Author responses and reviewer comments in regards to:

Hartery S. et al.: "*Estimating regional-scale methane flux and budgets using CARVE aircraft measurements over Alaska*," Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-72, in review, 2017.

Table of Contents

Response to Reviewer 1

1 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 dont 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. Ill 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 wont discuss the flux-soil relationships as until I can see the results from the revisions below I dont feel I have enough information to assess the later aspects of the paper.

We thank the reviewer for their very thorough reading of the paper and constructive comments and suggestions. We believe that the assumptions made in our analysis are not as extreme as the reviewer suggests and that our responses below address some of the confusion caused by our initial descriptions. We have added extra text in the manuscript to clarify these points. We agree that quantifying the uncertainties is very challenging. Since our analysis method is not a Bayesian inversion, rather we simply divide our observed column methane enhancements by the modelled column footprint sensitivity to estimate net fluxes, the model does not actually provide a posterior flux uncertainty. However, we now explore the uncertainties associated with our analysis through bootstrapping / Monte Carlo. This is explained in more detail below and is described in a new section (Sect. 3.7) of the manuscript.

To address the reviewer's suggestions, we have added a Supplement which includes tables listing all the profiles, whether they were included in our analysis, and the reason for exclusion if relevant, as well as monthly means and uncertainties calculated using different methods. It also includes additional figures showing the flux calculated from individual profiles with uncertainties as described above, the resulting residual when compared to the monthly means and background CH4 observed from the aircraft, at the BRW observatory and inferred from the CRV tower.

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.

The reviewer brings up a very important point and we have changed the wording throughout the text to reflect both that we are only calculating net emissions and that we are assuming that the emissions are originating from biogenic sources.

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: OShea, 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.

A citation to this study has now been included.

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?).

The reviewer brings up some very important concerns that face any receptor-oriented atmospheric model. The initial description in the manuscript was rather sparse so we hope that the discussion in response to this and the next two points will provide more clarity. The text in Sects. 3.1 and 3.4 of the main text have been edited to include more details.

In an early publication describing the development of STILT, Gerbig et al. (2003) investigated the effects of varying the assumed height at which the air is well-mixed. In their sensitivity analysis, they found no significant changes in their results when the fraction was varied from 10–100% of the boundary layer, although lower fractions resulted in less particles influenced by the surface and therefore increased noise in their results. We acknowledge that their findings were for continental North America using assimilated meteorological data, and may not be representative of Alaska, especially in the spring. Nevertheless, these findings suggest that this parameter is not the greatest source of uncertainty in our analysis. A sentence in the main text now directs the reader to the Supplement where this discussion is now included.

Air detrained into the residual layer at night is not influenced by the surface and should therefore not be included in the footprint sensitivity. As the 500 particles are traced backwards from the receptor point, a number will be in the PBL during the day. As detrainment occurs at night, some fraction of those particles will move into the residual layer while the rest remain in the nocturnal boundary layer. Of the particles remaining in the boundary layer, those that are in the lower half are the ones actually influenced by the surface and therefore included in the footprint sensitivity in the evening hours. We acknowledge, however, that defining the boundary layer at night is difficult and now mention it in this section. Nevertheless, we have fairly high confidence in the representativeness of our footprint sensitivities since past CARVE studies using the same WRF-STILT model coupled to a surface CO2 vegetative flux model have successfully predicted atmospheric CO2 mixing ratios in the mixed layer (Commane et al. 2017, Karion et al. 2016). Since CO2 fluxes are bidirectional and highly diurnal, these studies suggest that the transport model represents surface influences during both day and night reasonably well. This is now discussed in Sect. 3.1 and reference is made to these other studies.

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.

As described in the text, the Picarro measurements below 1 km were aggregated into discrete bins every 5 km in the horizontal and 50 m in the vertical and above 1 km they were aggregated every 5 km in the horizontal and 100 m in the vertical. This resulted in approximately 23 000, 32 000 and 36 000 receptor points in 2012–2014, respectively. At each receptor point, 500 particles were released and traced backwards in time for five days. At each hour backwards in time, a surface sensitivity map is generated that represents the surface that influenced the measurement from that hour. These 120 surface sensitivity maps

(24 h $\times$ 5 d) are added together to determine the cumulative 5-day footprint for a given receptor point. As described on P 8, lines 4–9 of the original submitted manuscript, the total footprint sensitivity associated with a vertical profile was determined by further binning these receptor points every 250 m and calculating the mean footprint at each height. The resulting profile of footprint sensitivities was then summed to determine the total surface influence on a given profile. The text in Sect. 3.4 has been improved to make this more clear.

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 days 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 percent of profiles were discarded from flux analysis because mixed layers could not be reliably separated or that PBL was over-estimated when compared to trace gas profiles. What does over-estimated mean here, and why are the remaining profiles more trustworthy for analysis than those retained? Discarding half of the dataset to select only those data that agree with a contrived and questionable model is a little worrying. I would have liked to know more as to why >50 percent of the dataset is at odds with the model/assumptions and to know what the sensitivity to including the data might be. Can you be confident that the way the remaining 50 percent has been retained hasnt led to some systematic bias in the data and model treatment?

We agree that the partial column is influenced by both local sources in the boundary layer and regional sources in the mixed layer. The important point to note is that the spatial distribution of the footprint sensitivities of air sampled in the boundary and mixed layers will reflect local and regional sources, respectively, thus correctly attributing surface influences to the source regions.

The 273 profiles quoted in the main text represents every instance of the aircraft flying from the surface to >3 km or vice versa. The reviewer will appreciate that not every aircraft transit between these heights properly sampled the entire mixed layer and sufficient height into the free troposphere to properly determine the background CH4 mixing ratio. In addition, there are times when the atmospheric structure does not follow the structure of a classic textbook profile, with mixed layer heights from observations of water vapour, potential temperature and other trace gasses differing, either with each other or with the refractivity method. The cause of these inconsistencies would warrant a study of its own in boundary layer meteorology and is beyond the scope of our study. We therefore limit our calculations to profiles that are well-defined and where we believe we understand the dynamics.

As requested by the reviewer, we now include calculated net fluxes for all identified profiles in the Supplement. However, we do not believe that all of these calculated fluxes are representative of actual net surface emissions and therefore do not include flagged profiles in any of our calculations. The effect of excluding these profiles from the monthly mean cannot be determined. It should be noted that for all 273 profiles, the mean 95% confidence interval (C.I.) for an individual CH4 flux estimate was 8.1 mg m^{-2} d^{-1}, ranging from 0.54–34 mg m^{-2} d^{-1} (2.5%–97.5% percentile). However, for the subset of 146 profiles used in our analysis, the average 95% C.I. was 5.2 mg m^{-2} d^{-1} and ranged from 0.6–15 mg m^{-2} d^{-1}. As stated at the end of Sect. 3.4, we can say that based on the model footprint sensitivities, our sampling region is not skewed by excluding these profiles.

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.

As the reviewer pointed out in the first specific comment, our calculations are only representative of net methane emissions in our study region, to which we attribute to biogenic emissions. These assumptions and other possible sources of methane are now discussed in Sects. 3.4 and 3.5. It should be noted that preliminary results by Floerchinger et al. (2016) on the North Slope of Alaska suggests that fugitive emissions from thermogenic sources are small compared to the net emissions and mostly confined to the region near Prudhoe Bay. Although these results are still preliminary, they indicate that our attribution of emissions to biogenic sources is reasonable.

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.

As discussed at the end of Sect. 3.4, 146 profiles are used to calculate the monthly means. As requested by the reviewer, we now include a time series of the flux calculated from each individual profile (Fig. S1) as well as the residuals from the monthly mean (Fig. S2) in the Supplement. Our inversion method is based on simple linear regression and not a Bayesian analysis, therefore there is no posterior flux uncertainty from the model. Instead, we derive our uncertainties in the estimated flux by bootstrapping each element that goes into the calculation (observed mixing ratio, pressure, temperature and footprint sensitivity at each height bin, as well as background methane mixing ratio and mixed layer height). Over 500 iterations, we arrived at 95% confidence intervals which are shown as the uncertainties in Fig. S1. A detailed description of this analysis is now included in Sect. 3.7 in the revised manuscript.

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.

We thank the reviewer for this useful suggestion. We averaged the 95% confidence intervals as the uncertainty in an individual flux estimate to determine the uncertainty in the monthly mean as well as calculating a footprint weighted average. These results can now be seen for each month and on the growing season budget in Tables S5–S7. We chose instead to represent uncertainty in the monthly mean with the standard deviation weighted by the inverse of the 95% confidence intervals because they tended to result in larger uncertainties in the budget and we felt represented the variability associated with each month's measurements as well as the uncertainty associated with individual flux estimates. We now discuss these findings in the newly

added Sect. 3.7 and have replaced the shading on Fig. 4 to reflect these new calculations.

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 judgment about what they trust and the implications of the assumptions used.

Tables S5–S7 in the Supplement now show the monthly mean net fluxes. As we discuss in the response to point 5 above, we do not believe that the excluded profiles provide meaningful estimates of net fluxes, and therefore do not include them in any monthly mean calculations. However, at the reviewer's suggestion, we now include flux estimates for all the individual profiles so that their quality can be judged by the reader.

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 wont 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.

These have been corrected in the revised manuscript.

References

Floerchinger, C. R., McKain, K., Newberger, T., Handley, P., Wofsy, S. C., Sweeney, C., and Benmergui, J. S.: A methane budget from the North Slope of the Alaskan Brooks Range including arctic tundra and proximate oil and gas operations, American Geophysical Union, San Francisco, CA, 12-16 December 2016, 2016.

Gerbig, C., Lin, J. C., Wofsy, S. C., Daube, B. C., Andrews, A. E., Stephens, B. B., Bakwin, P. S., and Grainger, C. A.: Toward constraining regional-scale fluxes of CO2 with atmospheric observations over a continent: 2. Analysis of COBRA data using a receptor-oriented framework, J. Geophys. Res., 108, n/a–n/a, doi:10.1029/2003JD003770, http://dx.doi.org/10.1029/2003JD003770, 4757, 2003.

Response to Reviewer 2

Summary

Hartery et al. presents CH4 fluxes from the Alaskan wetlands derived from aircraft measurements. The paper examines the relationship between the fluxes and several variables. The paper uses an extensive dataset, which has been presented by several previous studies. However, it is interesting to compare different techniques/models to derive fluxes. Examining the drivers of the CH4 flux at such a large scale is novel. The paper is interesting and well written with relatively few typos. The paper should be suitable for publication, but I do have several queries first.

My main concern is that the manuscript is missing a detailed description of how the uncertainties associated with the fluxes were calculated? How were the uncertainties for the individual components estimated and propagated.

I would expect that the choice of background would have a large impact on the calculated fluxes. However, I am not sure how appropriate the CH4 above the mixed layer is for the background. You are assuming that it is representative of air 5 days upwind of the measurements. Wouldn't you expect significant exchange between the PBL and free troposphere over a 5-day period? Is it possible that the free troposphere and mixed layer air could have different histories (e.g. due to long range transport)? You compare your background to those from Karion et al. (2016) who use 'pacific curtain', however I imagine this will only be valid if there is a west to east airflow. Also is Poker Flats in interior Alaska really observing background air? The choice of 5 day sensitivity footprints seems a bit arbitrary, does changing the length of time used to derive the footprints e.g. 5 to 10 days have much impact on the calculated fluxes? Would you use the same background if you used a 10 day sensitivity footprint in your flux calculation?

Figures 3 to 7 appear to have a lot of problems with missing units, axis labels, legends, etc. It looks like the figures have been corrupted. I don't know if this is just a problem for the pdf reader I am using, but please check them. This didn't seem to be a problem with the initial submission. But it does make reviewing some parts of this current version difficult (particularly sections 4.4 and 4.5).

We thank the reviewer for their thoughtful and detailed comments on the manuscript. The reviewer has insightfully pointed out the greatest uncertainty in our analysis - identifying background CH4 levels. As the reviewer summarizes, we assume that CH4 levels in the free troposphere are representative of background levels in the mixed layer. As we cited in the submitted manuscript, this is not unusual for aircraft studies (Chang et al., 2014; Gatti et al., 2014; Chou et al., 2002) and has been conducted successfully for the CO2 analysis from this campaign (Commane et al., 2017, now cited in the main text). As evidence that this is a valid assumption, we compared our free tropospheric levels with surface observations at the Barrow site in Alaska. This is illustrated in a new figure now included in the Supplement. As requested by the reviewer, backgrounds determined for the CRV tower are also included in this new figure. As can be seen, the Barrow observatory tends to observe

slightly higher mole fractions of CH4, although more variable, than observed in the free troposphere or from the pacific curtain used by Karion et al. One could conclude that the free tropospheric levels are too low, resulting in a higher flux estimate. However, the background CO2 measured at Barrow and the free tropospheric CO2 measured during CARVE, show very little, if any, bias (Commane et al., 2017). To be more transparent about these issues, we have expanded the text in Sect. 4.3 and direct the reader to the new figure in the Supplement so that they can judge our assumptions for themselves.

Another piece of evidence that air in the mixed layer does not mix up into the free troposphere is that vertical profiles of the cumulative 5-day footprint sensitivity determined by WRF-STILT tend towards zero in the free troposphere (e.g. Fig. 3 in the main text). This would suggest that the free troposphere is not being influenced by the surface in our domain, and is therefore not influenced by air from the mixed layer. As the reviewer suggests, however, it is possible that air in the mixed layer and the free troposphere have different transport histories outside of our domain and we now discuss this in Sect. 4.3.1. Unpublished data from measurements taken on the Alaskan Coast Guard aircrafts off the northern coast of Alaska show that average methane profiles are fairly constant from the surface to 7 km, suggesting that free tropospheric levels do represent surface background levels (measurements described by Karion et al., 2013). However, this may not be true for air transported from other directions. It should be noted that the majority of the air enters our domain from over the ocean, with only approximately 10% originating from the region east of our domain.

To address both reviewers' concerns about backgrounds, we now include a new section (Sect. 3.7) describing how we determine and propagate the uncertainties in our estimates. As the reviewer surmises, the background CH4 mixing ratio contributes the greatest uncertainty to our calculations, now shown in Table S4.

With regards to 10-day footprints, the observed free tropospheric mixing ratio that we use for our background obviously remains constant. The modelled footprint sensitivity does not significantly change if the run time is extended further back because the air will have moved outside of our study domain (50-75°N and 130-170°W) and only surface influences from inside our domain are included in our analysis of the modelled column enhancement. As such, modelled column enhancements from the 5-day cumulative footprints closely resemble those calculated from the 10-day cumulative footprints.

Finally, we apologize for the figures not displaying properly in the Discussions paper. We will ensure that we double check that there are no more compiling issues in the future.

Specific comments

Page 1 line 1. "emissions from northern regions is still poorly constrained". Change 'is' to 'are.'

This change has been implemented at the reviewer's suggestion.

Page 1, line 10. Change 'flux' to 'fluxes'.

This change has been implemented at the reviewer's suggestion.

Page 4, line 9 to 11. I find this sentence a bit confusing. I presume you interpolate the instrument's calibration curves between calibration times. I am not sure what the additional interpolation is?

The sentence was meant to convey that we interpolate using the calibration curve determined by the low and high spans. The reference to the second interpolation has been removed and the text now reads "`These in-flight calibrations were linearly interpolated between calibration times to generate time-varying calibration curves`".

Page 5, line 15. 'among other methodological improvements' is a bit vague. Either give more detail about these changes or remove.

This line has been removed as the relevant considerations are as already stated.

Page 5, line 31. You refer to a 5 day footprints here, but on page 5, line 17 you say that footprints were calculated over 10 days. Have I misunderstood something?

WRF-STILT footprints were calculated for a total of ten days by Henderson et al. (2015). However, our study only uses the first five days because the air had mostly exited our study domain by then. To clarify this, we now state this clearly in the last paragraph of Sect. 3.1.

Page 11, line 29. For 2012, are there any differences between this study and Chang et al., (2014)? They appear to use identical data and methods.

This study covers a slightly larger east-west domain (130–170°W vs 140–170°W) and only reports fluxes from non-mountainous land. The primary reason for including the 2012 analysis was so that the results could be compared over all three years using a consistent analysis method. It also served as a means of checking our current method.

Page 11, line 29. Can you also compare your fluxes with those from Karion et al (2016)?

In the submitted manuscript, we compared our fluxes with those from Karion et al. (2016) in the last paragraph of Sect. 4.1. The section that the reviewer listed is in the budget calculations where we do not compare with the study by Karion et al. because they do not report an estimated budget. It would be inappropriate to extrapolate their results to our entire domain since the CRV tower is not sensitive to the same surface types.

Page 12, line 27. Please be more specific about how the background for the CRV tower was calculated.

As discussed by Karion et al. (2016), backgrounds for the CRV tower were calculated by following particle trajectories in WRF-STILT backwards until they crossed a Pacific basin boundary "curtain", which is determined based on an interpolation of observations. These details have now been included in the text.

Page 12, line 20. How large was the difference between CARVE and BRW? Perhaps you could show a scatter plot showing the different methods used to derive the background. This comparison is a bit vague at the moment.

As recommended by the reviewer, a figure illustrating background CH4 determined from this study, the CRV tower and BRW observations is now included in the Supplement (Fig. S3). The comparison is quite favorable with the exception of those months already mentioned in the text.

Page 14, line 20. It is worth noting that wetland maps can show significant difference (e.g. Melton et al., Biogeosciences, 2013). I wonder if this would impact on your results in section 4?

Since different wetland maps have different spatial distributions, the results are dependent on the map that we use. In earlier iterations of this analysis we did try different maps, including those compared by Melton et al. (2013). However, we decided to only show results from the map by Bergamaschi et al. (2007) because the focus of the study was to quantify the methane budget and not to optimize wetland maps. The study by Miller et al. (2016) compares CH4 emissions estimated using different wetland maps extensively and we now direct the reader to this study in this section of the main text.

Page 15, lines 1-3. What do you mean by "diluted the contributions of other land types?" Does this increase the flux from other land types?

'Dilution' was intended to convey the fact that by averaging over a region where no flux is expected, that the estimated monthly flux will be underestimated in areas where the flux is actually occurring. This does not affect budget calculations and allows comparisons with tower and chamber studies more easily. This sentence has been changed to "as their inclusion would have led to an underestimation of the net CH_4 flux attributed to non-mountain land surfaces".

Page 15, line 7. Add correlation coefficients to main text.

These have now been added to the main text.

Page 16, line 20. Why was May 2014 suspected of being an overestimate?

As was discussed in the submitted manuscript on page 12, line 29–31, the flux estimates in May 2014 are possibly overestimated based on comparisons of observed free tropospheric mole fractions of CH4 by the CARVE aircraft being significantly lower than those observed at either the CRV Tower or at Barrow. At the reviewers request, this is now illustrated in Fig. S3 in the Supplement.

Page 17, line 33. Suggest you reword e.g. "... regional emissions can be determined by up-scaling local scale studies"

This has been changed to the reviewer's suggestion.

Page 18, line 1-5. Is 3 years really long enough to comment on a lack of inter-annual variability?

We agree with the reviewer here and have removed this entire paragraph.

Figure 3. The legend doesn't explain what the green and grey shading are, the units in the legend are missing the exponent, missing axis labels. Check formatting.

Figures 4 to 7. These appear to have similar problems to Fig 3. Please check.

We apologize that the figures in the discussion paper were not properly formatted. This will be fixed for the final publication.

References

Bergamaschi, P., Frankenberg, C., Meirink, J., Krol, M., Dentener, F., Wagner, T., Platt, U., Kaplan, J., Körner, S., Heimann, M., et al.: Satellite chartography of atmospheric methane from SCIAMACHY on board ENVISAT: 2. Evaluation based on inverse model simulations, Journal of Geophysical Research: Atmospheres, 112, 2007.

Chang, R. Y.-W., Miller, C. E., Dinardo, S. J., Karion, A., Sweeney, C., Daube, B. C., Henderson, J. M., Mountain, M. E., Eluszkiewicz, J., Miller, J. B., et al.: Methane emissions from Alaska in 2012 from CARVE airborne observations, Proceedings of the National Academy of Sciences, 111, 16 694–16 699, 2014.

Chou, W. W., Wofsy, S. C., Harriss, R. C., Lin, J. C., Gerbig, C., and Sachse, G. W.: Net fluxes of CO 2 in Amazonia derived from aircraft observations, J. Geophys. Rese., 107, 4614, doi:10.1029/2001JD001295, http://www.agu.org/pubs/crossref/2002/2001JD001295.shtml, 2002.

Commane, R., Lindaas, J., Benmergui, J., Luus, K. A., Chang, R. Y.-W., Daube, B. C., Euskirchen, E. S., Henderson, J. M., Karion, A., Miller, J. B., Miller, S. M., Parazoo, N. C., Randerson, J. T., Sweeney, C., Tans, P., Thoning, K., Veraverbeke, S., Miller, C. E., and Wofsy, S. C.: Carbon dioxide sources from Alaska driven by increasing early winter respiration from Arctic tundra, Proc. Nat. Acad. Sci., 114, 5361–5366, doi:10.1073/pnas.1618567114, http://www.pnas.org/content/114/21/5361.abstract, 2017.

Gatti, L. V., Gloor, M., Miller, J. B., Doughty, C. E., Malhi, Y., Domingues, L. G., Basso, L. S., Martinewski, a., Correia, C. S. C., Borges, V. F., Freitas, S., Braz, R., Anderson, L. O., Rocha, H., Grace, J., Phillips, O. L., and Lloyd, J.: Drought sensitivity of Amazonian carbon balance revealed by atmospheric measurements., Nature, 506, 76–80, doi:10.1038/nature12957, 2014.

Henderson, J. M., Eluszkiewicz, J., Mountain, M. E., Nehrkorn, T., Chang, R. Y.-W., Karion, A., Miller, J. B., Sweeney, C., Steiner, N., Wofsy, S. C., and Miller, C. E.: Atmospheric transport simulations in support of the Carbon in Arctic Reservoirs Vulnerability Experiment (CARVE), Atmospheric Chemistry and Physics, 15, 4093–4116, doi:doi:10.5194/acp-15-4093-2015, 2015.

Karion, A., Sweeney, C., Wolter, S., Newberger, T., Chen, H., Andrews, a., Kofler, J., Neff, D., and Tans, P.: Long-term greenhouse gas measurements from aircraft, Atmospheric Measurement Techniques, 6, 511–526, doi:10.5194/amt-6-511-2013, 2013.

Karion, A., Sweeney, C., Miller, J. B., Andrews, A. E., Commane, R., Dinardo, S., Henderson, J. M., Lindaas, J., Lin, J. C., Luus, K. A., Newberger, T., Tans, P., Wofsy, S. C., Wolter, S., and Miller, C. E.: Investigating Alaskan methane and carbon dioxide fluxes using measurements from the CARVE tower, Atmos. Chem. Phys., 16, 5383–5398, doi:10.5194/acp-16-5383-2016, http://www.atmos-chem-phys.net/16/5383/2016/, 2016.

Melton, J., Wania, R., Hodson, E., Poulter, B., Ringeval, B., Spahni, R., Bohn, T., Avis, C., Beerling, D., Chen, G., et al.: Present state of global wetland extent and wetland methane modelling: conclusions from a model intercomparison project (WETCHIMP), Biogeosciences, 10, 753–788, 2013.

Miller, S., Miller, C., Commane, R., Chang, R.-W., Dinardo, S., Henderson, J., Karion, A., Lindaas, J., Melton, J., Miller, J., Sweeney, C., Wofsy, S., and Michalak, A.: A multi-year estimate of methane fluxes in Alaska from CARVE atmospheric observations, Global Biogeochemical Cycles, 30, 1441–1453, 2016.

Estimating ~~regional scale~~ regional-scale methane flux and budgets using CARVE aircraft measurements over Alaska

Sean Hartery[1], Róisín Commane[2], Jakob Lindaas[2], Colm Sweeney[3,4], John Henderson[5], Marikate Mountain[5], Nicholas Steiner[6], Kyle McDonald[6], Steven J. Dinardo[7], Charles E. Miller[7], Steven C. Wofsy[2], and Rachel Y.-W. Chang[1,2]

[1]Department of Physics and Atmospheric Science, Dalhousie University, Halifax NS
[2]School of Engineering and Applied Sciences, Harvard University, Cambridge MA
[3]Global Monitoring Division, National Oceanic and Atmospheric Administration Earth System Research Laboratory, Boulder CO
[4]Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder CO
[5]Atmospheric and Environmental Research, Inc., Lexington, MA
[6]Department of Earth and Atmospheric Science, City College University of New York, New York NY
[7]Jet Propulsion Laboratory, California Institute of Technology, Pasadena CA

Correspondence to: R.Y.-W. Chang (rachel.chang@dal.ca)

Abstract. Methane (CH_4) is the second most important greenhouse gas but ~~their~~ its emissions from northern regions ~~is~~ are still poorly constrained. In this study, we analyze a subset of in situ CH_4 aircraft observations made over Alaska during the growing seasons of 2012–2014 as part of the Carbon in Arctic Reservoir Vulnerability Experiment (CARVE). ~~Surface~~ Net surface CH_4 fluxes are estimated using ~~an atmospheric particle transport~~ a Lagrangian particle dispersion model which quantitatively links surface emissions from Alaska and the western Yukon with observations of enhanced CH_4 in the ~~boundary~~ mixed layer. We estimate that between May and September, ~~2.1~~net CH_4 emissions from the region of interest were 2.2±0.5 Tg, ~~1.7~~1.9±0.4 Tg and ~~2.0~~ 2.3±~~0.3~~ 0.6 Tg of CH_4 ~~were emitted from the region of interest~~ for 2012–2014, respectively. ~~The predominant sources of the budget were two broadly classed~~ If emissions are only attributed to two biogenic eco-regions within our domain, ~~with from the tundra region~~ then tundra regions were the predominant source, accounting for over half of the overall budget ~~,~~despite only representing 18% of the total surface area. Boreal regions, which cover a large part of the study region, accounted for the remainder of the emissions. Simple multiple linear regression analysis revealed that overall, CH_4 ~~flux~~ fluxes were largely driven by soil temperature and elevation. In regions specifically dominated by wetlands, soil temperature and moisture at 10 cm depth were important explanatory variables while in regions that were not wetlands, soil temperature and moisture at 40 cm depth were more important, ~~reflecting the depth at which methanogenesis occurs~~suggesting deeper methanogenesis in drier soils. Although similar ~~variables~~ environmental drivers have been found in the past to control CH_4 emissions at local scales, this study shows that they can be used to generate a statistical model to estimate the ~~regional scale~~ regional-scale net CH_4 budget.

2 Introduction

Recent trends in observed global atmospheric methane (CH_4) mole fractions have shown increases since a short-lived stabilization period in the early 2000's and have increased by ~150% from pre-industrial values (Dlugokencky et al., 2011; Kirschke et al., 2013; Ciais et al., 2014). As the second most potent anthropogenically-emitted greenhouse gas after carbon dioxide (CO_2) in terms of total radiative forcing, CH_4 can account for 20% of recent trends in global surface air temperatures, which have risen approximately 0.6 K over the past century (Kirschke et al., 2013; Ciais et al., 2014). Since emissions of CH_4 from wetlands represent ~30% of the global CH_4 produced annually, it is of critical scientific interest to determine whether these sources will strengthen in a warming climate (Whalen, 2005; Kirschke et al., 2013; Ciais et al., 2014). In particular, it has been speculated that increased air temperatures in wetlands north of ~~40N~~40°N, combined with ecological, hydrological, and biogeochemical changes could couple into a temperature-emissions feedback, potentially leading to a depletion of organic carbon previously sequestered in below-ground permafrost and increased atmospheric CH_4 (Schuur et al., 2015; Tarnocai et al., 2009).

~~Temperature increases have been more drastic~~ Temperatures have increased dramatically in Arctic regions, which have seen almost 3 K increases in air temperatures since the beginning of the 20th ~~century~~ Century (Overland et al., 2015), with 1.8 K occurring over the past three decades (Collins et al., 2013). Rising ~~air surface~~ surface air temperatures have resulted in a response in soil temperatures, with winter-time observations of permafrost core temperatures in Alaska showing increases of approximately 3–4 K on the Arctic Coastal Plain of Alaska at borehole depths of 5–20 m and 1–2 K in the Brooks Range at depths of 20 m (Osterkamp, 2005). If current Arctic climate trends continue, up to 10–30% of permafrost in Arctic lowlands could significantly degrade, leading to measurable ecological shifts and adding new labile organic carbon to the carbon cycle (Jorgenson et al., 2006).

From a biogeochemical perspective, CH_4 is produced in soils below the water table, which provide the anaerobic conditions necessary for fermentation of soil organic carbon stocks (Whalen, 2005). The fermented organic carbon products are then consumed by methanogenic archaea within the soil column, producing CH_4 gas (Whalen, 2005). As CH_4 production is a biological process relying on microbial activity, it is commonly observed that high CH_4 emissions are coincident with warm, wetland soils (Sturtevant et al., 2012; Olefeldt et al., 2013; Kirschke et al., 2013; Christensen, 1993).

In Alaska and the neighbouring Yukon, seasonal wetlands make up nearly 12% of the total surface area (Bergamaschi et al., 2007), with 92% of soils within continuous permafrost zone (Hugelius et al., 2013). Since ~~Alaska~~ this region is frozen most of the year, its carbon stocks have long been preserved within permafrost, limiting carbon mobilization through respiration. In spite of recent warming and mobilization of sequestered carbon, a recent study of atmospheric mole fractions of CH_4 in the North Slope of Alaska found no significant increase in annual CH_4 emissions over the past 29 years (Sweeney et al., 2016). Field observations have also reported that microbial communities linked to CH_4 oxidation thrive in soils with low moisture content (Xue et al., 2016), highlighting that a warming climate may not have a one-to-one effect on biogenic CH_4 flux at a regional scale.

Numerous past studies have been conducted in permafrost regions such as Alaska at the scale of chambers (as summarized by Olefeldt et al., 2013) and eddy-covariance towers (e.g. Fan et al., 1992; Sturtevant et al., 2012; Zona et al., 2016), providing insight on factors controlling CH_4 emissions at the scale of $\sim$1 m to $\sim$1 km. These studies have revealed that CH_4 emissions are spatially inhomogeneous at those scales and are highly dependent on local conditions such as soil moisture, temperature, elevation and soil carbon. While these process-based studies are extremely important, extrapolating the results to larger scales can be challenging, although not impossible (e.g. O'Shea et al., 2014) . At the other extreme, top-down inversion studies estimate global CH_4 emissions using measurements from surface sites around the world and/or satellite observations coupled to sophisticated transport models (Bruhwiler et al., 2014; Bergamaschi et al., 2013; Chen and Prinn, 2006). These results provide insight on regional emissions; however, it is more difficult to understand local drivers that affect emission rates. More recently, tall towers, either alone or in a network, and aircraft observations ~~,~~ have been used to study regional emissions of CH_4 from permafrost areas (Karion et al., 2016; Sasakawa et al., 2010; Chang et al., 2014). Tall towers are advantageous since their operation is less dependent on weather conditions than static chamber measurements, they provide continuous measurements and their footprint spans a much larger area than a typical eddy-covariance tower. Despite this, in a region like Alaska, where high mountains significantly affect transport patterns, a single tower may not be sensitive to the entire region throughout the year (Karion et al., 2016). In contrast, aircraft observations, by virtue of their mobile platform, can sample larger regions and can periodically measure in the free troposphere to establish background levels. However, their coverage is dependent on weather conditions.

In this study, we estimate net surface CH_4 fluxes using in situ observations of CH_4 from an aircraft that flew in Alaska as part of the Carbon in Arctic Reservoirs Vulnerability Experiment (CARVE). This work uses similar methods as Chang et al. (2014), but extends the analysis to include the growing seasons of 2013 and 2014 and explores how their interannual and intra-annual variability can be explained by hydrological and environmental controls at a regional scale. A recent study by Miller et al. (2016) explores similar questions using a more complex geostatistical inversion model constrained by a much larger data set. Using ~~simple~~ multiple linear regression models, we investigate the relationship between land surface properties and observed atmospheric CH_4~~and find that the results are similar. This~~ . Our method finds similar results to the complex geostatistical inversion model employed by Miller et al. (2016) and could provide a ~~simpler~~ simple diagnostic tool for regional ~~carbon~~ methane cycle analysis.

3 Methods

3.1 Description of aircraft flights

The data presented in this study were collected on board a National Aeronautics and Space Agency (NASA) C23-B aircraft during the Alaskan growing seasons of 2012–2014. In 2012, flights occurred in the last two weeks of each month from May to September for a total of 31 flight days and 212 flight hours. In 2013 and 2014, flights occurred in the first two weeks of each month from April to October and May to November, resulting in 290 and 300 flight hours over 42 and 48 flight days, respectively. All flights analyzed in this study originated from Fairbanks, AK and stayed within Alaska. Not all regions and

altitudes were sampled every month since the flight routes were limited by icing concerns and weather conditions. Data were typically acquired at 150 m above ground level (AGL) to maximize sensitivity to local surface-atmosphere fluxes; however, periodic profiling of the atmosphere to 5–6 km AGL occurred throughout the flights to characterize planetary boundary layer height and mole fractions in the free troposphere.

3.2 Measurements

Mole fractions of CH_4, CO_2, carbon monoxide (CO) and water vapour (H_2O) were measured in situ every ∼2.5 s using two independent cavity ring-down spectrometers. Each system had an individual rear-facing inlet on the port side of the aircraft. Sample flow from the inlets passed through a length of Synflex tubing with an approximate transit time of 30±~~2s~~2 s. The first system (Picarro; G1301-m in 2012, G2401-m in 2013 and 2014) directly sampled the air without any pre-treatment and alternately sampled from one of two on-board calibration cylinders every 30 min. These on-board calibration cylinders were changed throughout the campaigns as the pressure approached 3.4 MPa (500 psi). They were calibrated by NOAA before and after each deployment. Calibration for the water vapour correction was conducted on the flight instrument before and after each year's campaign according to Chen et al. (2010). Further details of this system can be found in Karion et al. (2013).

The sample flow in the second system first passed through a 0.2 μm Teflon filter before passing through a Nafion dryer followed by a dry ice trap which effectively lowered the dewpoint temperature of the sample flow to ∼~~195K~~ 195 K before entering the spectrometer (Picarro; G2401-m). This reduced the water vapour of the sample flow to $< 0.001\%$ and allowed dry mole fractions of CH_4, CO_2 and CO to be directly measured. Two onboard 8 L calibration cylinders (independent of the first system's) were sampled at the beginning and end of every flight as well as every 30 min during flight. These in-flight calibrations were linearly interpolated between calibration times ~~, which were then additionally interpolated to calculate the measured mole fraction in between calibrations~~to generate time-varying calibration curves. In-flight calibrations that were greater than 2.5 standard deviations away from the mean of the calibrations for a given flight were excluded.

Before each year's campaign, the onboard calibration cylinders for the second system were flushed twice and then filled from 30 L fill tanks (Scott Marrin). Due to the longer sampling seasons in 2013 and 2014, the onboard calibration tanks were topped up once in each of those years with the original fill tanks before the pressure dropped below 3.4 MPa. Both the original fill tanks from Scott Marrin and the onboard calibration cylinders were calibrated in the laboratory with the spectrometer used in flight before and after each year's campaign using tanks with known mole fractions obtained from NOAA, tying our measurements to the WMO scales (Dlugokencky et al., 2005; Novelli et al., 1994; Zhao and Tans, 2006). The differences in mole fractions before and after each year's flight campaigns in both the ~~on-board~~ onboard calibration cylinder and the fill tanks were less than 0.15 ppm for CO_2, 0.8 ppb for CH_4 and 7 ppb for CO, all of which were within the precision of the system. The mole fraction in the on-board calibration cylinders was therefore treated as constant throughout each year's study and assumed to be unaffected by the addition of gas. It was also assumed that any changes within any of the tanks were negligible throughout each year's campaign.

The one deviation was that the onboard calibration cylinders were not calibrated before the 2012 deployment due to time constraints. However, these cylinders were calibrated after the 2012 campaign and compared to the fill tanks calibrated before

and after that year's campaign. Except for CO_2 in one of the onboard calibration cylinders, which was 0.29 ppm lower than the fill tank, the comparison with the fill tanks for the other gases all fell into the ranges given above. In the case of this exception, the post-mission calibration of the onboard calibration cylinder was used since the tank was not topped up during that year's mission.

Comparison of the two systems showed a mean difference of 0.8, 0.3 and 0.4 ppb for CH_4 for 2012–2014, respectively with no dependence on water vapour levels ($r^2 = 0.008$). The mole fractions presented in this study merges those measured by both systems, to fill in times when one of the systems was calibrating or functioning imperfectly. The system used to gap-fill is offset by the mean difference between the two systems for a given flight. Since our analysis is based on the difference between mole fractions in the free troposphere and the mixed layer (see below), this treatment does not bias our analysis. The greenhouse gas measurements were merged to a common 5 s time scale along with the location data, measured using a global positioning unit (GPS) (Crossbow; NAV420); outside air pressure (Parascientific; 745-15A); outside air temperature (Harco; 100366-18), dewpoint temperature (Edgetech; Vigilant) and ozone (2B Technologies; 205).

4 Analysis methods

4.1 Footprint sensitivity

To ~~relate the variation of~~ identify the contribution by upwind surface processes to observed mole fractions of ~~to surface processes, an adjoint atmospheric transport model was used to simulate the sensitivity of the air parcels measured by the aircraft to the surface influence~~CH4, source-receptor relationships ("footprints"), that represent the adjoint of the transport model, were computed using a Lagrangian Particle Dispersion Model (LPDM) driven by three-dimensional winds from a regional, high-resolution numerical weather prediction model. Specifically, the Weather Research and Forecasting (WRF; Skamarock et al. (2008)) regional numerical weather prediction model was coupled offline to the Stochastic Time-Inverted Lagrangian Transport model (STILT; Lin et al. (2003)). The coupling of the STILT model with WRF meteorological fields, hereafter WRF-STILT, is described by Nehrkorn et al. (2010) The polar variant of WRF (Hines et al., 2011; Bromwich et al., 2009; Hines and Bromwich, 2008) coupled with STILT was configured for the Arctic CARVE domain by Henderson et al. (2015), who provide a detailed description and meteorological validation of the high-resolution simulations performed on a 3.3-km grid suitable for simulating particle transport on regional scales in mountainous terrain. Their initial STILT footprints from WRF v3.4.1 were used in the analysis of Chang et al. (2014). The current investigation uses WRF v3.5.1 and incorporates recently-available PIOMAS cryosphere fields (Zhang and Rothrock, 2003; Hines et al., 2015) and footprints on an expanded circumpolar domain north of ~~30N, among other methodological improvements~~30°N.

To avoid unnecessary computational costs, measurement locations that occurred below 1 km AGL were placed into discrete bins on a grid of 5 km in the horizontal and 50 m in the vertical, while locations above 1 km AGL were similarly binned 5 km in the horizontal and 100 m in the vertical (Chang et al., 2014; Commane et al., 2017) . These spatial bins were then used as receptor points for each STILT simulation. Measurement times were also truncated to the hour to match the time resolution of

the meteorological fields. The use of WRF at high resolution to drive the STILT model results in higher fidelity meteorological fields and subsequent transport calculations than afforded by most global reanalysis (Nehrkorn et al., 2010; Henderson et al., 2015) .

To derive source-receptor relationships, each STILT simulation launched 500 particles into the atmosphere at the location of ~~a measurement and traced its~~ each receptor point and traced their dispersion through the atmosphere in reverse time over 10 days via 3D advective winds and stochastic (i.e. random/probabilistic) processes determined from the underlying meteorology (Lin et al., 2003; Henderson et al., 2015). At every ~~one-hour time step~~hour, particles that ~~are~~ were in the lower half of the boundary layer ~~are~~ were assumed to be influenced by the surface (Gerbig et al., 2003, see Supplement for further discussion about this assumption) and are gridded to generate a footprint sensitivity plot for that hour. The footprint used in this analysis was calculated on a $0.5° \times 0.5°$ lat/lon grid~~. The unit of the footprints is~~ , with a unit of abundance flux^{-1} (e.g. ppb (nmol m^{-2}s^{-1})$^{-1}$, where ppb is parts per billion) (Lin et al., 2003) and allows us to relate ~~surface emissions~~ upwind surface influence to an observation at a given time and location.

~~To avoid unnecessary computational costs, the measurement locations were first placed into discrete bins on a grid of 5 km in the horizontal, 50 m in the vertical below 1 km and 100 m in the vertical above 1 km (Chang et al., 2014) . Measurement times were also rounded to the hour to match the time resolution of the meteorological fields. The use of WRF at high resolution to drive the STILT model results in higher fidelity meteorological fields and subsequent transport calculations than afforded by most global reanalysis (Nehrkorn et al., 2010; Henderson et al., 2015)~~ These footprints are therefore species-independent. The original STILT runs simulated transport 10 days backward in time (Henderson et al., 2015) . However, as discussed in the next section, our analysis only uses the first five days. Past studies that have used WRF-STILT footprint sensitivities for this region have been successful in attributing fluxes of CO_2 (Commane et al., 2017; Karion et al., 2016) , a much harder problem due to its bi-directional flux and strong diurnal cycle. These studies were therefore much more reliant on WRF-STILT generating accurate footprints during both day and night compared to our study. The success of these past CO_2 studies gives us confidence in the footprint sensitivities used in this analysis.

4.2 Domain

This analysis focuses on the influence of surface emissions from Alaska and Yukon on observed CH_4 mole fractions. To this end, we restricted our region of interest to ~~50–75N and 130–170W~~ 50–75°N and 130–170°W (Fig. 1) and only consider footprint sensitivity derived from WRF-STILT during the first five days preceding each receptor point. The time-scale of five days was chosen ~~under the assumption that~~ because sources outside the domain ~~would have transport timescales~~ are expected to have transport time-scales roughly equal to that of free-tropospheric mixing and ~~would~~ therefore only contribute to background values of CH_4. Figure 2 shows the cumulative ~~5 day~~ five-day footprint of our analyzed times for each year averaged over the number of profiles in each year ~~,~~ and shows that our observations are most influenced by the boreal interior of Alaska as well as tundra regions. It should be noted that although our mean footprints were more sensitive to the North Slope of Alaska in 2012, our absolute sensitivity was comparable in all three years due to the increased number of flights in 2013 and 2014.

Our choice of the spatial domain was determined by the distribution ~~in~~ of five-day ~~footprint~~footprints. As will be defined in the subsequent section concerning the mixed layer, footprint sensitivity simulations that had less than 1.0 ppb (nmol m^{-2}

$s^{-1})^{-1}$ total land sensitivity were identified as being in the free troposphere, following Henderson et al. (2015). For an idealized homogeneous distribution of sensitivity, where an air parcel is equally sensitive to all land areas within the domain, this threshold would translate to $\sim 1 \times 10^{-4}$ ppb (nmol m^{-2} $s^{-1})^{-1}$ in any given grid cell. It was assumed that areas of the globe with sensitivities smaller than this threshold did not significantly contribute to CH_4 enhancements, since, even at extreme fluxes of 1000 mg m^{-2} d^{-1}, they would contribute <0.1 ppb to the observed enhancement. In the cumulative ~~5 day~~ 5-day footprints shown in Fig. 2, ~~71.2~~71% of the footprint that exceeded the grid-cell threshold was from land in the ~~domain of choice, while 28.5~~chosen domain, while 29% could be sourced to oceans. It is worth noting that the previous study of Chang et al. (2014) restricted ~~their~~ its domain to 135°W. While the May–September footprints showed that <2% of the footprint originated in regions east of 135°W, it was found that at least 5% of the total land sensitivity in April, October, and November came from the Canadian Yukon. Of all the land surface influence on our observations, only 0.3% originated from land outside of our study region, suggesting that our sampling strategy and choice of domain allowed our observations to be most sensitive to surface emissions from the study region.

4.3 Mixed layer

We use the term 'mixed layer' to refer to the combination of both the residual layer and the boundary layer throughout the remainder of this paper. As our ~~footprint sensitivity predicts that we are sensitive to surface fluxes 5 days previous to measurement, quantifying the total amount of emitted requires calculating the amount of~~ footprints predict the sensitivity to surface influences five days prior to a measurement, both local and regional sources need to be considered. As a general approximation, CH_4 ~~in both the boundary layer and the residual layers from previous days. By determining the altitude at which the residual layer is capped, it is possible to determine~~ enhancements in the boundary and residual layers reflect local and regional surface emissions, while the portion of a vertical profile ~~that is affected by surface emissions versus the portion that represents background levelsin the free troposphere. The column enhancement within the mixed layer will be used in Sect. 4.4 to estimate emissions~~above the mixed layer can be used to represent background levels. In practice, the footprint sensitivity maps attribute the spatial distribution of sources that influence a particular receptor, whether it is in the boundary layer or residual layer. It should be noted that particles in the residual layer do not contribute to the footprint at that specific time but carry surface influence from earlier time spent in the boundary layer and therefore may have elevated surface influence compared to background values.

We estimate the bottom of the free troposphere (h) by calculating the refractivity (N) using parameters measured at different altitudes on the aircraft (Chan and Wood, 2013; Bean and Dutton, 1966):

$$N = 77.6\frac{P}{T} + 3.77 \cdot 10^5 \frac{P_w}{T^2}, \quad (1)$$

where P and P_w are the atmospheric pressure of air and water in hPa, respectively, and T is the atmospheric temperature in K. The height at the minimum of the gradient in N is the top of the mixed layer.

~~In analyzing individual profiles~~Tables S1–S3 in the Supplement lists every time the aircraft flew between 200 m and at least 2.7 km as an individual profile. In analyzing these profiles, it was often found that this method over-estimated the ~~boundary~~

mixed layer height as compared to profiles of CH_4, O_3 and the total land surface sensitivity calculated by WRF-STILT; as a secondary calculation, the height at which WRF-STILT sensitivities dropped below 1.0 ppb (nmol m^{-2} $s^{-1})^{-1}$ was used to approximate h (Henderson et al., 2015). Upon individual inspection of profiles, it was found that in 25% of cases (i.e. 68 profiles) neither method captured an h which correctly separated the free troposphere and mixed layer. After removing these profiles, a further 25% of profiles within the remaining set (i.e. 50) had h differing by >750 m between the two methods (paired t-test; p<0.005). In these cases, either the boundary layer dynamics within WRF-STILT were not representative of local meteorology or there was too much variance about the minimum of the gradient in N. Therefore, these profiles were not included in the analysis. In the final set of profiles used for analysis, a paired t-test found that the two methods did not statistically differ from each other (p>0.1) and both agreed to within 500 m ~90% (r^2 = 0.6) of the time. We attempted to use other variables such as virtual potential temperature to define the mixed layer but the methods described above gave the most consistent results.

Across all years h was found to occur between 1.1 and 1.8 km with a median of 1.5 km above Alaska, consistent with estimates derived from satellite retrievals (Chan and Wood, 2013). On average, h in May–September was 30% higher than in April, October or November. These findings are consistent with lower solar zenith angles and shorter solar days in colder months, both of which reduce the strength of convective forces which form the boundary layer (Stull, 1988).

4.4 CH_4 flux estimates

In this analysis, we use column enhancements of CH_4 (ΔCH_4) to estimate ~~surface fluxes (Chang et al., 2014; Gatti et al., 2014; Chou et al., 2~~ surface fluxes (Commane et al., 2017; Chang et al., 2014; Gatti et al., 2014; Chou et al., 2002) . This method assumes that CH_4 enhancements below h that are above background levels are a result of interactions with the local surface that are not well-mixed with the free troposphere. ~~Surface~~ Net surface fluxes are determined by calculating the corresponding surface influence ~~throughout~~ at each height in the column using the footprint sensitivities determined from WRF-STILT. This method of calculating ΔCH_4 is less reliant on the accurate simulation of the vertical structure of the atmosphere as well as turbulent transport in the lower atmosphere. Instead, we rely on the integrated model simulation to match the integrated observations. The majority of the measurements from the aircraft flights were in the surface layer. As such, this method also reduces bias in regional emission estimates by focussing on measurements throughout the atmosphere.

To calculate ΔCH_4, measurements in a vertical profile were first block-averaged by altitude into 250 m bins. A sample of such a profile observed during the CARVE campaign is shown in Fig. 3 and shows median, minimum and maximum [CH_4] observed in each 250 m bin up to 3 km along with the estimated ΔCH_4 from the WRF-STILT footprint influence according to Eq. 3. In this case, the calculation of the refractive index (N) properly captured the transition from the mixed layer to the free troposphere and WRF-STILT adequately modelled the mixed layer, leading to a reasonable estimate of CH_4 flux. Using h as the upper bound on the section of the profile affected by local sources and sinks:

$$\Delta CH_4 = \int_0^h ([CH_4](z) - [CH_4]_0) \frac{P_d(z)}{R^* T(z)} dz, \quad (2)$$

where ΔCH_4 is the column enhancement in the mixed layer of CH_4 in ~~(ppb)~~nmol CH_4 m^{-2}, $[CH_4]_0$ is the ~~estimated background of~~ background CH_4 estimated by averaging the lowest 1 km in the free troposphere~~averaged in the first kilometer above *h*~~, P_d is the calculated atmospheric pressure of dry air using measurements of P and P_w, R^* is the universal gas constant, and z is the height ~~above ground level (AGL)~~ AGL as calculated from the height above sea level measured by the GPS and an underlying digital elevation map within the input meteorology of WRF. As an entrainment zone between the free troposphere and mixed layer was usually evident in profiles of CH_4, ~~where mixing ratios were between free troposphere and mixed layer values,~~ 500 m was added to ~~mixing layer height in order to estimate~~ h for our estimate of $[CH_4]_0$. Since $[CH_4]_0$ is stable with altitude in the free troposphere, this should not affect the integral in Eq. 2.

A similar calculation was undertaken for the corresponding footprint sensitivity, with the cumulative five-day footprint at each height replacing ($[CH_4](z)$ - $[CH_4]_0$) in Eq. 2, to calculate the sensitivity of the entire mixed layer to the ~~total~~ surface (ΔI with units of ~~ppb~~ nmol CH_4 m^{-2} (nmol m^{-2} s^{-1})$^{-1}$). As a first estimate, our a priori CH_4 flux is assumed to be uniformly emitting ~~over~~ from our entire domain. ~~Footprint sensitivities were restricted to~~ Surface contributions outside of the study domain ~~and masked from~~ were excluded as were seas and mountains, which we assume to be neither sources nor sinks of atmospheric CH_4 and therefore ~~did~~ to not contribute to ΔCH_4. The total mixed layer sensitivity, ΔI, is then scaled to match ΔCH_4 ~~. This scaling factor is the~~ to estimate CH_4 flux~~estimate~~, $\hat{E}$:

$$\hat{E} = \frac{\Delta CH_4}{\Delta I}, \tag{3}$$

with appropriate unit conversion to mg CH_4 m^{-2} d^{-1}. The estimation of uncertainties in CH_4 flux using Eq. 3 will be discussed in Sect. 4.7.

Monthly mean CH_4 flux ($\overline{E}$) were estimated by ~~weighing~~ weighting individual flux estimates by the total footprint sensitivity of each profile (Chang et al., 2014; Karion et al., 2016), so that:

$$\overline{E} = \frac{(\sum \hat{E} \times \Delta I)}{\sum \Delta I} = \frac{\sum \Delta CH_4}{\sum \Delta I}. \tag{4}$$

This ~~allowed~~ biased the average towards profiles with larger spatial coverage ~~to be more represented while~~ and helped normalize profiles with lower ΔI, which sometimes resulted in ~~a high~~ high estimates of $\hat{E}$~~, did not inadvertently skew the mean estimate~~. The residuals of individual estimates against these monthly means are shown in Figure S2 of the supplement. Their normal distribution about the trend in the monthly mean suggests that this method of weighted averaging was successful in mitigating biases in CH_4 flux stemming from differences in footprint sensitivity.

Growing season budgets were calculated by integrating the monthly emissions from May through September and over our study domain (excluding mountains). ~~Uncertainties in this budget are estimated using the standard error calculated from the weighted standard deviations in the monthly means.~~ The surface area of our domain was estimated using a digital elevation map from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) to estimate the enhancement of surface area in sloped terrain (Jenness, 2004). A land/ocean mask at 0.1°x0.1° was then used to approximate the fraction of land cover at 0.5°x0.5° to estimate surface area in coastal grid cells (i.e. $A_{land} = A^* f_{land}$). The total land surface area of our domain is estimated to be 2.1 million km^2 including mountains and 1.04 million km^2 without.

~~To ensure that our~~ Since the goal of this study is to understand biogenic CH_4 ~~flux estimates were not affected by combustion , measurements~~ emissions, observations influenced by combustion were excluded. Measurements of CO were used as an atmospheric tracer for air influenced by biomass burning events or oil development off-gassing, which co-emit CO and CH_4. Any profile which had measurements of CO exceeding 150 ppb within the mixed layer ~~were~~ was not used in the analysis. This threshold was determined by observing that the annual distribution of CO deviates from normal above 150 ppb (Chang et al., 2014). It should be noted that our calculations only estimate net CH_4 emissions from non-mountainous land surfaces. These estimates include all processes such as biogenic, thermogenic and non-combustion-related anthropogenic sources and sinks. We assume that the latter two processes are small compared to biogenic sources over the study domain, although this may not be the case and our values are in fact upper estimates for net biogenic emissions.

After CO screening, each individual vertical profile was assessed on a case-by-case basis to ensure that the WRF-STILT model ~~had~~ adequately captured a mixed layer, i.e. the footprint sensitivity attenuated to zero as altitude increased and that the estimated ~~mixed layer height~~ h from WRF-STILT clearly separated the mixed layer from the free troposphere. ~~In general, if h defined by WRF-STILT was below 500 m, the profile was rejected due to difficulty in calculating the column enhancement since there was only one binned point below h.~~ Of the 273 profiles identified, 9 were rejected due to excessive CO mole fractions~~, 68 were rejected due to a poor separation of the mixed layer and the free troposphere, and a further 50 were rejected when h differed by more than 750 m~~ . In addition, 118 were rejected using the two methods described in Sect. 4.3. The remaining 146 profiles, just under half of all the profiles, were kept for analysis. While this is a severe reduction in data, footprint sensitivity plots comparing the total footprint sensitivity of all the receptor points within all profiles against the total footprint sensitivity of those within the profiles kept for analysis showed no significant bias. As a result, this subset is thought to be both representative of the CARVE sampling campaign and free from errors in integration limits.

4.5 Eco-region dynamics

The previous assumption that the study region emits CH_4 uniformly is correct only to the zeroth-order~~and serves only~~. It serves as the best estimate for the total magnitude of ~~flux~~net flux and provides the most robust regional budget calculations. However, this assumption misses much of the spatiotemporal heterogeneity observed in emissions across and even within different biological and hydrological regimes (as summarized by Olefeldt et al., 2013). At regional scales, such as in this study, it is useful to separate the domain into eco-regions which group regions with similar vegetation, elevation, soil type and ~~hydrologic flow~~soil moisture dynamics. One can then use the different eco-regions as a basis set of independent sources and sinks in a linear inversion. The eco-regions used within this study were taken from the Environmental Protection Agencies (EPA) Level II map of eco-regions and grouped into the three following land types: tundra, which includes 'Alaska Tundra'; boreal, which includes 'Alaska Boreal Interior', and 'Taiga Cordillera'; and mountains, which includes 'Brooks Range Tundra', 'Marine West Coast Forest', and 'Boreal Cordillera' (Commission for Environmental Cooperation, 1997) (see Fig. 1). While 'Marine West Coast Forests' are not mountainous by definition, a study of CH_4 fluxes north of 50°N identified heavily forested areas as being negligible sinks of CH_4 (Olefeldt et al., 2013). Therefore these forests are suitably grouped with mountains under the

assumption that footprint sensitivity from these areas did not affect observed ΔCH_4. By using these predefined maps, we are attributing all net emissions to biogenic sources.

By calculating the mean fraction of the influence of each eco-region on our measurements using the footprint sensitivity maps, it is possible to estimate the net CH_4 flux from each eco-region using multiple linear least-squares regression according to the following:

$$\hat{E} = \sum_i f_i E_i, \tag{5}$$

where f_i are the fractions of influence from different eco-regions, $\hat{E}$ are the uniform CH_4 flux estimates, and E_i are the CH_4 fluxes from each eco-region. These estimates were grouped by month across all three years to maintain a sufficient sample size. To test the assumption that emissions from the mountain land type negligibly affected ΔCH_4, CH_4 flux was estimated with and without the mountain land type and presented in Sect. 5.4.2.

4.6 Regression analysis

To explore how the variability in CH_4 flux estimates is related to ecological, biological and hydrological parameters, the footprint sensitivity functions were also used to calculate weighted means of different variables common in process-based models of CH_4 flux. Maps that were included in this analysis were: surface, 10-cm and 40-cm daily soil temperature (T_x, where x is the depth [K]) and liquid soil moisture content (S_x [-]) from the North American Regional Reanalysis (NARR) project (Mesinger et al., 2006) ; digital elevation (z [m]) from ASTER; days since thaw (DST), derived from passive microwave satellite observations of surface thaw (Steiner et al., 2015); wetlands (%) (Bergamaschi et al., 2007); 30-cm and 100-cm soil organic carbon content (C_x, kg/m^2) from the Northern Circumpolar Soil Carbon Database (NCSCD) (Hugelius et al., 2013); and percent of soils classed as turbels, histels, gelisols or non-soils from NCSCD (Hugelius et al., 2013).

Previous correlation analyses have suggested that CH_4 flux varies non-linearly with variables such as soil temperature (Yvon-Durocher et al., 2014) and elevation (Karion et al., 2016). As such, the following functional forms were used to characterize variability in CH_4 flux:

$$E = e^{f(x)} \tag{6}$$

$$f(x) = \begin{cases} \mathrm{E}_a\left(\frac{1}{k\overline{x}} - \frac{1}{kx}\right), & \text{Boltzmann-Arrhenius-Type} \\ \left(\frac{1}{x+\overline{x}}\right), & \text{Inverse-Type} \\ (x - \overline{x}), & \text{Linear-Type} \end{cases} \tag{7}$$

where E is CH_4 flux, x is a hypothesized predictor variable, eV is the value of an electron volt in Joules (1.602x10^{-19} J), k is the Boltzmann constant (1.38x10^{-23} J/K), and $\overline{x}$ represents the mean sampled value of x. This transformation of units is performed such that when the predictor variable is soil temperature, the fitted parameter optimizes the activation energy (E_a) in units of eV (Yvon-Durocher et al., 2014).

Multi-variable fits were also performed using an equation of the form:

$$\hat{E} = e^{f(Z)\hat{\beta}}, \tag{8}$$

where $\hat{E}$ is an Nx1 vector of modelled CH_4 flux estimates, Z is an Nx(M+1) matrix with M columns of predictor variables transformed by $f(x)$ and one column of ones, and $\hat{\beta}$ is an (M+1)x1 vector of estimated parameters. We searched through all possible combinations of functional forms by allowing the M columns of Z to take either the Boltzmann-Arrhenius, inverse ~~,~~ or linear forms written in Eq. 7, leaving the column of ones to fit a constant. The parameter vector $\hat{\beta}$ is then constrained using the Levenberg-Marquardt algorithm for non-linear least-squares regression. Subsets of the data (n<N) were also fit to explore behaviour in limiting spatial and temporal cases. The best fit for each subset was chosen if it minimized the Aikake Information Criterion (AIC) (Burnham et al., 2011).

~~Since our estimates represent 5-day averages of~~ The footprint sensitivities are cumulative over five days, causing CH_4 ~~flux~~ flux estimates from individual profiles to represent five-day averages (see Sect. 4.1 & 4.4)~~, flux estimates from nearby days can be significantly correlated~~. Regression analysis using the CH_4 fluxes estimated from individual profiles resulted in a lot of noise and the majority of the environmental drivers could not explain the variability in these estimated CH_4 fluxes. To compensate for this~~correlation~~, CH_4 flux estimates were first averaged into ~~5-day~~ five-day bins before fitting to reduce issues of autocorrelation and to generate independent estimates of CH_4 flux. As with the monthly mean estimates, these were weighted by the total footprint influence. This temporal averaging reduced the total number of independent estimates used in the regression from 146 to ~~68.~~ 68, thus reducing the reliability of these results. It should be emphasized that the regression analyses are only reflective of seasonal variations on time-scales of five days.

4.7 Uncertainties

Uncertainties in the estimated CH_4 flux from individual profiles were determined through bootstrapping both observational measurements and the integrated model footprints involved in the its' calculation by sampling each variable with replacement 500 times. Explicit variables included were: 1) matching observations of CH_4, T, P, and P_w at each 250 m altitude bin to estimate uncertainties associated with the observations and spatial variability; 2) the footprint sensitivity at each 250 m altitude bin, approximating the uncertainty of releasing particles within neighbouring grid cells. While this is not a true posterior variability, and may under-estimate full modelling uncertainties, it is still a useful approximation of a type of model uncertainty. 3) The mixed layer height which we varied by $\pm$250 m and propagated through our calculations of $\hat{E}$ by either adding or removing another 250 m bin to the integration in Eq. 2 but not changing $[CH_4]_0$, representing the uncertainty due to h; 4) same as 3) but changing $[CH_4]_0$ and representing the uncertainty due to $[CH_4]_0$; and 5) all of the above propagated through the entire calculation, representing the total methodological uncertainty in our calculations of CH_4. The 95% confidence interval (C.I.) of the last step was used to estimate the uncertainty associated with the estimated CH_4 flux from individual profiles. The mean, minimum and maximum 95% C.I. (2.5–97.5 percentile) for each of these steps is shown in Table S4. They reveal that the largest methodological uncertainty associated with the results is the combined uncertainty in identifying both the mixed layer height and the $[CH_4]$ background.

Uncertainties in the monthly mean CH_4 flux were estimated in four ways: *average uncertainty*, the 95% C.I. for each individual estimate is averaged for each month with no weighting; *weighted average uncertainty*, the 95% C.I. for each individual estimate is averaged for each month and weighted by the column integrated total surface influence (ΔI); *standard deviation*, the normal standard deviation is calculated from the residuals of the weighted monthly mean; *weighted standard deviation*, the standard deviation of all CH_4 flux estimates within a given month is calculated, weighting each residual by ΔI and the inverse of the square of the 95% C.I. We present uncertainties calculated using the weighted standard deviation because it accounts for both monthly variability as well as the uncertainty in the flux estimated from each profile. Results estimated using the average uncertainty were slightly lower, likely because they do not account for the variability in a given month. Uncertainties in the budget are propagated from the average monthly flux uncertainty by quadrature. For reference, uncertainties in the budget and monthly mean estimates calculated using all four methods are shown in Table S4.

5 Results & discussion

In the following sections, we present estimates of ~~regionally averaged monthly~~ regionally-averaged monthly net CH_4 flux and the net total CH_4 emitted from our study domain from May–September. A discussion of the source of uncertainties in determining $[CH_4]_0$ will follow. Finally, the set of CH_4 fluxes estimated across 2012–2014 will be used to motivate discussions of how sampling different eco-regions across a dynamic range of soil conditions (e.g. soil temperature, soil moisture, soil type) affected the estimated CH_4 flux and budget.

5.1 CH_4 flux estimates

Using the methods described in Sect. 4.4, monthly mean net CH_4 fluxes for 2012–2014 were estimated from individual aircraft profiles of the atmosphere (Fig. 4). As will be shown in Sect. 5.4, we found that footprint sensitivity from oceans and mountains were poorly correlated with observations of ΔCH_4. Therefore the estimated fluxes presented in this subsection only ~~represent~~ attribute CH_4 flux from the boreal and tundra eco-regions defined in Sect. 4.5. These estimates ranged from 2–36 mg CH_4 m^{-2} d^{-1} (~~95%C.I.~~2.5%–97.5% percentile) for individual profiles and showed a distinct seasonal cycle that peaked in late July or early August across all years ~~,~~ (Fig. S1). Figure S2 illustrates that the residuals of the estimated net flux from individual profiles from the monthly mean are normally distributed, suggesting that our monthly mean estimates are not strongly biased. Overall, our results consistent with tall and eddy-covariance tower studies (e.g. Karion et al., 2016; Zona et al., 2016; Sweeney et al., 2016) ~~. Additionally, they~~ and provide observational evidence that region-wide CH_4 flux may on average be as high as 5 mg m^{-2} d^{-1} in the colder months of April and November, also consistent with previous studies on the North Slope (Zona et al., 2016) and interior (Karion et al., 2016) of Alaska.

Monthly mean flux estimates were found to be higher than those inferred from the CRV tower near Fairbanks for the same months (3–9 mg CH_4 m^{-2} d^{-1}) (Karion et al., 2016), although the ranges from the two studies overlap. As discussed by Karion et al. (2016), the tower observations likely underestimate CH_4 flux compared to the aircraft observations because of the aircraft's increased sensitivity to the North Slope and southwestern Alaska, regions that are known to be seasonal wetlands

(and therefore a significant CH_4 source (Bergamaschi et al., 2007)), as compared to the interior sites which were more sensitive to upland regions that are thought to emit less CH_4 (Olefeldt et al., 2013).

5.2 CH_4 budget calculations

Integrating over the months sampled every year (May–September), we estimate CH_4 emissions from our study region to be ~~2.1~~ 2.2 $\pm$ 0.5 Tg, ~~1.7~~ 1.9 $\pm$ 0.4 Tg and ~~2.0~~ 2.3 $\pm$ ~~0.3~~ 0.6 Tg for 2012–2014, respectively, with the assumption that all ~~land surfaces (excluding mountains)~~ non-mountainous land surfaces emit at a uniform rate over the entire month. As our observations do not extend throughout the colder months, we do not provide annual budget estimates since other studies have found that significant emissions of CH_4 are observed in the shoulder and cold seasons (Zona et al., 2016; Sweeney et al., 2016; Karion et al., 2016).

Previous estimates of emissions from Alaska during the growing season include the study by Chang et al. (2014), who used a similar method to estimate the May–September 2012 emissions to be 2.1 $\pm$ 0.5 Tg, as well as the geostatistical inversion of the CARVE observations by Miller et al. (2016), who estimated May–October CH_4 emissions of 1.80 $\pm$ 0.45, 1.65 $\pm$ 0.43 and 1.77 $\pm$ 0.45 Tg CH_4 for 2012–2014, respectively. Our mean estimates are within the uncertainties of these other studies, especially when we account for the $\sim$20% greater area in our study domain. It is promising that our relatively simple method for calculating budgets using selected profiles from the CARVE aircraft observations arrives at similar estimates to values derived from the much more complex geostatistical inverse model used by Miller et al. (2016), particularly as their study was constrained by all the aircraft observations as well as hourly averaged observations from the CRV tower.

Our estimates of May–September net CH_4 flux ~~,~~ have a mean of ~~2.0~~ 2.1 $\pm$ ~~0.4 Tg ,~~ 0.5 Tg and show no significant difference over the three years. These results are consistent with the findings reported by Miller et al. (2016), who suggest that regional CH_4 emissions would require decades to respond to changes in surface conditions. Similarly, a recent analysis of long-term measurements of CH_4 flux on the North Slope in Alaska observed little change in boundary layer CH_4 enhancement over the past 29 years, despite increases in air temperature (Sweeney et al., 2016). This lack of trend could potentially be related to methanogen community structure in the Arctic as recent microbiological research has found that communities from Arctic soils that were incubated at lower temperatures were insensitive to substrate manipulations, indicating that Arctic methanogens may not be sensitive to the addition of new labile carbon from thawing permafrost (Blake et al., 2015). While it is true that local-scale permafrost degradation patterns such as thermokarsts can result in local CH_4 fluxes of >100 mg m^{-2} d^{-1} (Johnston et al., 2014), multi-decadal studies such as Sweeney et al. (2016) suggest that at a regional scale, CH_4 fluxes in Alaska have been stable.

5.3 CH_4 background estimates

5.3.1 Comparison of background CH_4

~~Since our~~ Our calculations of ΔCH_4 rely on the free tropospheric $[CH_4]$ to represent background $[CH_4]$ in the mixed layer~~, it is important to demonstrate that these values are comparable. As seen in Fig. 5, a distinct seasonal cycle in $_0$ is evident and is consistent with cycles observed in the NOAA Global Network (Dlugogencky, 2016) . This gives us confidence that our~~

~~calculated~~ . It is possible that the transport history of the free troposphere is different than that of the mixed layer. In this case, $[CH_4]_0$ ~~is not strongly affected by changes in circulation due to long range transport or stratospheric intrusion of clean air.~~

measured in the free troposphere would not be representative of the background at the surface. To assess the accuracy of our estimates of the $[CH_4]_0$ background, we compare our free tropospheric values to in situ observations of the boundary layer CH_4 background observed at the Barrow Observatory ground station (BRW: 71.3°N, 156.6°W) in 2013 and 2014 (Dlugogencky, 2016). Because the station did not measure $[CH_4]$ in 2012 (Sweeney et al., 2016), we instead compared the 2012 CARVE $[CH_4]_0$ estimates to the free tropospheric backgrounds measured monthly by aircraft at the Poker Flats site in interior Alaska (Global Monitoring Division, 2016). We also compared our estimate of the $[CH_4]_0$ background to the backgrounds determined by Karion et al. (2016) for the CRV tower~~(Karion et al., 2016) . These~~ , who traced sampled air masses backward until they reached a boundary curtain at 170°W. Figure S3 shows that these other estimates of CH_4 background levels were generally within the standard deviation of $[CH_4]_0$ estimated from the CARVE profiles (±8 ppb in 2012 and ±6 ppb in 2013 and 2014). The exceptions were August 2013 and May 2014 when the CARVE observations were lower than the boundary layer CH_4 ~~background~~ observed at BRW from the clean air sector, possibly resulting in an overestimation of CH_4 flux. It should be noted that the BRW site is located on the northern shore of Alaska and is separated from the Alaskan interior by the Brooks Range so it is not always influenced by the same air mass that affects the remainder of the study region. As seen in Fig. 5, a distinct seasonal cycle in background $[CH_4]$ is also evident and is consistent with cycles observed in the NOAA Global CH_4 Network (Dlugogencky, 2016) . This gives us confidence that our calculated $[CH_4]_0$ is not strongly affected by changes in circulation due to long range transport or stratospheric intrusion of clean air.

While the backgrounds estimated from the CRV tower for August 2013 are within our variability, the May 2014 background estimate from the CRV tower is between our estimate and the levels observed at BRW. To evaluate the magnitude of this effect on the estimated May–September budgets, the net CH_4 ~~flux~~ fluxes using the BRW ground station observations as our background were calculated for those two months~~. Resulting~~ , resulting in May–September budgets ~~using these new values are~~ of 1.5±0.2 and 1.8±0.4 Tg CH_4 for 2013 and 2014, respectively. Since these values are within the uncertainties of our initial budget estimates and our background estimates correspond to those from the CRV tower, we believe that our estimates of $[CH_4]_0$ are representative of background levels in the mixed layer.

5.3.2 CH_4 background growth rate

Across the entire campaign, the estimated $[CH_4]_0$ had a median and standard deviation of 1880 ± 20 ppb, and a distinct seasonal cycle with higher mole fractions in colder months than in warmer months (Fig. 5). This ~~cycle is chemically driven~~ seasonal cycle is driven by large scale transport and chemical oxidation and is consistent with observations from global $[CH_4]$ observations (Dlugogencky, 2016). From Fig. 5, it is also evident that background CH_4 in the free troposphere rose from 2012–2014.

We estimate the atmospheric growth rate using monthly mean CARVE observed $[CH_4]_0$ fit to the function $y = \sum_{k=0}^{1} \sin(2\pi t + \frac{k\pi}{2}) + \sum_{k=0}^{1} t^k$, where time ($t$) is in units of years. This function is a simplified form of the function used by NOAA to estimate the global growth rate of CH_4 from their observation network (ESRL, 2016). Fitting with the Levenberg-Marquadt algorithm

for non-linear least-squares, we estimate the atmospheric growth rate in the free troposphere over Alaska to be 9 ± 2 ppb/yr (p<0.001) with a coefficient of determination of 0.77. The estimated growth rate is consistent with the 9 ± 1 ppb/yr and 8.6 ± 0.6 ppb/yr observed at Barrow (11 m a.s.l.) and Mauna Loa, HI (3397 m a.s.l.) respectively.

Overall we find that the monthly mean and annual growth rate determined from CARVE was the same as BRW within the variability of our observations, with the exception of a few months. These results indicate that the local mixed layer $[CH_4]_0$ in Alaska can be constrained from free tropospheric measurements and gives us confidence in our estimates of ΔCH_4.

5.4 Eco-region ~~Dynamics~~dynamics

5.4.1 Tundra & ~~Boreal Eco-regions~~boreal eco-regions

Figure 6 shows the monthly emissions estimated for the tundra and boreal eco-regions over all three years estimated using the linear system in Eq. 5. The seasonal average CH_4 flux from tundra regions was 21 ± 3 mg m^{-2} d^{-1} and ranged from 6–34 mg m^{-2} d^{-1}. ~~Although comparable~~ These estimates are comparable to flux observations in 2013–2014 from eddy-covariance towers on the North Slope, where the monthly mean emissions from May to September ranged from 6–21 mg m^{-2} d^{-1} (Zona et al., 2016) , as well as from the Yukon River Delta (25 mg m^{-2} d^{-1}) (Fan et al., 1992) . However, our mean is lower and our range is narrower than those reported in a database of flux observations at the chamber-scale compiled by Olefeldt et al. (2013), who found that average net CH_4 flux from wet tundra north of 50°N was 64.5 mg m^{-2} d^{-1} and ranged between 31.9 and 100.6 mg m^{-2} d^{-1}. This is likely a result of our estimation method, which relies on a large degree of spatial and temporal averaging which smooths out the high CH_4 bursts that can be captured by flux chambers. ~~Our estimates are also comparable to flux observations in 2013–2014 from eddy-covariance towers on the North Slope, where the monthly mean emissions from May to September ranged from 6–21 mg m^{-2} d^{-1} (Zona et al., 2016) , as well as from the Yukon River Delta (25 mg m^{-2} d^{-1}) (Fan et al., 1992) .~~

The mean monthly flux was much lower ~~from boreal regions~~ in the boreal regions than in the tundra. On average it was 10 ± 2 mg m^{-2} d^{-1}, and ranged from 4 – 21 mg m^{-2} d^{-1}. ~~The database by Olefeldt et al. (2013) variously classified areas within the boreal region as bogs, fens and palsas, which emit on average 7 – 37 mg m^{-2} d^{-1} (Olefeldt et al., 2013) . An additional study, which made flux measurements using~~ This average is at the higher end of fluxes observed by an eddy covariance tower within a poorly drained forested region in Alaska, which found that CH_4 flux varied from 3 – 11 mg m^{-2} d^{-1} in the wettest regions (Iwata et al., 2015). However, it is reasonably similar to fluxes from the database by Olefeldt et al. (2013) , which variously classified areas within the boreal region as bogs, fens and palsas, which emit on average 7 – 37 mg CH_4 m^{-2} d^{-1} (Olefeldt et al., 2013) . Again, as with the tundra ecosystem, our spatial and temporal averaging method will smooth any CH_4 bursts.

Considering the period of May–September, the total flux from both ecosystems (2.2 ± 0.2 Tg) was consistent with the mean of the net budget calculated in the previous section assuming the ecosystems were identical. This analysis also estimates that the tundra eco-region, which represented 18% of the total study area, accounted for more than half of the total net flux.

Nevertheless, emissions from boreal regions cannot be neglected in estimates of the regional budget since their spatial coverage is quite extensive.

In Fig. 6, emissions from boreal regions appear to lag tundra regions by one month. This offset could be due to a more rapid onset of the spring thaw in the tundra eco-region, with maps of freeze-thaw state showing that this area thawed 4–5 days earlier than the boreal eco-region in 2013 and 2014 (Steiner et al., 2015). While this is less than 1 week, a study relating the date of thaw to the annual radiation budget estimated that a 4-day shift in freeze/thaw date alters the annual radiation budget by 250 MJ m^{-2}, significantly altering the early season budget (Stone et al., 2002). However, in 2012, the boreal eco-region actually thawed 9 days earlier than the tundra. Considering that fluxes calculated from all years are included in this regression, it is not clear that the difference in the seasonal pattern could be explained by differences in thaw date alone. Instead, the difference in the seasonal cycle is more likely a result of the distribution of wetlands. Using a map of wetland extent, it is estimated that 30% of the tundra eco-region can be classified as a seasonal wetland, while only 15% of the boreal eco-region is similarly classified (Bergamaschi et al., 2007). As wetlands are defined by a near-surface water table, methanogenesis can begin when the depth of thaw is much shallower, resulting in ~~a~~ CH_4 emissions that ~~can~~ begin much earlier than in regions with a deeper water table. These results and those in Sect. 5.5 are dependent on the wetland map chosen for this analysis. A more in depth comparison and evaluation of wetland maps in this region can be found in Miller et al. (2016) . In Sect. 5.5, the effect of the water table on the relationship between temperature and net CH_4 flux ~~will be discussed~~ is explored further.

5.4.2 Mountain & ocean eco-regions

~~To test~~ Up to this point, we have made explicit reference to our assumption that ~~emissions from oceans and mountains did not significantly influence measured~~ net CH_4 ~~, we calculated fluxes assuming any surface (ocean or terrestrial) could emit .The~~ flux only originates from either the tundra or boreal eco-regions defined in Fig. 1. We tested this assumption by performing the same eco-region regression ~~was used but included the fraction of the footprint which covered the~~ but included all surfaces: tundra, boreal, mountain, and ocean regions.

~~Oceans~~ Regression analysis from these calculations found that oceans were a weak CH_4 source, emitting an average 2 ± 1 mg m^{-2} d^{-1} across the study months. However, Student t-testing found that the fraction of oceans sampled had a consistently large p-value across the season (p>0.05), suggesting that footprint sensitivity in oceans was not correlated with ΔCH_4. This is consistent with a recent study of summertime sea-air flux of CH_4 around Svalbard, Norway, which measured low boundary layer CH_4 enhancements despite substantial surface ocean concentrations of CH_4 from subsea clathrate deposits in the high ~~arctic~~ Arctic (Myhre et al., 2016).

~~This regression also indicated~~ Results from this regression also indicate that mountains might act as a weak seasonal sink of CH_4, with a mean strength of -1 $\pm$ 2 mg m^{-2} d^{-1}, consistent with recent regional observations of CH_4 in the mineral soils of Greenland (Jørgensen et al., 2015). However, the statistics (p>0.1) were again too weak to confirm this as a regional scale phenomenon in our study region. In light of these results, the oceans and mountains were masked from CH_4 flux estimates, as their inclusion would have ~~diluted the contributions of other land types~~led to an underestimation of the net CH_4 flux attributed to non-mountainous land surfaces.

5.5 Temperature dependence of CH_4 flux

The results of the regression analysis described in Sect. 4.6 are listed in Table 1 and indicate the ~~r²~~ r^2 of the best model, the driving variables, and the functional type for each subset (see Eq. 7). Overall, net CH_4 flux was seen to most strongly correlate to variance in a Boltzmann-Arrhenius function of T_{40} and the inverse of elevation ($r^2 = 0.36$). While these relationships have been observed in our study region in the past, primarily at the smaller scales of chambers and eddy-covariance towers (e.g. Olefeldt et al., 2013; Sturtevant et al., 2012; Zona et al., 2016), far fewer studies have reported them at a regional scale (Miller et al., 2016; Karion et al., 2016). These latter analyses also identified the inverse elevation (Karion et al., 2016) and sub-surface temperatures (Miller et al., 2016) to be important in explaining the variance of observed CH_4. As in the work of Karion et al. (2016), we used a relatively simple analysis method to arrive at similar explanatory variables ~~as~~ than the more sophisticated geostatistical inversion model used by Miller et al. (2016) but with only a subset of the data. Although the ~~r²~~ r^2 from our analysis may seem low, they are comparable to values derived from regression analysis conducted for chamber studies (Olefeldt et al., 2013). Our results suggest that soil conditions that affect CH_4 flux at the more local scale are also relevant at regional scales.

Since both CH_4 production and flux can be much stronger than in wetlands, (e.g. Olefeldt et al., 2013; Sturtevant et al., 2012), the profiles were subdivided into two groups based on the sensitivity of the profiles to wetlands in the domain. When a profile's footprint sensitivity was $\geq$20% wetland by area, the profile was categorized as 'Wetland Present'. Conversely, profiles which had sensitivities $\leq$10% to wetlands by area were categorized as 'Wetland Absent'. The percent of wetland per area sampled by a profile was calculated by averaging a map of wetland fraction (Bergamaschi et al., 2007) weighted by the footprint sensitivity. It was found that for profiles in the Wetlands Absent category, ~~T_{40} and S_{40}~~ soil temperature and soil moisture at 40-cm soil depth were the key predictors in understanding the variability in net CH_4 flux ($r^2 = 0.48$), suggesting that CH_4 was being formed ~~lower~~ deeper in the soil column. By contrast, ~~for profiles in the Wetlands Present category, increases in~~ net CH_4 flux ~~were correlated with increases in~~ was best correlated with T_{10} and S_{10} ($r^2 = 0.40$) for profiles in the Wetland Present category.

These two regressions highlight the importance of understanding the depth at which methanogenesis (and methanotrophy) occurs. For instance, in non-wetlands, the water table is deeper than in wetlands. As a result, the onset of CH_4 production in non-wetland regions can significantly lag wetland areas, as the lower section of the soil column will take much longer to thaw. In fact, maps of soil temperature from NARR estimate that, on average, soil at 10-cm in the Wetland Present region thawed nearly 3 weeks before soil at 40-cm in the Wetland Absent regions (Mesinger et al., 2006) . In addition, soils at these depths and in these regions warmed very differently: T_{10} in the Wetland Present regions increased at an average rate of 0.10 K day^{-1} from point of thaw to the point of annual maximum, while T_{40} in the Wetland Absent regions increased at only 0.06 K day^{-1}. While CH_4 production may occur at any depth within an inundated soil column, CH_4 produced at lower soil depths can be transport limited (depending on bubble formation or aerenchyma); it is, therefore, intuitive that CH_4 produced in wetlands will be more directly correlated to differences in temperature near the surface. Similarly, the dependence of CH_4 production

(and ultimately, CH_4 flux) on soil moisture will be most pronounced just beneath the water table, where soil moisture is more susceptible to variability than at lower depths.

~~These dynamics~~ The freeze-thaw processes and soil microphysics discussed above can be extended to explain the seasonal cycles presented in both Figs. 4 & 6. First, as remarked in Sect. 5.4.1, the tundra eco-region possesses more ~~wetlands~~ wetland extent than the boreal. As a result, the delayed thaw of the 40-cm soil in combination with slower warming likely explains the delay in boreal eco-region CH_4 flux. Second, while maps of footprint sensitivity across all years showed relatively consistent sensitivity to the tundra in the west and southwest of Alaska (see Figs. 1 & 2), they showed that our ~~measurements were most~~ 2012 measurements were more sensitive to the North Slope of Alaska ~~in 2012,~~ relative to the boreal eco-region than in other years. On average, ∼30%, 20%, and 20% of the footprint sensitivity was from regions classified as tundra in 2012 – 2014, respectively. Since the tundra region possesses more seasonal wetlands, it is therefore not surprising that the 2012 seasonal cycle of net CH_4 flux in Fig. 4 ~~is closer to~~ closely resembles the tundra eco-region in Fig. 6~~, while the 2013 and 2014 seasonal cycles more closely resemble that of the boreal eco-region~~. However, as the tundra only represents 18% of the total area of interest it is likely that the sampling in 2013 and 2014 ~~was representative of~~ more evenly sampled the eco-regions with respect to their total area, while the sampling in 2012 may have been over-sensitive to the tundra. In particular, if we consider that early season flux in 2014 is identical to that in 2012, it is obvious that soil moisture is not the only variable at play. In fact, seasonal cycles of domain average sub-surface temperature from NARR are very near identical in timing to the seasonal cycles observed in CH_4 flux. To highlight this timing, the start of the warm season shown in Fig. 4 was defined as the day when domain averaged soil temperatures at 40-cm exceeded 273 K. In this respect, the early season of 2012 and 2014 may be more similar than previously alluded. Since we postulate that CH_4 flux is occurring at 40-cm within the boreal eco-region, which is larger by area and more heavily sampled by the campaign, it is not surprising that domain average CH_4 flux rises at the same rate in the spring of 2012 and 2014, while showing significant delay in 2013.

Overall, sub-surface soil temperature was seen to be the single best explanatory variable throughout the regression analysis. Following the work of Zona et al. (2016), the seasonal cycles of CH_4 flux are plotted versus soil temperature and coloured by days since the thaw in Fig. 7. In this figure, CH_4 flux estimates from individual profiles are shown in black points. To highlight the average seasonal trend, these individual estimates were block-averaged into ~~5-day bins of days since thaw~~ five-day bins of DST based on satellite retrievals of thaw state (coloured points), before being smoothed by a ~~lowess-filter~~ lowess filter which locally averaged 35%, or 80-days, of the seasonal cycle (coloured line) (Steiner et al., 2015). Flux estimates from May 2014 were suspected of being over-estimated (Sect. 5.3.1) and were thus excluded from this ~~plot and the regression that follows~~analysis.

Of the three depths presented in Fig. 7, the seasonal cycle of CH_4 flux has the most consistent, monotonic relationship with T_{40} throughout the study, while the other depths, especially the individual profiles (small black points) show more scatter. We conjecture that the monotonicity of the relationship between T_{40} and CH_4 we observe for our study region is reflective of the fact that CH_4 production is, on average, taking place near this soil level. Evidence for the production of CH_4 at soil depths well below the surface have been reported at the eddy-covariance scale before. In particular, the counterclockwise hysteresis loop in our observations of T_{10} and CH_4 flux is very similar to the relationship observed by eddy-covariance towers

at Ivotuk between CH_4 and T_{15} (Zona et al., 2016). ~~Wheras~~ Whereas observations of CH_4 flux in wetland sites exhibited a clockwise "hysteresis" loop, the Ivotuk site itself was much drier than other eddy-covariance sites compared in the study, leading the authors to conclude that the direction of the loop was related to whether CH_4 production was ~~occuring~~ occurring above (clockwise) or below (counterclockwise) the soil depth at which temperature was measured (Zona et al., 2016). Similarly, an eddy-covariance tower near Fairbanks, AK, also observed a counterclockwise hysteresis loop when CH_4 flux was plotted against T_{20} (Iwata et al., 2015). It is~~therefore~~ , therefore, reasonable to believe that CH_4 production is ~~occuring~~ occurring at depths of ∼40-cm in the broader Wetland Absent regions as a result of a lower water table. In particular, these drier regions play a large role in the late season (September – December) budget, as the production of CH_4 at soil depths well beneath the surface enables CH_4 flux to continue even after the surface soil has begun to freeze (Zona et al., 2016). In the specific case of Ivotuk, CH_4 emitted after the surface had frozen represented nearly 30% of the total annual budget (Zona et al., 2016). Since "Wetland Absent" regions make up 50% of the surface area of the tundra and boreal eco-regions it is important that these regions are not overlooked when modelling CH_4 emissions at a regional scale.

From Fig. 7, we fit the mean CH_4 flux and T_{40} from the 5-day means to a Boltzmann-Arrhenius equation and determined an activation energy of 0.75 ± 0.20 eV. This value is slightly lower than but still within uncertainties of the global mean activation energy of 0.96 eV (0.86 – 1.07 eV, 95% C.I.) calculated using CH_4 flux measurements from static chambers (Yvon-Durocher et al., 2014). Using T_{40} from NARR in this parameterization, we estimate May–September emissions from our study region to be 2.1 ± 0.3 Tg, 1.8 ± 0.3 Tg and 2.0 ± 0.3 Tg for 2012–2014, respectively (only integrating over non-mountainous areas). This simple model does a remarkable job of capturing the growing season budget estimated from the aircraft observations as well as the timing of the peak in CH_4 emissions. As more ~~winter time~~ winter-time measurements would be necessary to properly constrain cold season fluxes, estimates of total annual budgets based on this model are not reported.

6 Conclusions

Analysis of CH_4 column enhancements supplemented by simulated atmospheric transport allowed us to estimate the monthly mean CH_4 fluxes from our study domain (50–75°N, 130–170°W). We estimate that domain averaged net CH_4 flux ranged from 2.0–36 mg m^{-2} d^{-1} and that ~~2.1~~ 2.2 ± 0.5 Tg, ~~1.7~~ 1.9 ± 0.4 Tg, and ~~2.0~~ 2.3 ± ~~0.3~~ 0.6 Tg CH_4 were emitted from our domain for 2012–2014, respectively. These estimates were consistent with more complex statistical methods, indicating that this relatively simple analytical technique, with only a subset of the data, is sufficient for determining regional scale CH_4 emissions. The methodology and analysis that followed are therefore useful guidelines for regional monitoring programs which suggest that short, regular profiling of different eco-regions supplemented by fine-scale meteorological modelling can be sufficient to characterize the regional dynamics of the carbon cycle.

Despite the lack of spatial resolution within the CH_4 flux estimates, we were able to leverage the atmospheric transport model to inform some basic regression models on how CH_4 flux co-varied with different soil variables and characteristics. We found that when we sampled wetlands, CH_4 flux co-varied most significantly with T_{10}. Conversely, when wetlands were absent, CH_4 flux co-varied with T_{40}. These two results are consistent with observations of how the water table affects the

anaerobic production of CH_4 at small spatial scales and emphasize that it is a relevant control at regional scales. Across our study region, we were able to reasonably predict the May–September CH_4 budget using a Boltzmann-Arrhenius model relating CH_4 flux to T_{40}.

Overall, these regressions provide insight ~~in~~ into the differences in seasonal cycles observed across the years and eco-regions. Methanogenesis in wetlands (like the tundra) occur closer to the surface since the water table depth is higher. By contrast, methanogenesis occurs lower in the soil column in regions with ~~less wetlands (i.e.~~ fewer wetlands (such as boreal regions). Since surface soils will thaw earlier in the season than deeper soils, CH_4 production begins earlier in wetland regions and ends later in drier regions. As a result, campaigns like CARVE, who sample across eco-regions need to be cautious to evenly sample regions with different subsurface hydrology. Overall our results show that factors found to affect CH_4 emissions at scales of 1 m to 1 km are still relevant at the regional scale, suggesting that regional emissions can ~~scale up local scale~~ be determined by upscaling local-scale studies.

~~Finally, while meteorological differences affected the springtime freeze-thaw transition and fluxes, there was a distinct lack of interannual variability in the overall May–September budget. While observations of of flux in northern Alasks have proven to be stable at the scale of eddy covariance towers, this study provides additional regional scale evidence that biogenic fluxes in the domain of study are stable at the multi-year time scale.~~

Acknowledgements. We thank the pilots, flight crews, and NASA Airborne Science staff from the Wallops Flight Facility for enabling the CARVE Science flights. We acknowledge funding from the National Oceanic and Atmospheric Administration and Natural Sciences and Engineering Research Council of Canada (postdoctoral fellowship to R.Y.-W.C.). Computing resources for this work were provided by the NASA High-End Computing Program through the NASA Advanced Supercomputing Division at the Ames Research Center as well as ACENET, the regional advanced research computing consortium for universities in Atlantic Canada. ACENET is funded by the Canada Foundation for Innovation, the Atlantic Canada Opportunities Agency, and the provinces of Newfoundland & Labrador, Nova Scotia, and New Brunswick. Additional thanks to A. Karion, B. Daube, J. Budney, A. Dayalu, E. Gottlieb, M. Pender, J. Pittman, J. Samra, J. Chen, T. Duck and C. Perro for their help. The research described in this paper was performed as part of CARVE, a NASA Earth Ventures investigation.

5 References

Bean, B. R. and Dutton, E.: Radio meteorology, Dover Publications, 1966.

Bergamaschi, P., Frankenberg, C., Meirink, J., Krol, M., Dentener, F., Wagner, T., Platt, U., Kaplan, J., Körner, S., Heimann, M., et al.: Satellite chartography of atmospheric methane from SCIAMACHY on board ENVISAT: 2. Evaluation based on inverse model simulations, Journal of Geophysical Research: Atmospheres, 112, 2007.

Bergamaschi, P., Houweling, S., Segers, A., Krol, M., Frankenberg, C., Scheepmaker, R. A., Dlugokencky, E., Wofsy, S. C., Kort, E. A., Sweeney, C., Schuck, T., Brenninkmeijer, C., Chen, H., Beck, V., and Gerbig, C.: Atmospheric CH4 in the first decade of the 21st century: Inverse modeling analysis using SCIAMACHY satellite retrievals and NOAA surface measurements, J. Geophys. Res., 118, 7350–7369, doi:10.1002/jgrd.50480, http://dx.doi.org/10.1002/jgrd.50480, 2013.

Blake, L., Ovreas, L., Head, I., and Gray, N.: Response of Methanogens in Arctic Sediments to Temperature and Methanogenic Substrate Availability, PLoS ONE, 10, 2015.

Bromwich, D. H., Hines, K. M., and Bai, L.: Development and testing of Polar Weather Research and Forecasting model: 2. Arctic Ocean, J. Geophys. Res., 114, D08 122, doi:10.1029/2008JD010300, 2009.

Bruhwiler, L., Dlugokencky, E., Masarie, K., Ishizawa, M., Andrews, A., Miller, J., Sweeney, C., Tans, P., and Worthy, D.: CarbonTracker-CH4: an assimilation system for estimating emissions of atmospheric methane, Atmospheric Chemistry and Physics, 14, 8269–8293, doi:10.5194/acp-14-8269-2014, http://www.atmos-chem-phys.net/14/8269/2014/, 2014.

Burnham, K. P., Anderson, D. R., and Huyvaert, K. P.: AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons, Behavioral Ecology and Sociobiology, 65, 23–35, 2011.

Chan, K. M. and Wood, R.: The seasonal cycle of planetary boundary layer depth determined using COSMIC radio occultation data, Journal of Geophysical Research: Atmospheres, 118, 2013.

Chang, R. Y.-W., Miller, C. E., Dinardo, S. J., Karion, A., Sweeney, C., Daube, B. C., Henderson, J. M., Mountain, M. E., Eluszkiewicz, J., Miller, J. B., et al.: Methane emissions from Alaska in 2012 from CARVE airborne observations, Proceedings of the National Academy of Sciences, 111, 16 694–16 699, 2014.

Chen, H., Winderlich, J., Gerbig, C., Hoefer, A., Rella, C. W., Crosson, E. R., Pelt, A. D. V., Steinbach, J., Kolle, O., Beck, V., Daube, B. C., Gottlieb, E. W., Chow, V. Y., Santoni, G. W., and Wofsy, S. C.: High-accuracy continuous airborne measurements of greenhouse gases (CO_2 and CH_4) using the cavity ring-down spectroscopy (CRDS) technique, Atmos. Meas. Tech., 3, 375–386, 2010.

Chen, Y.-H. and Prinn, R. G.: Estimation of atmospheric methane emissions between 1996 and 2001 using a three-dimensional global chemical transport model, J. Geophys. Res., 111, D10 307, doi:10.1029/2005JD006058, 2006.

Chou, W. W., Wofsy, S. C., Harriss, R. C., Lin, J. C., Gerbig, C., and Sachse, G. W.: Net fluxes of CO 2 in Amazonia derived from aircraft observations, J. Geophys. Rese., 107, 4614, doi:10.1029/2001JD001295, http://www.agu.org/pubs/crossref/2002/2001JD001295.shtml, 2002.

Christensen, T.: Methane emission from Arctic tundra, Biogeochemistry, 21, 117–139, doi:10.1007/BF00000874, http://dx.doi.org/10.1007/BF00000874, 1993.

Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., Chhabra, A., DeFries, R., Galloway, J., Heimann, M., et al.: Carbon and other biogeochemical cycles, in: Climate change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, pp. 465–570, Cambridge University Press, 2014.

Collins, M., Knutti, R., Arblaster, J., Dufresne, J.-L., Fichefet, T., Friedlingstein, P., Gao, X., Gutowski, W., Johns, T., Krinner, G., Shongwe, M., Tebaldi, C., Weaver, A., and Wehner, M.: Long-term Climate Change: Projections, Commitments and Irreversibility, book section 12, p. 10291136, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, doi:10.1017/CBO9781107415324.024, www.climatechange2013.org, 2013.

Commane, R., Lindaas, J., Benmergui, J., Luus, K. A., Chang, R. Y.-W., Daube, B. C., Euskirchen, E. S., Henderson, J. M., Karion, A., Miller, J. B., Miller, S. M., Parazoo, N. C., Randerson, J. T., Sweeney, C., Tans, P., Thoning, K., Veraverbeke, S., Miller, C. E., and Wofsy, S. C.: Carbon dioxide sources from Alaska driven by increasing early winter respiration from Arctic tundra, Proc. Nat. Acad. Sci., 114, 5361–5366, doi:10.1073/pnas.1618567114, http://www.pnas.org/content/114/21/5361.abstract, 2017.

Commission for Environmental Cooperation: Ecological regions of North America: toward a common perspective, Tech. rep., Commission for Environmental Cooperation, 1997.

Dlugogencky, E.: Trends in Atmospheric Methane, Tech. rep., National Oceanic and Atmospheric Administration, www.esrl.noaa.gov/gmd/ccgg/trends_ch4/, date accessed: 01/07/2016, 2016.

Dlugokencky, E., Myers, R., Lang, P., Masarie, K., Crotwell, A., Thoning, K., Hall, B., Elkins, J., and Steele, L.: Conversion of NOAA atmospheric dry air CH4 mole fractions to a gravimetrically prepared standard scale, Journal of Geophysical Research: Atmospheres, 110, 2005.

Dlugokencky, E. J., Nisbet, E. G., Fisher, R., and Lowry, D.: Global atmospheric methane: budget, changes and dangers, Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 369, 2058–2072, 2011.

Earth Systems Research Laboratory (ESRL): NOAA/ESRL calculation of global means, Tech. rep., National Oceanic and Atmospheric Administration, http://www.esrl.noaa.gov/gmd/ccgg/about/global_means.html, date accessed: 01/07/2016, 2016.

Fan, S. M., Wofsy, S. C., Bakwin, P. S., Jacob, D. J., Anderson, S. M., Kebabian, P. L., McManus, J. B., and Kolb, C. E.: Micrometerological Measurements of CH_4 and CO_2 Exchange Between the Atmosphere and Subarctic Tundra, J. Geophys. Res., 97, 16 627–16 643, 1992.

Gatti, L. V., Gloor, M., Miller, J. B., Doughty, C. E., Malhi, Y., Domingues, L. G., Basso, L. S., Martinewski, a., Correia, C. S. C., Borges, V. F., Freitas, S., Braz, R., Anderson, L. O., Rocha, H., Grace, J., Phillips, O. L., and Lloyd, J.: Drought sensitivity of Amazonian carbon balance revealed by atmospheric measurements., Nature, 506, 76–80, doi:10.1038/nature12957, 2014.

Gerbig, C., Lin, J. C., Wofsy, S. C., Daube, B. C., Andrews, A. E., Stephens, B. B., Bakwin, P. S., and Grainger, C. A.: Toward constraining regional-scale fluxes of CO2 with atmospheric observations over a continent: 2. Analysis of COBRA

data using a receptor-oriented framework, J. Geophys. Res., 108, n/a–n/a, doi:10.1029/2003JD003770, http://dx.doi.org/10.1029/2003JD003770, 4757, 2003.

Global Monitoring Division: Carbon Cycle Greenhouse Gas Reference Network, Tech. rep., National Oceanic and Atmospheric Administration, http://www.esrl.noaa.gov/gmd/dv/site/PFA.html, date accessed: 08/12/2016, 2016.

Henderson, J. M., Eluszkiewicz, J., Mountain, M. E., Nehrkorn, T., Chang, R. Y.-W., Karion, A., Miller, J. B., Sweeney, C., Steiner, N., Wofsy, S. C., and Miller, C. E.: Atmospheric transport simulations in support of the Carbon in Arctic Reservoirs Vulnerability Experiment (CARVE), Atmospheric Chemistry and Physics, 15, 4093–4116, doi:doi:10.5194/acp-15-4093-2015, 2015.

Hines, K., Bromwich, D., Bai, L., Bitz, C., Powers, J., and Manning, K.: Sea ice enhancements to Polar WRF, Monthly Weather Review, 143, 2363–2385, doi:doi: 10.1175/MWR-D-14-00344.1, 2015.

Hines, K. M. and Bromwich, D. H.: Development and Testing of Polar Weather Research and Forecasting (WRF) Model. Part I: Greenland Ice Sheet Meteorology*, Mon. Weather Rev., 136, 1971–1989, doi:10.1175/2007MWR2112.1, 2008.

Hines, K. M., Bromwich, D. H., Bai, L.-S., Barlage, M., and Slater, A. G.: Development and Testing of Polar WRF. Part III: Arctic Land, J. Climate, 24, 26–48, doi:10.1175/2010JCLI3460.1, 2011.

Hugelius, G., Tarnocai, C., Broll, G., Canadell, J., Kuhry, P., and Swanson, D.: The Northern Circumpolar Soil Carbon Database: spatially distributed datasets of soil coverage and soil carbon storage in the northern permafrost regions, Earth System Science Data, 5, 3–13, 2013.

Iwata, H., Harazono, Y., Ueyama, M., Sakabe, A., Nagano, H., Kosugi, Y., Takahashi, K., and Kim, Y.: Methane exchange in a poorly-drained black spruce forest over permafrost observed using the eddy covariance technique, Agricultural and Forest Meteorology, 214, 157–168, 2015.

Jenness, J. S.: Calculating landscape surface area from digital elevation models, Wildlife Society Bulletin, 32, 829–839, 2004.

Johnston, C. E., Ewing, S. A., Harden, J. W., Varner, R. K., Wickland, K. P., Koch, J. C., Fuller, C. C., Manies, K., and Jorgenson, M. T.: Effect of permafrost thaw on CO 2 and CH 4 exchange in a western Alaska peatland chronosequence, Environmental Research Letters, 9, 085 004, http://stacks.iop.org/1748-9326/9/i=8/a=085004, 2014.

Jørgensen, C. J., Johansen, K. M. L., Westergaard-Nielsen, A., and Elberling, B.: Net regional methane sink in High Arctic soils of northeast Greenland, Nature Geoscience, 8, 20–23, 2015.

Jorgenson, M. T., Shur, Y. L., and Pullman, E. R.: Abrupt increase in permafrost degradation in Arctic Alaska, Geophysical Research Letters, 33, 2006.

Karion, A., Sweeney, C., Wolter, S., Newberger, T., Chen, H., Andrews, a., Kofler, J., Neff, D., and Tans, P.: Long-term greenhouse gas measurements from aircraft, Atmospheric Measurement Techniques, 6, 511–526, doi:10.5194/amt-6-511-2013, 2013.

Karion, A., Sweeney, C., Miller, J. B., Andrews, A. E., Commane, R., Dinardo, S., Henderson, J. M., Lindaas, J., Lin, J. C., Luus, K. A., Newberger, T., Tans, P., Wofsy, S. C., Wolter, S., and Miller, C. E.: Investigating Alaskan methane and carbon dioxide fluxes using measurements from the CARVE tower, Atmospheric Chemistry and Physics, 16, 5383–5398, doi:10.5194/acp-16-5383-2016, http://www.atmos-chem-phys.net/16/5383/2016/, 2016.

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

Lin, J., Gerbig, C., Wofsy, S., Andrews, A., Daube, B., Davis, K., and Grainger, C.: A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model, Journal of Geophysical Research: Atmospheres (1984–2012), 108, 2003.

Melton, J., Wania, R., Hodson, E., Poulter, B., Ringeval, B., Spahni, R., Bohn, T., Avis, C., Beerling, D., Chen, G., et al.: Present state of global wetland extent and wetland methane modelling: conclusions from a model intercomparison project (WETCHIMP), Biogeosciences, 10, 753–788, 2013.

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P. C., Ebisuzaki, W., Jović, D., Woollen, J., Rogers, E., Berbery, E. H., et al.: North American regional reanalysis, Bulletin of the American Meteorological Society, 87, 343–360, 2006.

Miller, S., Miller, C., Commane, R., Chang, R.-W., Dinardo, S., Henderson, J., Karion, A., Lindaas, J., Melton, J., Miller, J., Sweeney, C., Wofsy, S., and Michalak, A.: A multi-year estimate of methane fluxes in Alaska from CARVE atmospheric observations, Global Biogeochemical Cycles, in review, 2016.

Myhre, C. L., Ferré, B., Platt, S., Silyakova, A., Hermansen, O., Allen, G., Pisso, I., Schmidbauer, N., Stohl, A., Pitt, J., et al.: Extensive release of methane from Arctic seabed west of Svalbard during summer 2014 does not influence the atmosphere, Geophysical Research Letters, 43, 4624–4631, 2016.

Nehrkorn, T., Eluszkiewicz, J., Wofsy, S. C., Lin, J. C., Gerbig, C., Longo, M., and Freitas, S.: Coupled weather research and forecasting–stochastic time-inverted lagrangian transport (WRF–STILT) model, Meteorology and Atmospheric Physics, 107, 51–64, 2010.

Novelli, P., Collins, J., Myers, R., Sachse, G., and Scheel, H.: Reevaluation of the NOAA/CMDL carbon monoxide reference scale and comparisons with CO reference gases at NASA-Langley and the Fraunhofer Institut, Journal of Geophysical Research: Atmospheres, 99, 12 833–12 839, 1994.

Olefeldt, D., Turetsky, M. R., Crill, P. M., and McGuire, A. D.: Environmental and physical controls on northern terrestrial methane emissions across permafrost zones, Global change biology, 19, 589–603, 2013.

O'Shea, S. J., Allen, G., Gallagher, M. W., Bower, K., Illingworth, S. M., Muller, J. B. A., Jones, B. T., Percival, C. J., Bauguitte, S. J.-B., Cain, M., Warwick, N., Quiquet, A., Skiba, U., Drewer, J., Dinsmore, K., Nisbet, E. G., Lowry, D., Fisher, R. E., France, J. L., Aurela, M., Lohila, A., Hayman, G., George, C., Clark, D. B., Manning, A. J., Friend, A. D., and Pyle, J.: 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, 13 159–13 174, doi:10.5194/acp-14-13159-2014, http://www.atmos-chem-phys.net/14/13159/2014/, 2014.

Osterkamp, T.: The recent warming of permafrost in Alaska, Global and Planetary Change, 49, 187–202, 2005.

Overland, J., Hanna, E., Hanssen-Bauer, I., Kim, S.-J., Walsh, J., Wang, M., Bhatt, U., and Thoman, R.: Arctic Report Card 2015: Surface Air Temperatures, Tech. rep., National Oceanic and Atmospheric Administration, http://www.arctic.noaa.gov/reportcard, 2015.

Sasakawa, M., Shimoyama, K., Machida, T., Tsuda, N., Suto, H., Arshinov, M., Davydov, D., Fofonov, A., Krasnov, O., Saeki, T., Koyama, Y., and Maksyutov, S.: Continuous measurements of methane from a tower network over Siberia, Tellus B, 62, 403–416, doi:10.1111/j.1600-0889.2010.00494.x, http://www.tellusb.net/index.php/tellusb/article/view/16583, 2010.

Schuur, E., McGuire, A., Schädel, C., Grosse, G., Harden, J., Hayes, D., Hugelius, G., Koven, C., Kuhry, P., Lawrence, D., et al.: Climate change and the permafrost carbon feedback, Nature, 520, 171–179, 2015.

Skamarock, W., Klemp, J., Dudhia, J., Gill, D., Barker, D., Wang, X.-Y., Wang, W., and Powers, J.: A Description of the Advanced Research WRF Version 3, Tech. rep., MMM Division, NCAR, http://www.mmm.ucar.edu/wrf/users/docs/arw_v3.pdf, 475+STR, 2008.

Steiner, N., McDonald, K., Dinardo, S., and Miller, C.: Snowmelt and Surface Freeze/Thaw Timings over Alaska derived from Passive Microwave Observations using a Wavelet Classifier, American Geophysical Union, San Francisco, CA, 14-18 December 2015, 2015.

Stone, R. S., Dutton, E. G., Harris, J. M., and Longenecker, D.: Earlier spring snowmelt in northern Alaska as an indicator of climate change, Journal of Geophysical Research: Atmospheres (1984–2012), 107, ACL–10, 2002.

Stull, R. B.: An introduction to boundary layer meteorology, vol. 13, Springer Science & Business Media, 1988.

Sturtevant, C., Oechel, W., Zona, D., Kim, Y., and Emerson, C.: Soil moisture control over autumn season methane flux, Arctic Coastal Plain of Alaska, Biogeosciences, 9, 1423–1440, 2012.

Sweeney, C., Dlugokencky, E., Miller, C., Wofsy, S., Karion, A., Dinardo, S., Chang, R. Y.-W., Miller, J., Bruhwiler, L., Crotwell, A., et al.: No significant increase in long-term CH4 emissions on North Slope of Alaska despite significant increase in air temperature, Geophysical Research Letters, 2016.

Tarnocai, C., Canadell, J., Schuur, E., Kuhry, P., Mazhitova, G., and Zimov, S.: Soil organic carbon pools in the northern circumpolar permafrost region, Global biogeochemical cycles, 23, 2009.

Whalen, S.: Biogeochemistry of methane exchange between natural wetlands and the atmosphere, Environmental Engineering Science, 22, 73–94, 2005.

Xue, K., M. Yuan, M., J. Shi, Z., Qin, Y., Deng, Y., Cheng, L., Wu, L., He, Z., Van Nostrand, J. D., Bracho, R., Natali, S., Schuur, E. A. G., Luo, C., Konstantinidis, K. T., Wang, Q., Cole, J. R., Tiedje, J. M., Luo, Y., and Zhou, J.: Tundra soil carbon is vulnerable to rapid microbial decomposition under climate warming, Nature Climate Change, http://dx.doi.org/10.1038/nclimate2940, 2016.

Yvon-Durocher, G., Allen, A. P., Bastviken, D., Conrad, R., Gudasz, C., St-Pierre, A., Thanh-Duc, N., and Del Giorgio, P. A.: Methane fluxes show consistent temperature dependence across microbial to ecosystem scales, Nature, 507, 488–491, 2014.

Zhang, J. and Rothrock, D.: Modeling global sea ice with a thickness and enthalpy distribution model in generalized curvilinear coordinates, Monthly Weather Review, 131, 845–861, 2003.

Zhao, C. L. and Tans, P. P.: Estimating uncertainty of the WMO mole fraction scale for carbon dioxide in air, Journal of Geophysical Research: Atmospheres, 111, 2006.

Zona, D., Gioli, B., Commane, R., Lindaas, J., Wofsy, S. C., Miller, C. E., Dinardo, S. J., Dengel, S., Sweeney, C., Karion, A., et al.: Cold season emissions dominate the Arctic tundra methane budget, Proceedings of the National Academy of Sciences, 113, 40–45, 2016.

Table 1. Correlation coefficients and predictors of CH_4 emissions for various linear regression models

N	Subset Condition	r^2	Predictors	Type
		All		
68	–	0.36	T_{40}, Z	Boltz.-Arrh., Inv.
		Sub-sets		
27	Wetlands Present	0.40	T_{10}, S_{10}	Boltz.-Arrh., Lin.
28	Wetlands Absent	0.48	T_{40}, S_{40}	Boltz.-Arrh., Lin.

T_x: mean sampled x-cm subsoil temperature from NARR (K); S_x: mean sampled x-cm subsoil liquid moisture fraction from NARR (-); Z: soil surface elevation above sea-level (km).

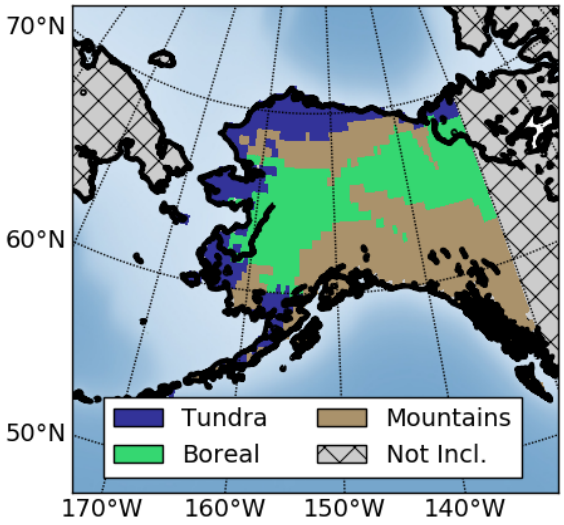

Figure 1. Eco-regions derived from the Commission for Environmental Cooperation Level II Terrestrial Eco-regions. The study region is defined by the coastline and filled eco-regions.

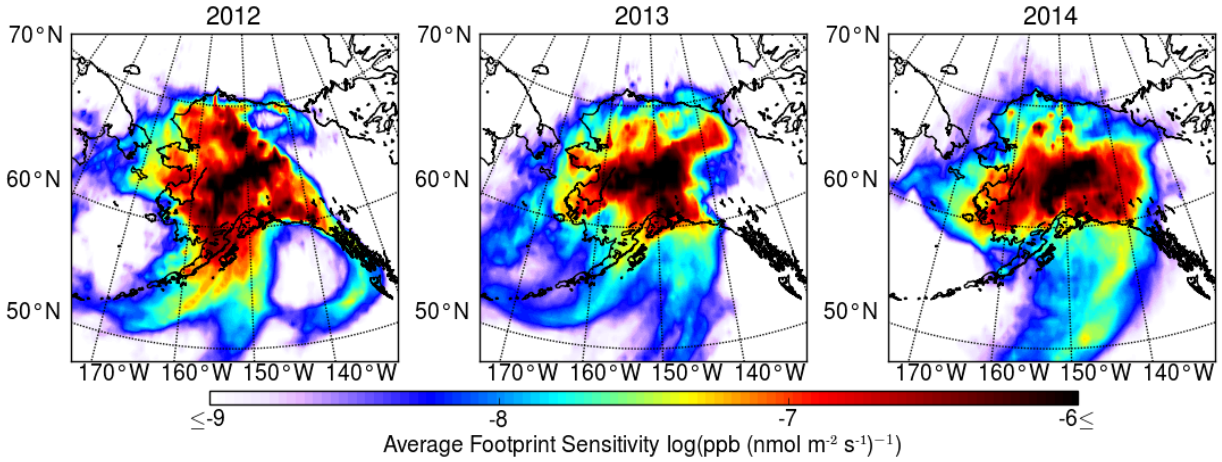

Figure 2. The average footprint sensitivity calculated from WRF-STILT is shown for all receptor points modelled within the profiles included in the analysis (2012–2014).

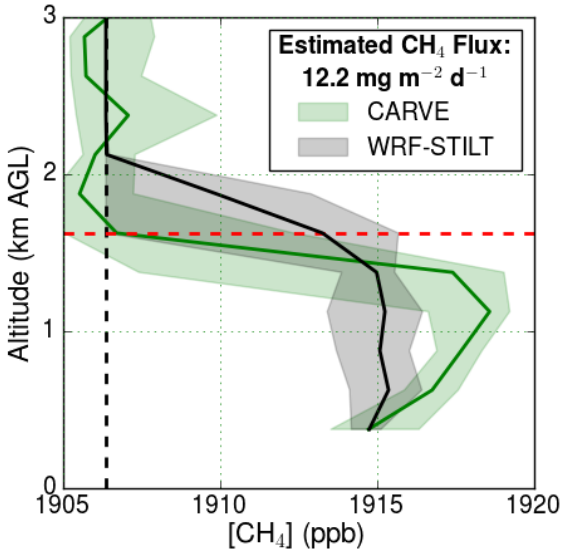

Figure 3. A sample profile from 3 Sep 2014 at 65°N, 148.6°W. Shaded regions denote the minimum/maximum of the ranges and the solid line is the median. The dashed red line represents h, and the dashed black line is $[\mathrm{CH_4}]_0$.

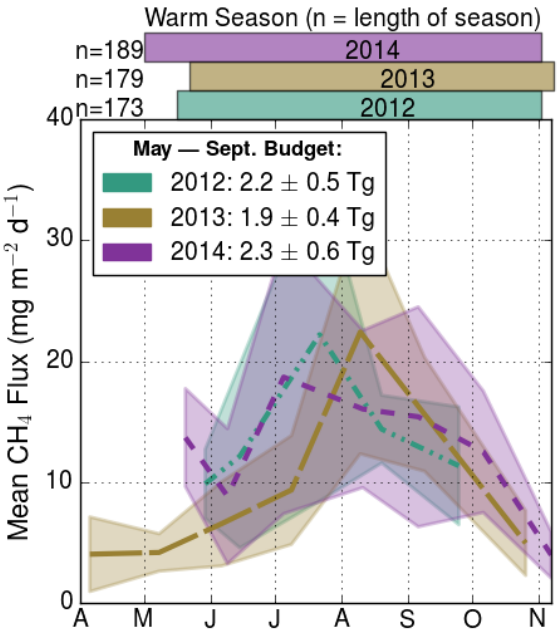

Figure 4. Monthly mean CH_4 flux estimates. Values are centered on the mean measurement date for a given month and the shaded regions are the standard ~~error~~ deviations of the ~~mean~~individual flux estimates weighted by the 95% C.I. for each estimate. See text for additional details. The bar plot above the graph marks the days when average soil temperatures from NARR were above zero, with the total number of unthawed days given by n.

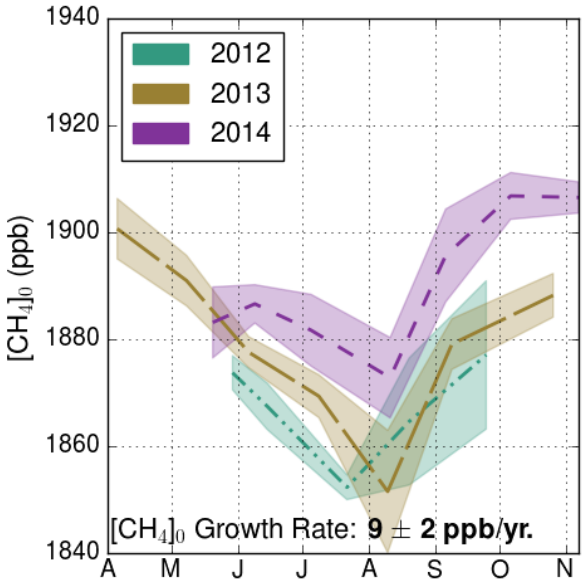

Figure 5. The mean tropospheric $[CH_4]_0$ estimated from profiles (shaded areas indicate 1σ). Values are centered on the mean measurement date.

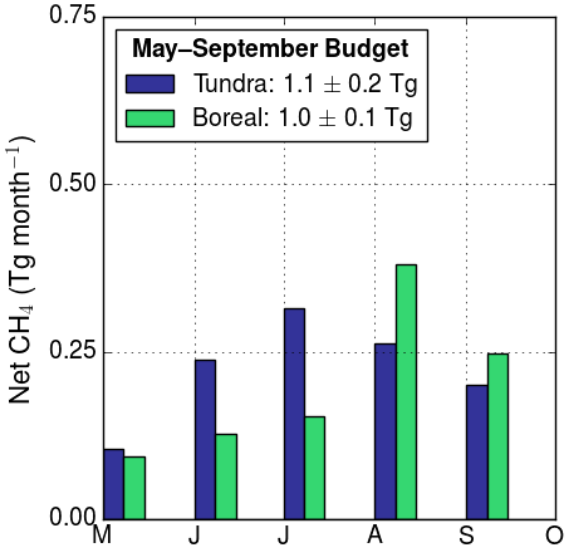

Figure 6. Estimated CH_4 emissions from tundra and boreal eco-regions averaged over all years.

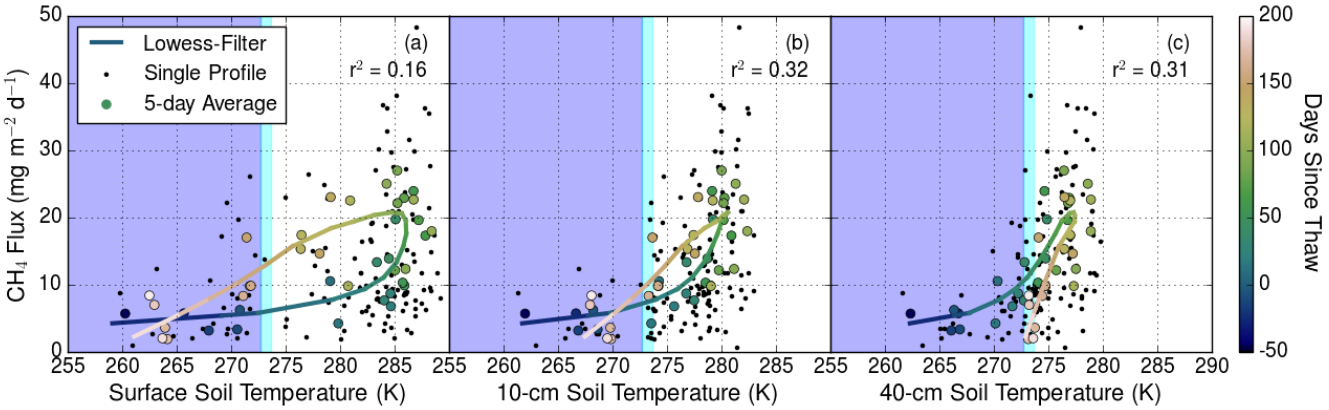

Figure 7. (a – c) CH_4 flux estimates from atmospheric profiles are shown versus footprint-weighted mean soil temperatures at different depths. The shaded background denotes when the soil temperature was at (cyan) or below (blue) the fusion point of water.