

We thank the reviewers for their helpful comments that have improved our paper.

Anonymous Referee #1

Received and published: 1 February 2018

<General Comments> Satellite observation is the most important method to provide decade long and global data of anthropogenic methane emission. In 2010s, GOSAT is the only satellite to provide column CH₄ density but its spatial coverage is limited and a single data has large fluctuation. Therefore, statistical analysis is important. In addition, selection of reference point together with emission point or estimation of the background is critical for quantitative analysis. This paper proposed and described new analytical method clearly. The trend data from different emission source by this work is innovative. It is worth publication after minor revision.

<Specific Comments>

(1) Page 3, Line 1, Proxy method. “The proxy method uses prior knowledge of carbon dioxide” Brief description of prior knowledge is needed. Does it include seasonal variation plus annual growth only or anomalies such as caused by heat wave in 2010 and El Nino?

We added the description.

(2) Page 3, Line 21 “instrument error” Page 3 line 23 “instrument noise” Page 3 Line 25 “Local instrument bias” Supplemental material, Page 3, Line 12, “instrument error” Supplemental material, Page 4, Figure S2 caption, “instrument noises”

Do these terms have the same meaning? TANSO-FTS onboard GOSAT has two major random error sources and there are also several systematic errors. Detector noise and pointing fluctuation in 4 sec to acquire single interferogram creates random noise. Radiometric calibration error due to degradation after launch, spectral calibration and spectral line shape error, radiative transfer calculation error, molecule parameter cause systematic bias.

Here instrument error/noise means random error, and instrument bias means systematic error.

We updated the text accordingly.

(3) Page 5, Line 26, Gulf of Mexico observation by GOSAT “are not directly detectable by GOSAT because the nadir measurements are only over land” It should be described more accurately. Over ocean including Gulf of Mexico, GOSAT can observe column averaged CH₄ using glint mode by tracking specular reflection point but the data are sparser.

We added this in the text: “Glint observations are available over the ocean but are much sparser.”

<Technical Corrections> (1) Supplemental material, Page 7, Figure S5, Description of blue, black and red lines in the figure caption will help readers’ understanding even though they are described in the text.

We added the description in the figure caption.

Anonymous Referee #2

Received and published: 20 March 2018

General comments —————

This paper is mostly an update of the Turner et al. paper of 2016 aiming at estimating trends in methane emissions over North America as inferred from inversion of GOSAT satellite atmospheric weighted columns. Basically, two more years of data are assimilated and the method to estimate the background is revised. The methodology used here has been criticized in details in Bruhwiler et al (JGR, 2017), main arguments being a too short time window for data assimilation making the GOSAT trends sensitive for instance to changes in atmospheric transport, seasonal biases in GOSAT data towards summer months (less clouds = more data), and influence of the choice of the background. In this paper, the authors address only partly these criticisms and add an original sectorial analysis of the inferred trend.

My main concern on this paper is that it does not fully address the extensive criticisms made in Bruhwiler et al. A window of 6 years is still very short to make a robust trend analysis for a species like methane with a 9-year lifetime and I am not sure that adding 23 months compared to Turner is enough. The inferred trend is very noisy (0.2 ± 0.7 ppb a⁻¹) and moving to percentages is a bit misleading considering the very low value inferred especially when considering the remaining bias of GOSAT data of 4-6 ppb (PVIR4 report from Buchwitz et al., 2016). Nothing seems to be done for the seasonal bias and only the question of backgrounds is addressed in detail. The authors may consider looking at the Cressot et al paper (ACP 2016) on the detectability of emissions at regional scale to figure that trends are very hard to detect with the not-so-dense and biased GOSAT data. The text also lack precision in many places (see specific comments). Some part of the work is interesting such as the methodology for the sectorial analysis but I think that more time is needed to extend the timeseries and be able to use this approach more safely and provide a reliable update of the Turner et al. paper addressing all the issues raised since they published it.

Reviewer #2 picks up on the criticisms made by Bruhwiler et al. (2017) of the Turner et al. (2016) paper. Our work has made an honest attempt to address these criticisms (definition of the background, length of the record, inconsistency with surface network) and we have made a good-faith effort to further address the reviewer's concerns in revision. Point-by-point responses are below. It is very doubtful that we can fully satisfy the reviewer but we hope that he/she will let us "agree to disagree" in an open spirit and carry out the discussion in the literature. In answer to the criticisms above:

- **To dismiss the paper as simply an update to Turner et al with two more years of data and different definition of background seems very unfair. This paper adds (1) sectoral breakdown, (2) Canada and Mexico, (3) validation with TCCON, (4) relations of trends to activity data, (5) examination of consistency in the trend with surface sites. These are important advances. In addition, we have extended the GOSAT trend analysis till**

year 2016, by adding an additional year of analysis to what was submitted in the first version to ACPD.

- Not clear why lifetime is relevant here. The 9-year lifetime of methane is not relevant to the length of the record needed to diagnose a trend. The relevant time scale for a trend in enhancement over background is how long it takes for the enhancement signal to dilute into the background - and that time scale is a few weeks.
- The bias in GOSAT data is not relevant since there is no reason to think that it would affect the local background and the enhancement differently (we now make that point in the revised manuscript).
- Seasonal bias in the GOSAT data only affects Canada as stated in the text. This doesn't invalidate the trend analysis; it just means that (for Canada) the trend is more of a summertime one.

We expanded our discussion to address criticisms made in Bruhwiler et al. (see response below and in Specific comments).

Regarding the length of the GOSAT record, we now expand our analysis to 2016 (latest available GOSAT data) for what is now a 7-year record. The addition of 2016 supports the trend previously observed for 2010-2015. We think that a window of 7 years is reasonable to infer methane trends. Lifetime has little to do with it. Methane trend analysis has previously been done using SCIAMACHY with 7 years of data, e.g., Frankenberg et al. (2011). We agree a longer time period would lead to more robust results. We have mentioned this limitation in the conclusion.

Frankenberg, C., I. Aben, P. Bergamaschi, E. J. Dlugokencky, R. van Hees, S. Houweling, P. van der Meer, R. Snel, and P. Tol (2011), Global column-averaged methane mixing ratios from 2003 to 2009 as derived from SCIAMACHY: Trends and variability, J. Geophys. Res., 116, D04302, doi:10.1029/2010JD014849.

Regarding the inferred trend (0.25 ± 0.48 ppb a⁻¹ with the addition of 2016 data). This trend is significant but it is indeed noisy, which is precisely why we move our analysis to the aggregated enhancement. Our conclusion is based on the trends in the aggregated enhancement. We edited our text accordingly (see the response in Specific comment).

The bias in GOSAT data is removed in our approach. We define our local background as low percentiles, and the resulting local enhancement is unbiased as we stated in the text: "This approach removes any local instrument bias (systematic error) because the bias is expected to similarly affect all percentiles of the methane observations."

Regarding GOSAT seasonal bias, we already mentioned this in the text: "GOSAT observes in all seasons with near-uniform frequency south of 45°N (CONUS and Mexico), but observations further north (Canada) are biased toward summer. The number of successful retrievals over

Canada is 2-3 times less in winter than in summer (see Supplemental Material).” We now mention this explicitly again in the conclusion “... variations in wetland areal extent, though this trend is weighted toward summer because of the seasonal bias in observation frequency (less observations in winter)”.

Cressot et al. (2016) found GOSAT performed better than surface observations and IASI for detecting methane anomalies at global and regional scales. The poor rate to detect the methane anomalies at the regional scale as stated by Cressot et al. may be due to that (1) they were conservative to estimate the noise (possibly leading to its overestimation); (2) the time period of GOSAT is 2009-2011, a time period with relatively flat methane signal as seen in our trend analysis. We now mention Cressot et al. in the text.

We do not agree that our text lacks precision in many places (there are 5 places in specific comments related to precision, which are all minor).

Specific comments —————

P2 - L10: you may also mention decreasing BBG and quote Worden et al (2018) paper in Nature Comm.

We have added this in the text.

P2 – l14: please add that, contrary to surface networks, the GOSAT data have residual biases of 4-6 ppb as stated in the PVIR reports (Buchwitz et al). Also, the spatial coverage is enhanced by GOSAT but the number of clear-sky scenes is so huge, and temporal coverage is probably smaller than continuous surface in-situ measurements

We now mention the bias in Methods:

“The resulting GOSAT XCH₄ data have been validated against the ground-based Total Carbon Column Observing Network (TCCON), and found to be of high quality with a single-scene precision of 0.7% (random error) and a systematic error of 4-6 ppb (Parker et al., 2015; Buchwitz et al., 2015, 2016).”

We have deleted the text about spatial coverage being enhanced by GOSAT.

P2 – l16-17: there are other reason in Bruhwiler’s paper to be added here: impact variations of atmospheric transport linked to short-term window of assimilated data (6- 7 years is still short to me), seasonal bias of GOSAT data. You cannot only pickup what arrange you and have to address all limitations raised by previous work.

Here we updated the text as

“...been biased by the brevity of the GOSAT record, atmospheric transport variability, seasonal bias in GOSAT sampling frequency, and the use of Pacific data as background.”

We actually addressed all the limitations later in the text. We now expand these discussions (also see response below).

P2 – I19 : This is not precise enough. short-term trend may depend on local to regional conditions but longer trend is a global signal and one station is enough to get it.

We updated the text as “...local or regional trend detectability from the surface data may be limited by their sparsity”.

P2 – I20: lack of precision. which version of EDGAR ? 4.2 has too large emission and trend especially in Asia. EDGAR4.3.2 partly corrects this issue. Please be more precise. Also, the dependency to prior assumption may be loose or tight depending on the associated error structure.

Asia is not relevant here, and EDGAR4.3.2 has its own problems, but we deleted that text as non-essential.

P2 - I22-23 : Adding 2 years compared to Turner et al., 2016 does not convince me that the time period will be long enough to overcome the issues raised in Bruhwiler et al (2017). 10 years (~ methane lifetime) would be a minimum to start extracting reliable information on methane trends to my opinion.

Not clear why lifetime is relevant here. If it was we couldn't say anything about trends of CO2 on decadal scales...

P3 – I6 : 0.7% is 12 ppb. Are you talking of random error or systematic errors ? please be more precise as systematic errors (estimated at 4-6 ppb from PVIR report of Buchwitz et al) ultimately limit the use of GOSAT to estimate emission trends of a few ppb/yr or less.

As we said in the text, it's instrument precision so here we mean random error. We updated the text as “...a single-scene precision of 0.7% (random error) ...”. Systematic errors (or bias) are irrelevant for methane enhancement in our approach. We have already discussed this in the text (see P3, L24-26): “... This approach removes any local instrument bias because the bias can be expected to similarly affect all percentiles of the methane observations.”

P3 – I9-10 : the opposite is clearly shown in Bruhwiler's paper which surface emission changes appear only weakly sensitive to surface emissions. Please rephrase.

Replaced “given source region” by “strong source region”

P3 – I16 : “ the low (10th -25th) percentiles of the deseasonalized GOSAT methane Observations” unclear to me. Which observations? on which area? how is it specific to the 0.5x0.5 location. Please rephrase to be more clear and explain what you do exactly.

We updated the text as

“Here we define local background methane for a given CONUS location (0.5°x0.5° grid cell, typically including a single repeated GOSAT measurement location) and for a given year as the low (10th-25th) percentiles of the deseasonalized GOSAT methane observations in the given 0.5°x0.5° grid cell and year, ...”.

P3 – I20 : how did you choose these upper bound 25th percentile ? did you try other range and how sensitive is this choice on your results ?

We consider values below 25th percentile to be low percentiles. Results are only weakly sensitive to the choice of different ranges as stated in the text. We also did a sensitivity test on this (see Fig. S5 in supplementary material).

P4 – I3-4 : the trend on enhancements does not seem to be significant considering the error bars. Please provide more quantitative results on this.

Here significance for a single site is not relevant because we only focused on the aggregated enhancement trends. We updated the text as “... although the error standard deviations defined by the ranges of the 10th-25th percentiles are large and the trends at this single site are significant ($p = 0.07$). Below we will use enhancement statistics aggregated over a large number of sites in order to reduce that uncertainty and quantify trends.”

P4-I8: Is EDGAR 4.3.2 very different than 4.2 over North and central America ?

No. They are similar. We added “Compared to EDGAR v4.2, the more recent EDGAR v4.3.2 (Janssens-Maenhout et al., 2017) has similar national totals and spatial patterns for non-oil/gas anthropogenic methane emissions.”

P4 – I19-20: did you try not doing so as it reduces largely the number of wetland- dominated pixels.

Using either wetland inventory alone would bias our results because they differ significantly in space (see Supplement Material). We updated the text as

“Wetland-dominated areas determined by the WETCHIMP mean and WetCHARTs inventories differ significantly (see Supplemental Material). Using either of the two inventories alone may bias our results, and thus we conservatively require wetland-dominated areas to be determined as such in both inventories.”

P4 – I24-25 : what about atmospheric transport ? summing only columns above the high emitting pixels does not account for transport and the potential plume sampling by other GOSAT data. It would be worth mentioning this to clarify what is it you do here.

As we mentioned earlier, the local background range (10th-25th percentiles) accounts for atmospheric transport. We updated the text as:

“To account for background variation due to atmospheric transport, the summation in Equation (1) is conducted for 1000 Monte Carlo realizations where the background XCH_4,b,i for each grid cell and for individual years is obtained by random sampling of percentiles in the 10th-25th range.”

P5 – I15 : are they all supposed independent ? How robust is this significance ? Although tighter than in Tiuner et al., the PDF is still broad with a sigma of 0.66

With addition of 2016 the sigma has decreased to 0.48. Each local enhancement is calculated independently. Significance is indicated by p-value <0.01. We agree the sigma is broad, but that's why we move our analysis to the aggregated enhancement that significantly reduced the uncertainty. We now mention that in the text:

"Below we will use the aggregated enhancement (Equation 1) to infer the trends and reduce the uncertainty."

P5 – I16-17 : 10.8 ppb enhancement might be due to other causes as stated in Bruhwiler et al. Please mention that this is a maximum and which of the causes raised in Bruhwiler's paper may still apply here. I strongly recommend to add in the following that inferred numbers are maximum number, potentially smaller because of limitations raised in Bruhwiler's paper.

We do not think that 10.8 ppb is a maximum. As we discussed earlier, local background is statistically not affected by random error, and should account for transport variability to some extent. Janardanan et al. (2017) found a large number of observed and simulated enhancements in the range of 10 to 20 ppb in North America using GOSAT observations and a Lagrangian particle dispersion model. We updated the text accordingly. "The mean 2010 methane enhancement for high-emitting grid cells in CONUS relative to local background is 10.8 ppb, comparable to that found by Janardanan et al. (2017)."

Figure 3 : just ot be sure : the grey bars for wetwhimp and Bloom do reflect the totals for the common pixels ? if not please correct.

No, those are national totals as stated in the figure title and caption, but we now add the totals for common pixels.

P5 – I29 : what about pixelrIs emitting a lot but with a balanced share of emissions (ivestock & oil&gas) ? Yopur method should discard them. How does it influence your results ?

Here for national trends we do not discard those pixels. We only discard them when we do sectoral trend analysis. So it does not influence national trends. For sectoral analysis, it does not make sense to use grid cells that are not dominated by any source sector. We updated the text as

"Here the trends are calculated for the summed enhancement Δ in Equation (1) calculated for individual years and for high-emitting grid cells in individual countries or high-emitting sectors."

P6 – I1 : replace ambiguous "interannual" by "year-to-year" or equivalent

We replaced it by "year-to-year".

P6 – I14-15 : US oil&gas activity (figure 5) show stalled variations in 2014-15 whereas your analysis find a fast increase from 10 to 20% (figure 4). Isn't that contradictory ? Please comment in the main text.

No, that's not contradictory. In the text, we have already mentioned that "..., though production rate is not necessarily a predictor of emissions (Peischl et al., 2015)."

P6 – I20-22 : how do emission factors for swine and cattle compare ? it would be worth to add the cattle number in comparison with the swine emission factor range given. Is this increase really significant for methane emissions (uncertain range of small emissions of 0.01-0.2 Tg/yr)?

We don't think it's worth to compare cattle population with swine emission factors as it will not convey any new information. We already provided emissions in Midwest from enteric fermentation (cattle) and manure management (swine): "These grid cells emit 0.95 Tg CH₄ a⁻¹ from enteric fermentation (cattle) and 0.55 Tg CH₄ a⁻¹ from manure management (swine) according to the gridded EPA inventory (Maasakkers et al., 2016)."

This increase is significant and not small for Midwest as we stated in the text that the trend largely reflects Midwest. The uncertainty range here is due to the choice of emission factors.

We added "A larger value of the emission factors is more likely. The emission may increase ..."

P6 – I28 : interannual → year-to-year or equivalent

We replaced "interannual" by "year-to-year".

P6 – I30 : "wetland areal extent" : this is very controversial and there is no consensus of wetland extent and their evolution (see Poulter et al., 2017 also). Please mention this controversy here.

We updated the text as

"..., though the definition of wetland areal extent may vary significantly (Poulter et al., 2017). Here the WetCHARTs extended ensemble used GLOBCOVER land cover data (Bontemps et al., 2011) and the Global Lakes and Wetlands Database (GLWD Lehner and Dölla, 2004) to represent spatial wetland extent, and ERA-interim precipitation to account for temporal wetland extent (Bloom et al., 2017)."

P6 – I33 : please note in the text that the "trend" you infer for CONUS is mostly after 2012 ("total" line on figure 4). The inversions reported in Bruhwiler 2017 stop in 2012. Please mention these two elements in the main text. Again, waiting more time to get longer time series would avoid limitations in trend analysis. . .

We removed Bruhwiler et al. (2017) here. It's irrelevant for our residual test. We updated the text accordingly.

P7 – I1 : Are the stations shown on figure 7 used in the CT inversion? please precise. Do some other surface stations not shown here show some trend? If not please mention it at it reinforce your point.

Yes, they are used in the CTL inversion. We updated the text accordingly.

P7 – l8-9 : But this does not discard the possibility that the trend found in your paper is not due to emissions but to other factors as stated in Bruhwiler's paper. Please mention this here as well.
The detected trends have already accounted for the factors as stated in Bruhwiler's paper.

P7-l12 : I recommend to change " significant increase in US methane emissions" into "significant increase in total US methane emissions after 2012"
We changed the text accordingly.

Conclusion : please develop more the main limitations of your study either at the end of result section or in the conclusion.
We added limitations of our study in the text accordingly.

What about OH changes in your method ? you do not mention your assumptions on OH. Please specify them somewhere in the text.
OH is irrelevant here as we already discussed in the text (P3, L28-30): "Any trends in OH concentrations would also not affect the enhancement because the lifetime of methane against oxidation is 9-10 years (Prather et al., 2012; Kirschke et al., 2013), very long compared to the timescale for ventilation from the source region."