathbf{S}_O$ with only one year of data. These tests show that using one year of data at a time and calculating the average and variance of the repeated results is reasonable considering the limits of the observation system towards resolve low magnitude emissions.

Table S2: Sensitivity test against independent observations

		NOAA Aircraft Observations ETL			ECCC Surface Observations CHA		
		Slope	R^2	Mean Bias (ppb)	Slope	R^2	Mean Bias (ppb)
Sectoral	Prior	1.15	0.14	–6.8	1.17	0.36	–16.4
	Posterior (1 yr)	0.91	0.20	–0.4	0.98	0.44	–5.9
	Posterior (6 yr)	0.96	0.20	–4.3	1.01	0.43	–10.6
Provincial	Posterior (1 yr)	0.86	0.22	–0.5	0.91	0.47	–4.9
	Posterior (6 yr)	0.83	0.27	–3.2	0.87	0.51	–4.1

We show a comparison of emissions estimates and methods to derive errors for the sectoral inversion in Table S3 and for the provincial inversion in Table S4. The tables compare two error estimates to three sensitivity tests. They show the error estimates from the diagonal elements of the posterior error covariance matrix $\hat{\mathbf{S}}$ and compares to the 1σ variance in the repeated yearly inversions. In both the sectoral and the provincial inversions, the error estimates from the diagonal elements of $\hat{\mathbf{S}}$ often show a more optimistic estimate of the uncertainties. This is likely due to spatial and temporal correlations in the daily-mean ECCC in situ observations and correlations in the GOSAT data that are difficult to quantify in the absence of a full OSSE study. We compare the 1σ variance from repeated yearly inversions from 2010–2015 to the relative change in posterior emissions from using only ECCC data, only GOSAT data, and using 6 years of ECCC+GOSAT data simultaneously. The 1σ yearly variance captures these differences except for state vector elements that were shown to be below the noise and highly correlated with other emissions in Figure S4 and S5. The lack of improvement against the comparison to independent data in Table S2 suggests that this may be suggestive of overfitting. We consider the agreement between the independent use of ECCC and GOSAT data to be a reliable sensitivity test to check the robustness of our results.

Table S3: Sensitivity analysis of the Sectoral (5 state vector) inversion. The error estimates from the posterior error covariance matrix are compared to the yearly variance and the change in emissions using alternative observation vectors.

	Prior ($Tg\ a^{-1}$)	Posterior ($Tg\ a^{-1}$)	Posterior $\hat{S}$ Relative Error (%)	1σ Yearly Variance Relative Error (%)	ECCC-only (% change)	GOS-only (% change)	6-year (% change)
Energy	2.4	3.6	±11	±25	+6	–7	–21
Agriculture	1.0	1.5	±29	±27	–0	–16	**+57**
Waste	0.9	0.6	±31	±47	–1	**+84**	–33
Wetlands	14.0	9.4	±4	±12	–4	+4	+3
Other Natural	0.8	1.7	±20	±56	–37	–6	**+78**

Table S4: Sensitivity analysis of the Provincial (16 state vector) inversion. As per S3 error estimates from the posterior error covariance matrix are compared to the yearly variance and the change in emissions using alternative observation vectors.

	Prior ($Tg\ a^{-1}$)	Posterior ($Tg\ a^{-1}$)	Posterior $\hat{S}$ Relative Error (%)	1σ Yearly Variance Relative Error (%)	ECCC-only (% change)	GOS-only (% change)	6-year (% change)
BCA	0.5	0.8	±24	±41	–26	–11	**+117**
ABA	2.3	3.2	±5	±14	–6	+5	–2
MBA	0.3	0.3	±44	±40	+11	+1	+4
SKA	0.2	0.2	±49	±26	–3	+5	**+33**
ONA	0.5	0.4	±20	±25	–1	+**27**	+2
QCA	0.4	0.3	±51	±42	–11	+17	+23
ATLA	0.0	0.0	±52	±4	+1	+3	–9
NORA	0.0	0.0	±50	±1	0	0	+1
BCN	0.4	0.5	±35	±53	–7	+13	**–80**
ABN	2.4	1.9	±14	±34	+**59**	–30	–26
MBN	1.6	0.7	±31	±46	+7	+4	–4
SKN	1.5	1.4	±21	±33	+13	–9	–13
ONN	3.5	0.9	±38	±57	+9	+13	–18
QCN	1.6	1.2	±38	±38	+15	–30	–37
ATLN	0.7	0.8	±40	±27	–**36**	+24	**+58**
NORN	0.7	1.9	±14	±35	–**45**	–3	**+73**

References

Heald, C. L., Jacob, D. J., Jones, D. B. A., Palmer, P. I., Logan, J. A., Streets, D. G., 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: COMPARATIVE INVERSE ANALYSIS, J. Geophys. Res., 109(D23), doi:10.1029/2004JD005185, 2004.

Hutchinson, M. F., McKenney, D. W., Lawrence, K., Pedlar, J. H., Hopkinson, R. F., Milewska, E. and Papadopol, P.: Development and Testing of Canada-Wide Interpolated Spatial Models of Daily Minimum–Maximum Temperature and Precipitation for 1961–2003, Journal of Applied Meteorology and Climatology, 48(4), 725–741, doi:10.1175/2008JAMC1979.1, 2009.

Miller, S. M., Commane, R., Melton, J. R., Andrews, A. E., Benmergui, J., Dlugokencky, E. J., Janssens-Maenhout, G., Michalak, A. M., Sweeney, C. and Worthy, D. E. J.: Evaluation of wetland methane emissions across North America using atmospheric data and inverse modeling, Biogeosciences, 13(4), 1329–1339, doi:10.5194/bg-13-1329-2016, 2016.

Stanevich, I., Jones, D. B. A., Strong, K., Keller, M., Henze, D. K., Parker, R. J., Boesch, H., Wunch, D., Notholt, J., Petri, C., Warneke, T., Sussmann, R., Schneider, M., Hase, F., Kivi, R., Deutscher, N. M., Velazco, V. A., Walker, K. A. and Deng, F.: Characterizing model errors in chemical transport modelling of methane: Using GOSAT XCH4 data with weak constraint four-dimensional variational data assimilation, preprint, Gases/Atmospheric Modelling/Troposphere/Chemistry (chemical composition and reactions)., 2019.

Zona, D., Gioli, B., Commane, R., Lindaas, J., Wofsy, S. C., Miller, C. E., Dinardo, S. J., Dengel, S., Sweeney, C., Karion, A., Chang, R. Y.-W., Henderson, J. M., Murphy, P. C., Goodrich, J. P., Moreaux, V., Liljedahl, A., Watts, J. D., Kimball, J. S., Lipson, D. A. and Oechel, W. C.: Cold season emissions dominate the Arctic tundra methane budget, Proc Natl Acad Sci USA, 113(1), 40–45, doi:10.1073/pnas.1516017113, 2016.