Atmos. Chem. Phys. Discuss.,
doi:10.5194/acp-2017-72-RC2, 2017

Atmospheric Chemistry and Physics
Open Access
Discussions

Interactive comment on "Estimating regional scale methane flux and budgets using CARVE aircraft measurements over Alaska" *by* Sean Hartery et al.

Anonymous Referee #1

Received and published: 28 March 2017

Review of Hartery et al: Estimating regional scale methane flux and budgets using CARVE aircraft measurements over Alaska

Summary:

This paper uses aircraft observations of in situ trace gas concentrations and thermodynamics to constrain Lagrangian-transport-inversions of CH4 flux during a campaign over the Alaskan region as part of the CARVE project. More specifically, spatially resolved fluxes of biogenic CH4 flux during the growing seasons between 2012-2014 have been derived. The work uses a rich measurement dataset that has been carefully obtained and calibrated to a high standard.

After a description of the measured dataset, the work goes on to describe a flux derivation method using footprint sensitivities and inversions calculated using WRF-STILT.

These methods are interesting and adapt existing conventional optimal flux transport-inversion approaches to attempt to link spatially-resolved flux to soil temperature and depth as a function of time. Such an attempt is highly challenging and this paper is a trailblazer in terms of attempting to do this from a long-term aircraft campaign.

That said, I do have some major questions and concerns about the methods and the conclusions drawn from them. At the moment, the flux work seems to be predicated on assumption after assumption, followed by a procedure of arbitrarily discarding data (over half of it in the end), temporal and spatial averaging, and averaging some more, and then discarding more data, before using only 68 (or is it 146?) measured aircraft profiles (averaged to mixed-layer partial columns) to obtain biogenic growing season flux as monthly-averages over 3 years. This equates to about 4.5 profiles per month (though this is only a rough average calculation as there is no information on the sampling statistics per month other than the total number of profiles used across the whole study). I find it hard to accept that such a limited dataset can derive robust flux statistics representative of regional monthly means, especially where the only uncertainty given on the fluxes is the standard error on the mean of the already monthly-averaged fluxes. Such an error statistic is useless – it neither represents the systematic error associated with the method, nor represents the natural variability of CH4 flux in the region between each independent flux calculation. Instead, it convolves the two with no possible determination of which dominates.

What I instead believe the authors have here are a set of independent flux retrievals (one per profile/flight) and independent posterior flux uncertainties using their method. The extensive averaging of these independent retrievals in the paper make it very difficult to assess the performance of the method; and the current error budget is meaningless. Until I can see more about the performance of the method and the statistics of flux retrieval-by-retrieval, I don't have any confidence in the conclusions and discussion later on (e.g. on soil temperature relationships).

The paper does present some very interesting analysis and I believe there is some really great science to come from the work. Therefore, I definitely recommend publication in ACP as the methane flux problem is a key topic in atmospheric and geoscience at the moment and this paper represents an exciting way to make use of long-term aircraft datasets. However, I do have to recommend major revisions at present as I think the way the analysis has been done needs to be extensively rewired to present more meaningful data that the reader can more transparently assess, especially with regard to independent flux calculations and uncertainty and error budgets. I'll try to give some specific constructive guidance on this below, which I hope would help will turn a questionable analysis into something really quite interesting and useful. My review won't discuss the flux-soil relationships as until I can see the results from the revisions below I don't feel I have enough information to assess the later aspects of the paper.

Specific comments:

1. Abstract, line 9 etc: It is very important to note that the whole analysis of the paper derives "net emissions", not "emissions", e.g. the statement that ". . ..Boreal emissions. . ..accounted for the remainder of the emissions" should contain the word "net" as the study does not address local or regional sinks (albeit potentially small). This important point needs to be kept in mind throughout the paper when discussing flux and needs to be very clear to the reader early on.

2. P.2. line 31-32: I agree that scaling local fluxes to regions is challenging, even in areas where it may be argued it is possible to derive meaningful regional statistical parameterizations (such as this paper sets out to do). However, there are some studies (not currently cited) that have attempted to do this (also at high latitudes) using a combination of aircraft, chamber and eddy covariance measurements. It would be useful to discuss and cite such work in this paper (as it seems very relevant to the introduction, and later discussion, in this paper). Please see: O'Shea, S. J et al.: Methane and carbon dioxide fluxes and their regional scalability for the European Arctic wetlands during the MAMM project in summer 2012, Atmos. Chem. Phys., 14, 13159-13174, doi:10.5194/acp-14-13159-2014, 2014.

3. P.5: Footprint method: The authors have used WRF-STILT to derive a surface sensitivity footprint. The footprint seems to have been derived using a grid with frequency/counts equal to the summed incidences of residence of 500 reverse-Lagrangian particles per measurement in the “lower half” of the boundary layer. This is a can of worms and it is glossed over far too quickly here, and later on. I guess this definition is a bit arbitrary and I have plenty of sympathy with it, as defining surface influence in Lagrangian trajectories is very difficult to quantify. However, I would have liked to have seen more discussion on the uncertainty or sensitivity that this arbitrary definition of surface contact may have on the derived flux footprint. The authors at the very least need to clearly acknowledge that there may be an unquantified source of model transport error coupled with the use of their lower half PBL definition; and - more usefully - they could examine footprint sensitivity to different surface contact definitions (e.g. as other percentages of the boundary layer height). I raise this simply because there is no basis to believe that the lower half of the PBL is in dynamical contact with the surface, especially in enormously diurnally-variant boundary layers such as those in Arctic spring. It could be argued that the diurnal ventilation and contraction (i.e. entrainment and detrainment) of the boundary layer could skew the footprint derived here to one more biased toward representing daytime flux (as daytime PBL trajectories that are isentropically detrained into a PBL-residual layer at night-time would not be counted at night-time in the footprint using the authors’ method along the 5-day history used). This could have implications for the fluxes that are derived and their biogenic interpretation and quantification. I have no major problem with the use of an arbitrary definition such as this, as it attempts to do the best it can with the information available, but I do think the reader needs to be made more aware of the potential limitations and issues with it. And some of this may be quantifiable with a sensitivity analysis to PBL depth-contact versus footprint. Without this, I would have some outstanding questions about the numerical validity of the later flux calculations and what they truly represent. I liked the discussion on what the footprints represent more globally on Page 6 but more needs to be added. In summary, I suggest an easy (making it clear as to the limitations) fix and a bigger effort (sensitivity) fix to help those following the work to make their own informed judgment. Page 9 lines 1-5 seem to suggest that some effort has been made to examine footprint sensitivity to the inclusion (or not) of discarded profiles so perhaps it could be simple and useful to add some of this to the paper to convince the reader that there is no important bias (even if as an Appendix?).

4. Related to the comment above, were 500 particles released per singular Picarro measurement, or were 500 particle released per mixed-layer column used in the later inversion? If the latter, were these equally spaced with altitude in the mixed layer? A sentence to make this clear would help.

5. P.6. Section 3.3 - Mixed layers: This relates very closely to the comment above. If mixed layer height (i.e. instantaneous local PBL height plus residual layer height) is used to derive the fluxes later on, how do you differentiate between the true PBL (which is in contact with the local surface) and the residual layer (which may represent the previous day's local PBL ventilation, or advection from a more distant regional source)? This is very confusing. The method states (page 6, line 26) that the authors use the integrated mixed-layer partial column concentration (not the vertically-resolved measurements) to calculate flux in Section 3.4. This must then surely convolve any local emissions (in the true local PBL) and non-local emission (in any residual layer). The authors later go to great lengths to show that any residual layer is not influence by long-range (non-regional) transport but this does not solve the problem of varying airmass histories for the true PBL versus the residual layer when these get clumped into a partial column for the purposes of the inversion. When this singular column concentration is used in STILT and coupled to the footprint described in Section 3.1 (and the issues alluded to in the previous comment), it seems impossible to deconvolve spatially-resolved flux with any true or traceable footprint sensitivity as the column represents an unknown mix of local and non-local surface contact. Again, I have a lot of sympathy (more than it may sound like) with the approach and doing the best job possible with adjoint models. But there is currently no awareness or clarity of these issues in the text which would alert the reader to the challenges and limitations in the approach. Page 7 goes on to demonstrate that 50

6. P.8. Using CO as a tracer for combustion CH4 sources: What about pure fugitive emissions of thermogenic CH4 (where there is no combustion)? This could conceivably lead to an over-estimate of biogenic flux if the remaining profiles contain any significant non-biogenic CH4 from sources not co-emitted with (potentially large fluxes of) CO. I see that only 9 of the profiles were discarded by this definition and the analysis of sensitivity by including them to derive a different flux is useful. This style of analysis starts to give the reader what they need to assess things. With this in mind, I have a few suggestions below to consider:

Suggested principal corrections:

7. The monthly-averages given may be hiding a wealth of useful data. A time series of biogenic flux (as an area-normalised quanitity – i.e. as biogenic flux per unit time per unit area) for each independent flux retrieval (I believe there are 68 of these?) would be useful. The posterior flux uncertainty (that STILT should yield as output) could be plotted as an error bar on each data point on such a plot.

8. Error/uncertainty analysis: As discussed above, the current tolerance placed on the derived fluxes is meaningless and does not represent either systematic error (flux inversion uncertainty) or natural variability. The seasonal trend plotted in Figures 4 and 5 do not convince me that natural regional variability dominates the mean as this could simply be a manifestation of the changing northern hemispheric seasonal background and priors used. I would suggest that the posterior flux uncertainty of independent retrievals/footprints is used instead as this captures the uncertainty on each retrieval. And then, rather than a standard error on the mean flux (taken from the spread of the averaged inverted fluxes), which would clearly be an incorrect (and much reduced) error, I would recommend quoting the posterior flux uncertainty (calculated as an average of the posterior uncertainties across all inversion that contribute to the final monthly mean). It would be important to give the average of the posterior uncertainties (not their standard deviation or standard error), to yield a meaningful uncertainty on the monthly mean flux. Such an error will still convolve natural flux variability but at least it would be a more accurate measure of the systematic uncertainty in the method used. This should replace the shading (error bars) used in Figure 4.

9 A table of the derived mean fluxes (and their corrected uncertainties) could be presented, which displays flux with and without the sensitivities/assumptions that have been used (i.e. masking seas and mountains, removing elevated CO profiles). These fluxes are currently in the body of the text, making it hard to compare them. And perhaps a new flux could be calculated where all 248 profiles are used in the inversion? A table would add (at a glance) the comparison between these sensitivities. This latter sensitivity test would then give the reader all the information they need to compare the information and make their own judgement about what they trust and the implications of the assumptions used.

Technical corrections:

10. 1/ Title – hyphenate “regional-scale” 11. 2/ Abstract line 1: change to “…gas but its emissions….” Not “their”. 12. 3/ P.1. Line 10: change to “…CH4 flux was…” or “…CH4 fluxes were…” - There seems to be some confusion between the use of singular and plural references throughout when referring to flux and fluxes, respectively. Please check as I won’t list any further instances. 13. 4/ P.2 line 5: change to “40 °N” and check throughout. “40N” is not acceptable. 14. P.2. Line 9: century should always be capitalized when referring to a specific century. 15. P.4. Line 6: add space to “195 K”. 16. P. 8, lime 11: typo - change to “weighting”.

Interactive comment on Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2017-72, 2017.