Final author response to referee comments on **"Arctic Regional Methane Fluxes by Ecotope as Derived Using Eddy Covariance from a Low Flying Aircraft" by David S. Sayres et al.**

1. ***Comment from Referee 1:*** *The chosen flux fragment method disregards a fundamental micrometeorological underpinning that permits relaxing what is formally a continuity equation to a single "eddy covariance" term: Atmospheric turbulence needs to be sampled from the high-frequency dissipation range to the first low-frequency spectral gap (e.g., Foken, 2008). As mentioned by the authors, the latter would necessitate the analysis of several kilometres of flight data at once. Nevertheless, the authors revert to the use of 60 meters at a time, thus disembodying a few numbers at a time from their theoretical foundation (here: discarding the spatially varying "base state").*
   Comment from Referee 2: *The authors used a method called flux fragment method to explore the heterogeneity of the fluxes. But this method is questionable, as each flux calculation only consider data points in a very short period (1 s) and low frequency parts of the fluxes are totally ignored in the calculation.*

   1.1. **Author response:** A common misconception about the Flux-Fragment Method (FFM) is that the fragments contain no information on scales larger than their length (FOCAL used 60 m). To be sure, fragments formed using departures from local 60-m averages would jettison all larger-scale contributions, but these fragments use departures from the 3-km base state not local averages. Sections 2.2 and 2.3 as written described the method correctly, but insufficiently emphasized this point. The scale of the base state is determined by ogive analysis (Foken, 2008) to be an upper limit for the turbulence present at the time of measurement. The fragments therefore contain information on all scales from the Nyquist wavelength of the sample rate up to the 3-km scale of the spectral gap determined from the ogive analysis. Yet, the air packets quantified by the fragments are also short enough to have likely interacted with a single class of surface. All fluxes defined in the paper are formed from sums of at least 50 fragments, enough to have a cumulative length of at least 3 km, usually more.

   1.2. **Changes to manuscript:** A statement about how the base state for the fragment is made will be added to Page 6, Line 28. "Departure quantities used to form the fragments are relative to a base-state of 3-km scale or more, a scale determined by ogive analysis (Foken, 2008) to be an upper limit for the turbulence present at the time of measurement. The fragments therefore contain information on all scales from the Nyquist wavelength of the sample rate up to the 3-km scale of the spectral gap determined from the ogive analysis. Yet, the air packets quantified by the fragments are also short enough to have likely interacted with a single class of surface."
   Add Page 7, Line 4. "Fluxes are calculated only for those surface class groups whose total length is greater than 3 km."

2. ***Comment from Referee 1:*** *Superior space-frequency decomposition techniques are widely available and in use (e.g., Barnhart et al., 2012; Strunin and Hiyama, 2004; Thomas and Foken, 2005; van den Kroonenberg and Bange, 2007). These are not only theoretically sound, but provide better spatial resolution down to meters and do not suffer from the loss of low-frequency contributions.*
   Comment from Referee 2: *As pointed by the other reviewer, other promising methods are available for investigate heterogeneous fluxes, such as wavelet analysis.*

   2.1. **Author response:** The space-frequency decomposition techniques mentioned by reviewer 1 are based in well-developed mathematical theory. Such work as Farge (1993) and Torrence and

Compo (1999) have made multi-resolution continuous-wavelet transforms highly useful to the treatment of turbulence in general and atmospheric turbulence in particular. However, eddy covariance is also theoretically sound and has long been treated successfully (Foken (2008)). The eddy covariance approach remains useful in providing a different, more directly intuitive, physical perspective in the space/time domain. The FFM is a modification of the canonical eddy-covariance. It is unorthodox in its use of conditional sampling to pluck individual fragments from the data stream at will to be combined into a mean covariance. This produces gaps not normally tolerated in space/time eddy-covariance work. The traditional analysis takes advantage of the autocorrelation of the data stream. This advantage is to some extent sacrificed in the FFM, but a large-enough random sample of departure quantities, defined as in response 1.1 above, will produce a meaningful estimate of the flux on all scales of turbulence present in the boundary layer.

Procedures exist to estimate the uncertainty in averages computed over a serially correlated, unevenly spaced data stream (eg. Mudelsee, 2010, Chapter 3).

So long as any significant secondary circulations are accounted in the base-state, the turbulent atmosphere on all its scales can be postulated to repeat over the landscape in a fairly random fashion. A contiguous sample (i.e, without gaps) should not therefore be required. The sample only need be sufficiently large to include multiple instances of boundary-layer structures at each scale. An aircraft moving at airspeed 60 m s^{-1} covers 216 km in an hour encountering 72 instances of 3-km turbulence structure. A sufficiently prevalent class of land surface, whether found in large or small patches is very likely to provide a sufficient sample. Samples which are too short can be discovered in confidence intervals developed by bootstrap resampling as was done by Kirby et al (2008). A more sophisticated bootstrap procedure developed in conjunction with analysis of these 2013 data follows Mudelsee (2010, Chapter 3). A manuscript describing the approach in detail has been submitted to the Journal of Atmospheric and Oceanic Technology and is in review.

In drawing randomly spaced samples from an autocorrelated data series the FFM does sacrifice some efficiency. A contiguous series (or a multi-scale wavelet reconstruction thereof) can take advantage of whatever coherency is contained in the feature it is sampling, though it still must sample several such features to provide an adequate estimator of the mean turbulence and flux found in the study area.

The FFM in the space/time domain is a statistical approach as opposed to decomposition approaches like the wavelet and Empirical-Mode decompositions. Also, being totally in the space/time domain FFM can in principle provide spatial resolution down to whatever scale is required so long as the small-scale features are repeated sufficiently often. Of course, the range of practically realizable scales will depend on the instrumentation used to make the measurements which is the same for wavelet analysis.

Therefore, we see no justification for the referee's conclusion that wavelet analysis is superior to eddy covariance/FFM. The FFM provides the same spatial resolution, and it does not suffer from the loss of low-frequency contributions suggested by the reviewer.

2.2. **Changes to manuscript:** We do not agree with the reviewers that a change in methodology is

needed. While it might be interesting for a future paper to compare the various approaches and their results, that is not the intent of our present manuscript.

3. ***Comment from Referee 1:*** *Next, relating atmospheric fluxes to discrete land cover classes alone neglects intra-class variability (e.g., Beyrich et al., 2006; Ogunjemiyo et al., 2003). This is better expressed with continuous land cover properties such as temperature, vegetation indices etc. (e.g., Glenn et al.,2008; Ogunjemiyo et al., 1997).*

3.1. **Author response:** "Better expressed" is a relative assessment, dependent on the question being asked. One may in fact want to determine the intra-class variability to assess the representativeness of a surface site located in a particular land-use or land-cover class identifiable by remote sensing. Intra-class variability is expressed in our results by the confidence intervals, which as long as the number of fragments is large, mostly represents the variability within that class. One of the goals of the paper is to compare with towers and other published measurements which classify methane flux based on surface classes similar to those used in this paper.

Specifically, the reviewers' suggestion to use NDVI would be inappropriate for methane measurements. It works somewhat well for CO2 flux because CO2 has a known causal relationship with photosynthesis and plant respiration. Methane is not primarily controlled by the physiology of the vegetation. Vegetation type may, however, serve as a proxy reflecting different soil moisture and other properties. Also the roots of sedge are known to act as a passive transport for methane bypassing any oxidation that might otherwise occur in the surface soil. Perhaps a different interval quantity can provide a meaningful correlation to methane flux, e.g. soil moisture, water-table height, or (sub-canopy) soil temperature. These are hard to measure remotely, especially with the accuracy needed. They were not available during the mission, nor do the authors know of a way to do this remotely at the spatial scale necessary. Failing that, we are using surface cover as a proxy for subsurface hydrology. Ignoring any assumptions about subsurface features, our results still show what sort of surface cover is associated with the strongest methane flux. We found wet sedge to dominate CH4 emission when the soil was warm. In particular, it was much more important than open water such as thermokarst lakes, which have garnered much attention based on the work of Walters-Anthony and others.

3.2. **Changes to manuscript:** modify sentence Page 7, Line 26. "These classifications, assigned based on remotely sensed data, are plausible proxies for properties that have been shown to be primary drivers of methane production and emission such as water table height, soil temperature, and emission pathways such as sedge roots. Interval quantities sensible remotely such as NDVI, air temperature, or other vegetative indexes which correlate with carbon dioxide do not correlate with methane (Olefeldt, 2013). Vegetation classifications such as these have been shown to be useful for estimating regional methane emissions from other regions (eg. Schneider, 2009) though those were based on upscaling from ground measurements."

Page 6, line 23. "The Flux Fragment Method (FFM) was conceived to answer questions concerning the homogeneity of land classes defined by some remotely sensible measurement in areas where the land classes vary on lengths short compared to what would be needed for a traditional running flux calculation. "

Page 3, line 14. “How representative were towers’ footprints of the class of land cover, as identified by remote sensing, in which they were placed? In principle a stationary site can measure all manner of properties and state variables in the soil, the vegetation, and the air within and above the canopy. Much can be learned about the bacteria, soil chemistry, canopy storage, and other quantities relevant to the exchange of mass, momentum, and energy with the surface. But all of this is known only at the one site. How representative is that site of other locations that to remote sensors appear similar? Are there land-cover types that are particularly indicative of emission of a given trace gas? Can the class so identified be used as a quantitative predictor of a particular type of soil chemistry. This is relevant in assessing the regional methane emission from remote sensing. Methane in particular has a fairly complex chemistry in the soil involving state quantities such as the (sub-canopy) soil temperature and the height of the water table. These are measurable only in situ so that having a proxy indicator such as vegetation cover would be valuable.

Aircraft, though more limited in what they can measure than fixed sites, are very mobile providing the opportunity to sample many instances of the same remotely sensed class over the landscape. From this multi-instance sample one can assess how representative the single fixed site is. One can also assess the strength of the variability within the given land-surface class for later investigation from the surface. In remote parts of the earth, in particular, a determination of near homogeneity of emission properties from many instances of a recognizably similar surface class can save considerable effort over a surface-based survey. Alternatively, large variation within a class that is not well predicted by some practically measurable interval quantity will be seen as requiring additional effort for in-situ measurements to find an effective monitoring program for methane emission from that surface class.”

Page 7, Line 21 “The questions to be answered by the FFM, using a fuzzy-logic approach (Nguyen and Walker, 2000) to assign surface classes to fragments and then to conditionally sample them based on those classes include:

a) What is the mean flux over all measured instances of each surface class?
b) What surface classes dominate the methane emission, and by how much?
c) How much does the flux over each class vary? Is there a spatial pattern to the variation. The variability will come both from the prevailing atmospheric environment and the heterogeneity of the emission within the same class.
d) How representative is a particular instance of all similar instances over the landscape?”

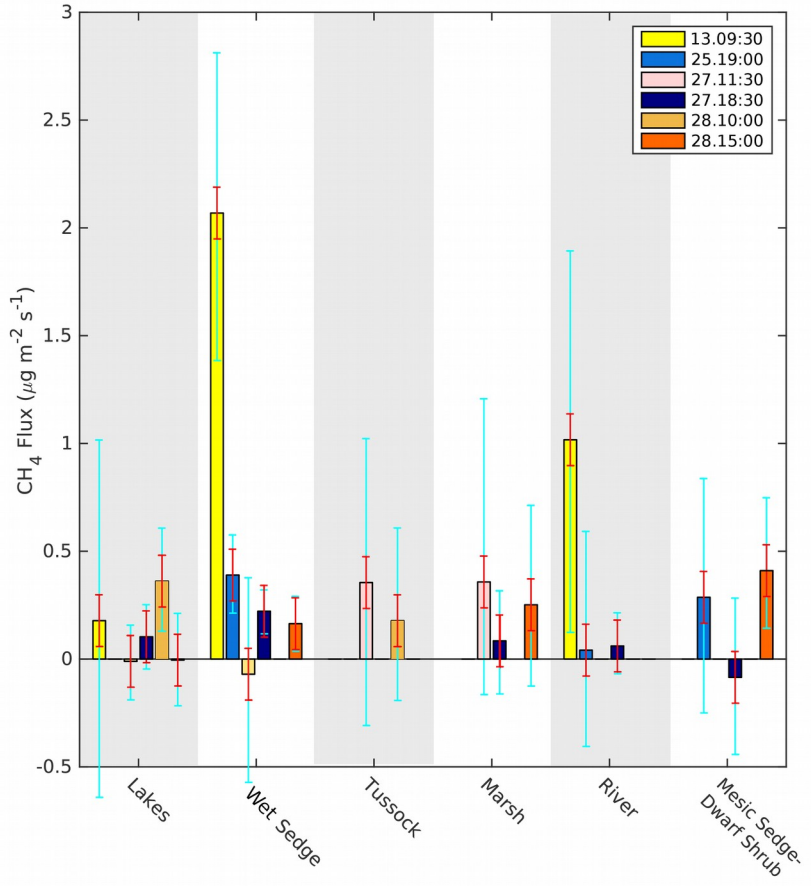
4. ***Comment from Referee 1:*** *In addition, FFM results for individual land covers are not comparable across flight days, as day-to-day synoptic variations and different flight times within the diurnal cycle are not taken into account.*

4.1. **Author response:** This comment has nothing to do specifically with FFM because the same question could be asked about eddy covariance in general from a tower or aircraft. Again, FFM is a specific implementation of eddy covariance. Synoptic variations are ideally removed by the base state, except as they affect the turbulence. Unlike CO2, methane has a weak to non-existent diurnal cycle. Our tower data do show a weak cycle, most likely caused by near surface soil temperature changes through the day. However, this diurnal cycle is an order of magnitude less than the variation due to other causes including deeper-soil temperature. It is also smaller than the difference observed between surface classes and therefore comparing flights even though

they were at different times of day is justified.

4.2. **Changes to manuscript:** Page 10, Line 4 add "Though some of the flights were in the evening (1800 – 1900 local time) and some in the morning, these data are still comparable. Unlike CO2, methane has a weak to non-existent diurnal cycle (Figure 4). Based on our tower data we do show a very weak cycle, most likely caused by near-surface soil-temperature changes through the day. However, this diurnal cycle is much weaker (<0.2 ug m-2 s-1) than the class to class variations, seasonal variations, or variations due to other factors. Therefore, comparing flights even though they were at different times of day is justified. The sharp feature in the tower trace on August 13 (DOY 225) probably has a diurnal component, The important comparison, however, is between the strong methane flux in the summer regime of first half of August and the much weaker flux in the autumn regime of later August after the major reduction in soil temperature."

Figure 7 has been altered to give better evidence of which flights occurred in the daytime and which in the evening.



5. ***Comment from Referee 1:*** *Moreover, FFM acts as a filter reducing the use of available data by order 50%, i.e. it is wasteful with respect to data use efficiency.*

5.1. **Author response:** The FFM retains all data suitable for flux calculation. The data are simply

stored and used as 60-s sums of cross products, a convenient form flexible enough to allow many different treatments. The particular approach used in this paper selects a subset of these fragments to address the question being asked, which is to identify discrete land-cover classes that stand out in their contribution to landscape-wide emission of methane. The focus of the current analysis is to examine the spatially dominant land classes in their "pure" form, so rather stringent criteria were applied which, it is true, removed about half of the fragments from the analysis. The FFM was conceived to answer this question: how representative is a single fixed site of other locations on a heterogeneous surface that to remote sensors appear similar? How good is the land-cover class occupied by that site as a proxy for methane flux? The more representative of a single land class the fragment is, the more significant the differences between land classes becomes.

The FFM, however, is not limited to addressing this question alone. Fragments could just as well be associated with values of some interval quantity such as a carbon-isotope ratio, NDVI, or the fraction of footprint occupied by each of several land classes. For this study we wanted to compare to other published measurements and assess the intra-class variability. Limiting the results to a few well sampled classes was better suited to that purpose.

5.2. **Changes to manuscript:** No change.

6. ***Comment from Referee 1:*** *Figures 4, 7 show that at a 5% significance level the FFM-derived fluxes do not actually differ between land cover classes, i.e. there is more unexplained variation in the error bars than there is explained variation in the land cover means. Also here, techniques overcoming these systemic deficiencies are available and in use (Jung et al., 2011; Yang et al., 2007)*

6.1. **Author response:** 1. Not all land-surface types emit significantly different amount of methane. This is not a problem given the questions we are asking. And while many land classes have similar methane emissions, others have significant difference, e.g. mesic sedge and wet sedge on August 13, or lakes and wet sedge on August 13. For the most part, after the soil cooling, the various land classes are not distinguished in their methane release. Keep in mind the goal is not to come up with a criterion that distinguishes land class by its methane emission (or to predict land class based on methane emissions), but to measure regionally aggregated methane emissions from each of a limited number of land classes. It is reasonable that some land classes will have similar methane emissions, especially for land classes that emit little methane.

2. Broader confidence intervals reveal lower statistical power. Typically for our data set, the broader confidence intervals are associated with the shorter samples (which reduces the power). A statistical sample, to the extent that it is independent and identically distributed is a repeated drawing from the population. If a particular outcome happens only 5% of the time, then at each drawing it has a 5% chance of being realized. But with repeated drawing, the chance increases of getting at least once some outcome having a 5% chance or less. In a very large sample, each outcome having a 5% chance will occur 5% of the time. But more than 5% of a small sample will comprise some outcomes individually having a 5% chance. If one uses a bootstrap method, which assumes the realized sample to be the entire population, a disproportionate number of population members will be outcomes that in the full population would be much less likely to occur. Of course, a new measurement set will contain a comparable number of unlikely outcomes, but they will be different from those in the earlier set of measurements. Adding new

data thus reduces the overall likelihood of all low-probability events and increases the power. Unfortunately, getting a new set of measurements is **expensive.** So the tails of the distribution developed using a relatively short sample of actual measurements will be biased toward greater probability than the true population. It will therefore have wider confidence intervals (which depend on the weakness of the tails) than would the true population. Techniques have been developed to address this issue, but their implementation is not trivial. They belong to the next generation of the FFM.

3. Our measurement of wet sedge has the greatest power, Second greatest is often lakes, but may be another land-cover type. Sedge is a strong emitter, but its confidence interval is shorter in part because we have a longer sample from it.

6.2. **Changes to manuscript:** Page 10, Line 23 add "Wet sedge, followed by the Sag river, had the largest observed flux of any of the land classes sampled during the first half of August. The other land classes have smaller, more variable fluxes on most flights so that surface class alone does not distinguish them. Most likely the true variability, contributing to the large confidence intervals, is caused by heterogeneity within the surface class in sub-surface soil temperature and water table height. However, within that we can still derive a mean flux based on a large regional sample. Once the soil cools, wet sedge shows reduced, though still positive, flux of methane consistent with the other surface classes measured such as mesic sedge and lakes. The Sag river shows close to zero methane flux.

Page 11, Line 9 add "The mean methane flux from lakes sampled on a flight by flight basis shows little flux on average, except for the lakes sampled on 130828.3, which are in a different area 250 km west of the tower. Those lakes show an aggregate mean of 0.36 ug m-2 s-1 (Figure 7)"

7. ***Comment from Referee 1:*** *I suggest the authors to consider a combination of above methodologies. In fact in their introduction the authors cite Metzger et al. (2013), who demonstrate such combination specifically for the use case of airborne flux measurements.*
Comment from Referee 2: *A few recent papers have used the wavelet analysis method to determine fluxes of air pollutants in urban and oil/gas regions (Karl et al., 2009;Vaughan et al., 2015;Yuan et al., 2015). The authors are encouraged to try this method.*

7.1. **Author response:** These papers look promising as discussions of how one can operate in urban and fracking regions. As explained in the first few responses, the FFM is a reasonable and sound method for analyzing these data. Its value derives from its position as an alternate approach from a different perspective (space/time domain). A comparison of the different methodologies is an activity we hope to pursue, but that is outside the scope of this present paper.

7.2. **Changes to manuscript:** No change.

8. ***Comment from Referee 1:*** *This hemispherical model requires calibration, in the case of very low-level flight in particular to offset dynamic upwash and ground effect which otherwise affect the covariance calculation (e.g., Crawford et al., 1996; Garman et al., 2008). Have these calibrations and corrections been performed, and if so to within which residual error?*

8.1. **Author response:** We flew multiple calibration maneuvers both in preparing for and during the Alaska campaign. Before assembling the FOCAL system, we characterized the BAT (gust) probe in a wind tunnel (Dobosy et al., 2013). We also tested a similar BAT probe in flight on a different aircraft (Vellinga et al., 2013, hereafter V2013). After calibration derived from a flight taken on the evening of August 27 in Alaska, we performed the yaw maneuver described by V2013 and obtained a residual contamination within 10%, as described there. A pitch maneuver described by V2013 was performed resulting in contamination of 10% for the high-frequency pitching (1.6 s period), which was the best executed of the pitch test's three parts and is the severest test.

8.2. **Changes to manuscript:** Page 4, Line 9 insert new paragraph "Before assembling the FOCAL system, we characterized the BAT probe in a wind tunnel (Dobosy et al., 2013). We also tested a similar BAT probe in flight on a different aircraft (Vellinga et al., 2013, hereafter V2013). After the FOCAL system was assembled, similar calibration maneuvers were flown in preparation for and during the Alaska campaign. As part of a calibration flight on the evening of August 27 in Alaska, we performed the yaw maneuver described by V2013 and obtained a residual contamination within 10%, as described there. A pitch maneuver described by V2013 was performed resulting in contamination of 10% for the high-frequency pitching (1.6 s period), which was the best executed of the pitch test's three parts and is the severest test.

9. ***Comment from Referee 1:*** *Please confirm that you use CH4 dry mole fraction for the covariance / flux calculation. •In case your calculation is based on partial density, how do you correct for density variations due to temperature and humidity fluctuations (WPL), as well as variations in pressure-altitude and corresponding changes in temperature and pressure, and thus partial density (Poisson equation)?*

9.1. **Author response:** We had provided this confirmation in the manuscript, page 4, line 25, and also on page 5, line 13 citing both Webb et al. (1980) and its update, Gu et al., (2012). We will move this citation back to the first mention of the gas measurements.

9.2. **Changes to manuscript:** Page 4, Line 25, add (Webb et al., 1980; Gu et al., 2012).

10. ***Comment from Referee 1:*** *At 5 m above ground this is approximately the eddy wavelength contributing most to the turbulent vertical transport. •Using the power law of spectral decay, for this platform the need for high-frequency spectral correction of the vertical turbulent flux would be minimal only at measurement heights of 50 m above ground and higher. •As you are focusing on measurements below 25 m above ground, which high-frequency spectral correction did you use, and how large was the correction?*

10.1. **Author response:** Plots of the spectra and cospectra of the data streams of vertical air motion and the dry-air mixing ratios of the trace gases were prepared and presented in a paper that was submitted to J. Ocean. Atmos. Tech. We have not used high-frequency spectral corrections as long as the highest wavenumber for vertical wind was clearly in the inertial subrange, i.e. following the -5/3 power of the wavenumber, and clearly above the wavenumber of the maximum spectral density. A data-starvation test using the flux runs from the evening of August 25 yielded an estimated loss of about 10% in fluxes computed with a coarser sample rate. Presenting a long discussion of the the spectra and cospectra seemed out of scope for the current paper. A discussion is included in a separate paper submitted to the Journal of Atmospheric and

Oceanic Technology (JTECH).

A regression of 3-km running flux (see Section 2.3.1) against the height above ground for flight 13.09:30 was run to assess the correlation of flux with altitude. A quadratic regression was required yielding significant positive slope but significant negative curvature. The regression line reached a maximum at an intermediate point before the maximum height above ground. Furthermore, the regression explained only 10% of the variance.

10.2. **Changes to manuscript:** add above to Page 4, line 9 and after additions from 8.2 above.

11. ***Comment from Referee 1:*** *Correct, this method neglects low-frequency contributions to the vertical turbulent flux. •As mentioned above, Ogive analysis typically saturates at 100 - 1000 x measurement height. •How did you correct low-frequency loss and how large were the contributions to the flux*

11.1. **Author response:** This comment was addressed in the response the reviewers' objection 1 above.

11.2. **Changes to manuscript:** See changes from comment one above.

12. ***Comment from Referee 1:*** *How was the turbulence statistics for a robust application of the flux footprint model calculated? A 1 s flux fragment has far too large random error to assume upstream isotropy of the wind field.*

12.1. **Author response:** The turbulent statistics required to parameterize the model of Kljun et al. (2004) were computed from averages taken over the length of each flight leg, where the flight leg was defined as the straight segment between turns over which the collected data were used. The detrending (subtracting the base state from the original series) was done over each flight leg. Typically the flight legs were 15 km to 20 km.

12.2. **Changes to manuscript:** Page 7, Line 17 [Start in 13.2] "We use the parameterization scheme described in Kljun et al. (2004) which uses a backward Lagrangian model (Kljun et al., 2002) for a range of heights, stability measures and other turbulence quantities that are measured from the aircraft. The turbulence quantities are computed from averages taken over the length of each flight leg, where the flight leg is defined as the straight segment, between turns, over which the collected data were used. [continue at 13.2, second part]

13. ***Comment from Referee 1:*** *This model is 1-D and does not resolve the cross-wind distribution of the influence area – how did you take this into account? •An updated 2-D version of this model is available (Kljun et al., 2015). Why was this model not used?*

13.1. **Author response:**

Since we use the surface class as a categorical quantity the crosswind-integrated form of the footprint model of Kljun et al. (2004, KCRS04) was considered appropriate for our use as a membership function for the fuzzy set (Nguyen and Walker, 2000) of a particular surface class. The selected 85% membership criterion is strict so as to admit only particularly representative instances of the surfaces encountered.

The more recent work of Kljun & Co. (2015, KCRS15) became known to us late in our investigation. In providing an explicit crosswind distribution to the footprint it represents significant advance over KCRS04. However, the crosswind-integrated footprint of KCRS04, fundamentally unchanged, provides the backbone for the two-dimensional footprint of KCRS15. The crosswind spread may be important, for example, where an interval quantity, such as NDVI is to be calculated from the footprint of each unit of flux in order to train a regression or machine-learning model, such as done by Metzger & Co, (2013), or Ogunjemiyo & Co (2003).

The present study was not intended to produce a regression scheme. It is about the role of each surface class (as a category) in the emission of methane. Since the footprint is computed every 60 m the procedure will identify all instances of the surface classes present except for the very smallest. Expanding the footprints to two dimensions does not appear to add sufficient value to justify recalculation. The results would be unlikely to produce any changes in the results.

13.2. **Changes to manuscript:** Page 7, Line 14 "Finally, a footprint model is applied to estimate the level of influence of each surface type on each fragment. This provides a measure of membership of that fragment in the fuzzy set (Nguyen and Walker, 2000) associated with each surface type, treated as a categorical variable. Fragments having a sufficient level of membership for a particular surface class are assigned to that class. A membership level above 0.5 restricts all fragments to no more than one class. Fragments can thus be grouped into sets all members of which have a measure greater than a prespecified level of the probability that they came from the same surface type (see sec. 3.2 for examples of how FFM is used to interpret these data)" [continue at 12.2]

[second part, continued from 12.2] The more recent two-dimensional version (Kljun et al., 2015) was not considered necessary because of the footprint's restricted use as a membership criterion to assign a selected subset of fragments to the surface categories.

14. ***Comment from Referee 1:*** *There is no such website. Where can the data (incl. raw data) be accessed?*

14.1. **Author response:** The URL was missing an 's'. Should have been https://. Thanks for pointing this out.

14.2. **Changes to manuscript:** Page 12, Line 2 "https://arcticdata.io"

15. ***Comment from Referee 1:*** *There are more intuitive ways to visualize the footprint influence area. I am wondering why the authors did not use standard contour plots. Also, it is not apparent from the display whether cross-wind dispersion has been taken into consideration – the individual sequences of dots simply extend in the along-wind direction, which is only half the truth.*

15.1. **Author response:** With crosswind integrated footprints, it makes sense to plot them as lines, rather than as 2-D contour plots. However, to show the full footprint area along the flight track we have modified figure 4 to show a ribbon of footprint probabilities for one leg of each flight track for each day. Arrows have been added to show the dominate wind direction, which was observable before from the individual footprints. Hopefully this will be clearer for the reader.

15.2. **Changes to manuscript:** Figure 5 has been modified as described above.

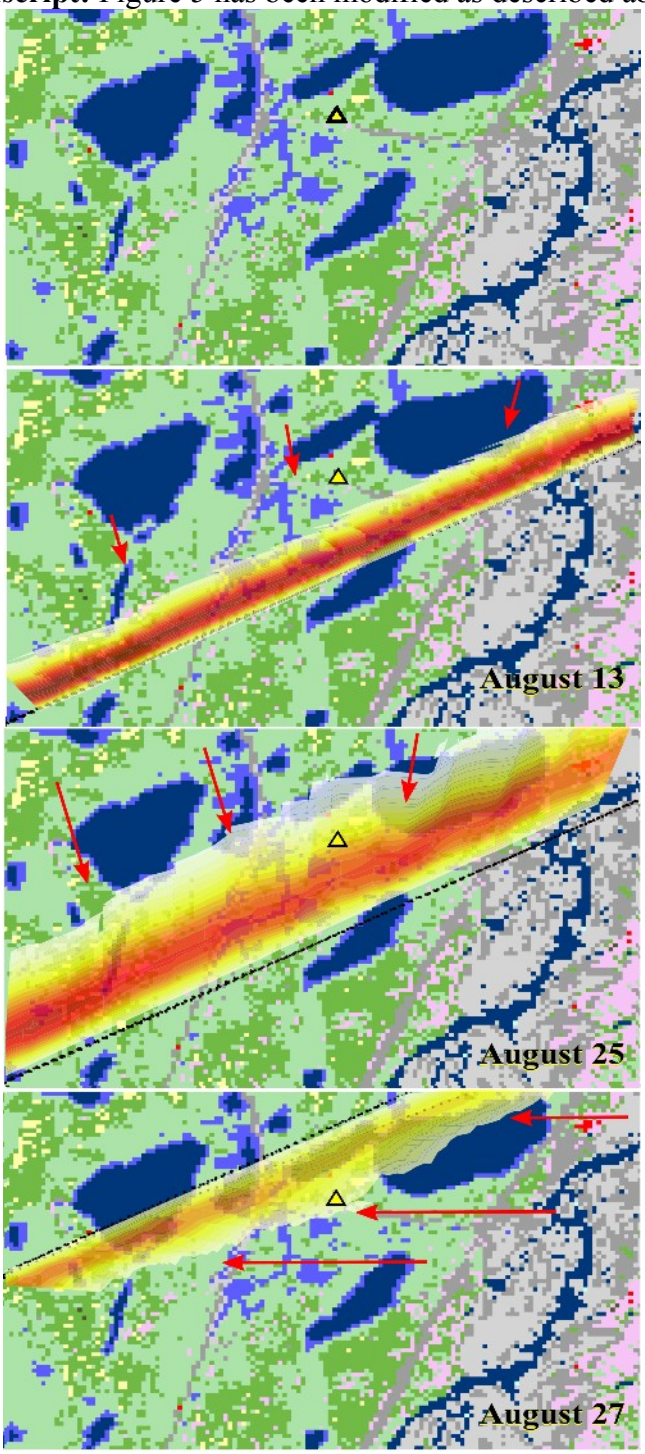


Figure 5. Map of area surrounding the flux tower (yellow triangle) with false color map representing different land classes defined as in Fig. 2. Bottom three plots show three days when data was taken near the tower. The flight track for each flight is shown as black points, where each point is the start position of a flux fragment. Colored ribbon shows the flux footprints along the flight track. The darker and redder color of the ribbon represents larger probability of contribution to the total flux as described in the text. Red arrows indicate the mean direction of the wind.

16. ***Comment from Referee 2:*** *The authors spent some time to introduce the fast measurement system of wind and CH4. Could you add some spectral analysis for measured data.*

16.1. **Author response:** See 10.1

16.2. **Changes to manuscript:** See 10.1

17. ***Comment from Referee 2:*** *Figure 4. Can you show the graph as 2*2 layout? The inserts are somewhat misleading and are hard to follow at present layout.*

17.1. **Author response:** We have modified Figure 4 by breaking it into four panels. One long panel displays the tower data, locating the three near-tower flights as before. Temporal resolution was improved by displaying only the periods when the aircraft was operating. The three insets have been relocated as individual panels underneath the tower data and are labeled by flight day instead of a,b,c. The abscissa of each is now given as (local) time of day to show the actual time of flight.

17.2. **Changes to manuscript:**

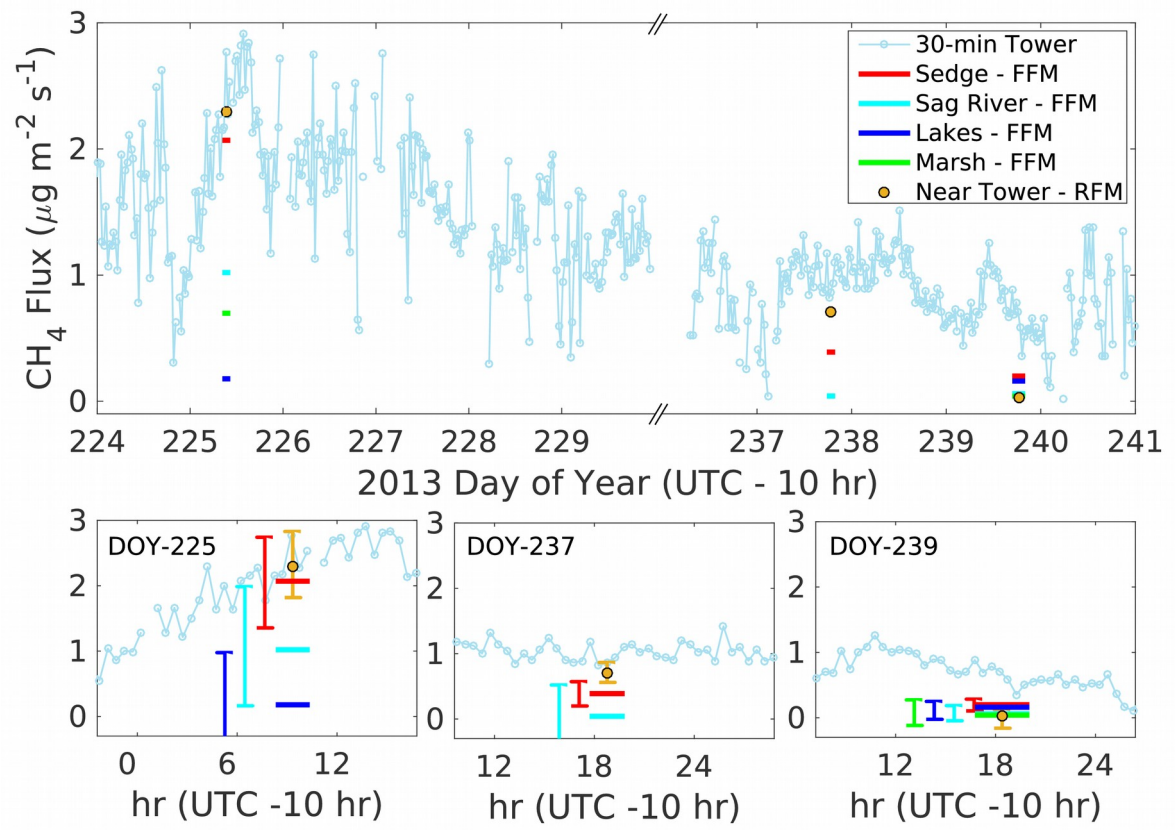


Figure 4. Comparison of methane flux measured by the flux tower with fluxes measured by the FOCAL system. Tower methane fluxes (top plot) are 30-minute means plotted versus day of year. Three flights (Aug. 13, 25, and 27) made repeated flight transects near the tower. A running mean flux, using the nearest 3 km of flight track to the tower for each leg, was calculated and the mean of these fluxes is plotted for each day as an orange circle. Fluxes for wet sedge, marsh, lakes, and the Sag river were calculated using FFM using data from the whole

flight and are plotted for each day, color coded according to the legend, with the length of the line along the time axis representing the time over which the data were taken. Bottom plots show details for each flight day, labeled by day of year (DOY), with bars showing the 95% confidence interval based on bootstrap analysis. Bars are offset along the x-axis for clarity.

18. ***Comment from Referee 2:*** *Figure 2 and Figure 7: Could you use a consistent way to indicate flight numbers conducted at the same days. Please include this information in the figure caption.*

18.1. **Author response:** It is consistent, Figure 7 just leaves off the common 1308 part, but we can add that back into the figure legend. The information is already included in the captions of fig 2 and 7 and table 1.

18.2. **Changes to manuscript:** We have modified the date convention to include the flight time and changed, Table 1 and Figures 2 and 7. The new convention uses DD.HH:MM.

References added:

Dobosy, R., E.J. Dumas, D.L. Senn, B. Baker, D.S. Sayres, M.F. Witinski, C.E. Healy, J. Munster. and J.G. Anderson, 2013: Calibration and quality assurance of an airborne turbulence probe in an aeronautical wind tunnel. Journal of Atmospheric and Oceanic Technology, 30 (2), 182–196.

Gioli, B., Miglietta, F., De Martino, B., Hutjes, R. W. A., Dolman, H. A. J., Lindroth, A., Schumacher, M., Sanz, M. J., Manca, G., Peressotti, A., and Dumas, E. J., 2004: Comparison between tower and aircraft-based eddy covariance fluxes in five European regions, *Agricultural and Forest Meteorology*, **127**, 1–16.

Kljun, N., P. Calanca, M.W. Rotach, H.P Schmid, 2015: A simple two-dimensional parameterization for Flux Footprint Prediction (FFP), *Geoscientific Model Development*, **8(11)**, 3695-3713

LeMone, M., R. Grossman, F. Chen, K. Ikeda, and D. Yates, 2003: Choosing the averaging interval for comparison of observed and modeled fluxes along aircraft transects over a heterogeneous surface. *Journal of Hydrometeorology*, **4**, 179–195.

Metzger, S., Junkermann, W., Mauder, M., Butterbach-Bahl, K., Trancón y Widemann, B., Neidl, F., Schäfer, K., Wieneke, S., Zheng, X. H., Schmid, H. P., and Foken, T. 2013: Spatially explicit regionalization of airborne flux measurements using environmental response functions, *Biogeosciences*, **10**, 2193-2217, doi:10.5194/bg-10-2193-2013.

Mudelsee, M., 2002: TAUEST: A computer program for estimating persistence in unevenly spaced weather/climate time series. *Computers& Geosciences*, **28** (1), 69–72.

Mudelsee, M., 2010: *Climate time series analysis*. Springer, 474 pp.

Nguyen, H.T., and E. A. Walker, 2000: *A First Course In Fuzzy Logic,* Chapman & Hall/CRC, ISBN 0-8493-1659-6, 373 pg.

Ogunjemiyo, S. O., Kaharabata, S. K., Schuepp, P. H., MacPherson, I. J., Desjardins, R. L., and Roberts, D.A. 2003: Methods of estimating CO2, latent heat and sensible heat fluxes from estimates of land cover fractions in the flux footprint, *Agric. For. Meteorol.*, **117**, 125-144, doi:10.1016/S0168-1923(03)00061-3.

Schneider, J., G. Grosse, D. Wagner Land cover classification of tundra environments in the Arctic Lena Delta based on Landsat 7 ETM+ data and its application for upscaling of methane emissions, *Remote Sensing of Environment*, **113**, 380-391, 2009.

Torrence, C., and G. P. Compo, 1998: A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society*, **79** (1), 61–78.

Vellinga, O. S., R. J. Dobosy, E. J. Dumas, B. Gioli, J. A. Elgers, and R. W. A. Hutjes, 2013: Calibration and quality assurance of flux observations from a small research aircraft. Journal of Atmospheric and Oceanic Technology, 30 (2), 161–181.

Arctic Regional Methane Fluxes by Ecotope as Derived Using Eddy Covariance from a ~~Low Flying~~ Low-Flying Aircraft

David S. Sayres[1], Ronald Dobosy[2,3], Claire Healy[4], Edward Dumas[2,3], John Kochendorfer[2], Jason Munster[1], Jordan Wilkerson[5], Bruce Baker[2], and James G. Anderson[1,4,5]

[1]Paulson School of Engineering and Applied Sciences, Harvard University, 12 Oxford Street, Cambridge, MA 02138
[2]Atmospheric Turbulence and Diffusion Division, NOAA/ARL, Oak Ridge, TN 37830
[3]Oak Ridge Associated Universities (ORAU), Oak Ridge, TN 37830
[4]Department of Earth and Planetary Sciences, Harvard University, 12 Oxford Street, Cambridge, MA 02138
[5]Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, MA 02138

Correspondence to: David Sayres (sayres@huarp.harvard.edu)

Abstract. The Arctic terrestrial and subsea permafrost region contains approximately 30% of the global carbon stock, and therefore understanding Arctic methane emissions and how they might change with a changing climate is important for quantifying the global methane budget and understanding its growth in the atmosphere. Here we present measurements from a new *in situ* flux observation system designed for use on a small, low-flying aircraft that was deployed over the North Slope of Alaska during August~~of~~ , 2013. The system combines a small methane instrument based on Integrated Cavity Output Spectroscopy (ICOS) with an air turbulence probe to calculate methane fluxes based on eddy covariance. We group surface fluxes by land class using a map based on LandSat Thematic Mapper (TM) ~~30 meter resolution data~~ data having 30-meter resolution. We find that wet sedge areas dominate the methane fluxes with a mean flux of 2.1 $\mu g \cdot m^{-2} \cdot s^{-1}$ during the first part of August, with methane emissions from the Sagavanirktok ~~river~~ River being the second highest at almost 1 $\mu g \cdot m^{-2} \cdot s^{-1}$. During the second half of August, after soil temperatures had cooled by 7 °C, methane emissions fell to between 0 and 0.5 $\mu g \cdot m^{-2} \cdot s^{-1}$ for all areas measured. We compare the aircraft measurements with an eddy covariance flux tower located in a wet sedge area and show that the two measurements agree quantitatively when the footprints of both overlap. However, fluxes from sedge vary at times by a factor of two or more even within a few kilometers of the tower demonstrating the importance of making regional measurements to map out methane emission spatial heterogeneity. Aircraft measurements of surface flux can play an important role in bridging the gap between ground-based measurements and regional measurements from remote sensing instruments and models.

1 Introduction

Methane is the third most important greenhouse gas after water vapor and carbon dioxide, and its concentration in the atmosphere has increased from a pre-industrial value of 0.7 parts per million by volume (ppmv) to its current value of approximately 1.85 ppmv. Methane sources are varied, with major contributors being anthropogenic (including fossil and agricultural) as well

as natural. Often multiple sources occur in the same vicinity, for example emissions from gas wells collocated with agricultural fields or pasture for grazing livestock.

In the past few years there have been increased efforts to understand how methane emissions, as well as carbon dioxide, might change from the Arctic region in response to warmer temperatures (Yvon-Durocher et al., 2014; Sturtevant et al., 2012; Sturtevant and Oechel, 2013; Walter et al., 2007b, and references therein). For example, temperatures in the Alaskan North Slope have increased 0.6 °C per decade for the last 30 years. Likewise, in that same time period the minimum extent of Arctic sea ice at the end of the summer has decreased from 8 million km^2 to 5 million km^2. ~~Late summer sea ice extent until~~ Until this past century late-summer sea ice extent was 10 million $\pm$ 1 million km^2 over the past 1500 years (Kinnard et al., 2011). Global methane concentrations have also varied during this time period, with atmospheric increases slowing down in the 1990s, leveling off in the early part of the 21st century and then increasing again since 2007 with concentrations reaching 1.8 ppmv in 2010 based on several surface based observation networks (Kirschke et al., 2013). It has been postulated that the increase could be from Arctic wetlands (Koven et al., 2011; Walter et al., 2007b).

A brief look at the carbon stock in the Arctic reveals why it has garnered so much attention. The Arctic permafrost region contains between 1330 and 1580 Pg of carbon in the tundra surface layer (0-3 meters depth), Yedoma ~~region~~ deposits, and rivers. An additional quantity is contained in deeper deposits and subsea permafrost (Tarnocai et al., 2009). Arctic carbon stock represents about a third of the total global surface carbon pool and increases to 50% when accounting for the deeper soils (Schuur et al., 2015). As the climate continues to warm, this carbon is vulnerable to thaw and decomposition by microbes, potentially leading to large increases in methane and carbon dioxide emissions. ~~Particularly important are~~ Methane from anaerobic reduction of organic carbon stocks in permafrost ~~reduced anaerobically to methane, as the latter has~~ is particularly important, having a warming potential more than twenty times ~~larger than~~ that of carbon dioxide on a 100-year time scale ~~which increases~~ and greater yet over shorter time periods (Boucher et al., 2009). The correlation between a warming Arctic and the release of methane and carbon dioxide from northern wetlands and ocean clathrates ~~in response to a warming Arctic exhibits strong evidence~~ is strongly evident in the paleoclimate record (Zachos et al., 2008; Whiticar and Schaefer, 2007). ~~It is also supported by~~ This relation is also seen 1) in current observations of methane release from thermokarst lakes formed from melting Arctic permafrost each spring and summer (Sepulveda-Jauregui et al., 2015; Walter et al., 2007b; Bastviken et al., 2004; Casper et al., 2000), ~~from~~ 2) in ebullition from deep sea sediments (Shakhova et al., 2014; Reagan et al., 2011; Damm et al., 2010), and 3) from airborne campaigns (Wofsy, 2011; Chang et al., 2014).

The North Slope of Alaska is covered by several different land classes though dominated by permafrost~~and the interior is mostly accessible only by airplane or helicopter with the exception of Prudhoe Bay which sits at the end of~~ . Access to the interior normally requires aircraft, except along the Dalton Highway (Rt. 11) ~~which extends~~ from Fairbanks to Prudhoe Bay. The lack of infrastructure, especially roads, makes continuous ~~ground based~~ ground-based measurements difficult except near the major settlements~~, which in turn makes~~ . This sparsity of data increases the uncertainty in regional bottom-up estimates of carbon flux~~uncertain owing to the sparsity of data~~. At the same time top-down estimates based on inversion modeling ~~rely on some~~ from measured concentration profiles rely on knowledge of flux sources on the ground ~~and lack the spacial resolution to discriminate~~ to determine which sources are dominating the emissions in areas like the North Slope ~~where there are a~~ with

its multitude of broad-scale emitters and point sources. ~~Airborne measurements, especially from low flying aircraft, have the potential to bridge the scale gap between process level~~ A scale gap exists between process-level studies on the ground and ~~large scale~~ large-scale regional estimates from ~~remote sensing data or inversion modeling of concentration profiles~~remotely sensed data or inversion-model results. Airborne measurements, especially from low-flying aircraft, have the potential to bridge this gap. Flux measurements from low-flying aircraft coordinated with surface measurements ~~allow for the detailed surface flux measurements to be extended to~~ promote extension of the detailed surface-flux measurements to the larger regional scale by mapping the heterogeneity in the fluxes over these larger areas.

Eddy covariance is a direct way to determine ~~the exchange~~ *in situ* the exchange (flux) of mass, momentum, and energy between the atmosphere and the surface. ~~The covariance between turbulent~~ Turbulent wind and concentration ~~is calculated directly from the individual variances from the mean and as~~ are measured at high sample rate, and their covariance yields the flux. With stationary instruments the wind and concentration measurements can be routinely obtained, and eddy covariance from ~~the ground or towers~~ fixed sites is widely represented in the literature as a way of obtaining the flux of a quantity between the surface and atmosphere. Obtaining eddy covariance measurements from a moving aircraft presents some unique challenges including accurately measuring turbulent wind velocity relative to the ground and measuring concentration at a sufficiently high data rate~~(~10 Hz for a slow moving aircraft). Even with these, in order for~~ . Furthermore, if the flux from the aircraft is to be a good proxy for a measurement taken at the surface, it needs to be sampled close to the ground~~, with the exact distance varying~~ . The appropriate distance varies depending on boundary layer height, turbulence, and the footprint size of interest. ~~Even with these difficulties, several~~ Several groups have successfully measured carbon dioxide and heat flux from low flying aircraft in the Arctic (Zulueta et al., 2011; Oechel et al., 2000, 1998; Gioli et al., 2004), Europe (Bange et al., 2007; Vellinga et al., 2010; Hutjes et al., 2010; Gioli et al., 2006), Asia (Metzger et al., 2013), and continental US (Kirby et al., 2008; LeMone et al., 2003; Avissar et al., 2009).

Here we present ~~*in situ*~~ methane fluxes taken during the summer of 2013 in the North Slope of Alaska ~~. Using two different methods for calculating flux we show that airborne flux measurements can compare with tower measurements when attention is paid to properly overlapping the footprints of the airplane and the tower. We also derive fluxes from several different land classes and compare with estimates using more traditional ground based techniques~~and use the data to explore several questions. For example, how representative are towers' footprints of other instances of the remotely determined class of land cover in which they were placed? In principle a stationary site can measure all manner of properties and state variables in the soil, the vegetation, and the air within and above the canopy. Much can be learned about the bacteria, soil chemistry, canopy storage, and other quantities relevant to the exchange of mass, momentum, and energy with the surface. But all of this is known only at the one site. How representative is that site of other locations that to remote sensors appear similar? Are there land-cover types that are particularly indicative of emission of a given trace gas? Can the class so identified be used as a quantitative predictor of a particular type of soil chemistry? This is relevant in assessing the regional methane emission from remote sensing. Methane in particular has a fairly complex chemistry in the soil involving state quantities such as the (sub-canopy) soil temperature and the height of the water table. These are measurable only in situ so that having a proxy indicator such as

vegetation cover would be valuable. Interval quantities[1] sensible remotely, such as NDVI, air temperature, and other vegetative indexes that correlate with carbon dioxide do not correlate with methane (Olefeldt et al., 2013) . Vegetation classifications determined remotely, however, have been shown in other regions to be useful for estimating regional methane emissions (eg. Schneider et al., 2009) in upscaling from ground measurements.

Aircraft, though more limited in what they can measure than fixed sites, are very mobile providing the opportunity to sample many instances of the same remotely sensed class over the landscape. From this multi-instance sample one can assess how representative the single fixed site is. One can also assess the strength of the variability within the given land-surface class for later investigation from the surface. In remote parts of the earth, in particular, a determination of near homogeneity of emission properties from multiple surfaces of recognizably similar character (class) can save considerable effort over a surface-based survey. Alternatively, large variation within a class that is not currently well predicted by some remotely measurable interval quantity will be seen as requiring additional effort for in-situ measurements to find an effective monitoring program for methane emission from that surface class.

2 Methods

To measure methane emissions over large areas of the North Slope, the Flux Observations of Carbon from an Airborne Laboratory (FOCAL) system was flown during August ~~,~~ 2013 out of Deadhorse Airport, Prudhoe Bay, AK. FOCAL, pictured in Fig. 1 flying near the NOAA Atmospheric Turbulence and Diffusion Division (NOAA/ATDD) flux tower, consisted of three main parts: the aircraft, a Diamond DA-42 from Aurora Flight Sciences, a turbulence probe, the Best Airborne Turbulence (BAT) Probe from NOAA/ATDD, and a fast methane and water instrument from the Anderson Group at Harvard University. Data presented in the results section ~~was~~ were obtained during six flights between August 13 and August 28 (Fig. 2 and Table 1). During three of these flights the aircraft made repeated passes near the NOAA/ATDD tower that was set up for comparisons. The other three flights were flown as grid patterns over large regional areas (~~~~50x50 km^2) to better sample the heterogeneity of different land types over a large region. These flights consisted of both profiles from the bottom of the boundary layer (~5-10 m) up to ~1500 m altitude and long transects (~50 km) at low altitudes (<25 m) that are used to access surface flux using eddy covariance.

2.1 FOCAL instrumentation

The airborne methane flux calculations rely on having fast ~~(10 Hz)~~ measurements of both turbulent wind velocity and dry-air mixing ratio, with the two quantities being coordinated in time and space to ~~better~~ within an error much smaller than the measurement ~~time, which in this case is 20 milliseconds~~interval. NOAA/ATDD developed the BAT probe in the 1990s as a pioneering low-cost solution for mobile atmospheric turbulence measurements (Crawford et al., 1996, 1993; Crawford and Dobosy, 1992). The BAT probe consists of a hemisphere, 15.5-cm in diameter, with nine pressure ports located at selected

[1] An interval quantity such as temperature can take an ordered range of values the length of which has meaning, as opposed to a set of categories such as surface classes having no notion of order or length.

positions on the probe head. The vertical and horizontal pairs of ports measure the differential pressure between them to calculate the angle of attack and side slip, respectively. Static pressure is taken ~~as~~ from the average of the pressures measured at the four diagonal pressure ports ~~correcting for the~~ corrected for nonzero attack and sideslip angles. Dynamic pressure is measured ~~as~~ from the difference between the pressure measured at the center hole and the static pressure, again ~~correcting for small errors as~~ adjusted for nonzero values of the angles of attack and sideslip~~are not truly zero~~. These pressure measurements are combined with a known model for flow over a hemisphere to determine three dimensional wind direction and speed relative to the probe. In order to get the velocity of the probe relative ~~the ground~~ to the ground, a GPS/INS system located near the center of gravity (CG) of the aircraft~~and two~~ , accelerometers located in the probe, and two additional GPS antennas, one ~~located~~ on the BAT probe and the other ~~located~~ on top of the main cabin, provide for calculation of the movement of the BAT probe relative to the aircraft and the aircraft relative to the Earth's surface (Crawford and Dobosy, 1997, 1992).

Fluxes of trace gases are covariances between turbulent winds and fluctuations in gas concentration. The BAT probe was designed to accurately measure turbulent winds from a moving aircraft and, using accelerometers and GPS/INS, relate those winds to the surface. The BAT probe digitizes samples at 1600~~Hz~~ s^{-1} for low-pass filtering and subsampling at 50~~Hz~~ s^{-1} to suppress aliasing. The wind measurements are synchronized with the 50~~Hz~~ s^{-1} signal from the GPS/INS system. In calculating fluxes, the 10~~Hz~~ s^{-1} data from the spectrometers discussed below are interpolated to 50~~Hz~~ s^{-1} and synchronized with the data from the BAT probe.

Before assembling the FOCAL system, the BAT probe was characterized in a wind tunnel (Dobosy et al., 2013) . A similar BAT probe was also tested in flight on a different aircraft (Vellinga et al., 2013, hereafter V2013) . After the FOCAL system was assembled, similar calibration maneuvers were flown in preparation for and during the Alaska campaign. As part of a calibration flight on the evening of August 27 in Alaska, we performed the yaw maneuver described by V2013 and obtained a residual contamination less than 10%, as described there. A pitch maneuver described by V2013 was performed resulting in contamination of 10% for the high-frequency pitching (1.6 s period), which was the best executed of the pitch test's three parts and is the severest test.

Plots of spectra and cospectra of data streams of vertical air motion and trace gases' dry-air mixing ratios were prepared and are presented in Dobosy et al. (2017) . High-frequency spectral corrections are not used for the data presented here because the highest wavenumber measured for vertical wind is above the wavenumber of the maximum spectral density and is in the inertial subrange, i.e., the spectral density follows the -5/3 power of the wavenumber. A data-starvation test using the flux runs from the evening of August 25 yielded an estimated loss of about 10% in fluxes computed at one-third the sample rate.

To evaluate the dependence of the measured methane flux on the height above the ground, a regression of 3-km running flux (see Sec. 2.3.1) against the height above ground for flight 13.09:30 was run to assess the correlation of flux with altitudes ranging from 5 to 45 m. A quadratic regression was required yielding significant positive slope but significant negative curvature. The regression line reached a maximum at an intermediate point before the maximum height above ground. Furthermore, the regression explained only 10% of the variance.

The methane instrument draws air from an inlet located 8 cm aft of the BAT probe turbulence measurements. Flow of air through the axis is controlled by a dry scroll pump located in the back of the aircraft. Air from the inlet passes through 1.25

cm diameter tubes into the nose and forward luggage bay sections of the aircraft. The pressure of the air is controlled by a proportional solenoid valve and a pressure control board that uses pressure measured at the detection axis to feed back on the valve orifice position. The actual detection axis is located in the ~~port side~~ port-side forward luggage bay. The methane instrument uses Integrated Cavity Output Spectroscopy (ICOS) to measure CH_4, H_2O and N_2O (Witinski et al., 2011). The ICOS instrument uses a ~~high finesse~~ high-finesse optical cavity composed of two high-reflectivity mirrors (R = 0.9996) to trap laser light for a period on the order of 2 µs producing effective path lengths 10^3 times the mirror separation. For the fast methane sensor used in this deployment a small ICOS cell (25 cm in length; mirrors 5 cm ~~diameter mirrors~~in diameter) was built that combines the sensitivity and stability of ICOS with a small sample volume to attain high flush rates (17~~Hz)~~ s^{-1}), which permits a sample rate of 10 s^{-1}. Using the wavelength region around 1292 cm^{-1} (7.74 µm), measurements of methane achieved a precision of 7 ppbv ($1-\sigma$, $1-\mathrm{s}$). Due to the high variability of water in the troposphere, water vapor measurements are required with any trace gas measurements in order to quantify dilution effects caused by changes in water vapor content as well as changes to spectroscopic line broadening (Webb et al., 1980; Gu et al., 2012) . Well defined absorption features of water vapor and its isotopologues as well as nitrous oxide are obtained in the same sweep of the laser, therefore the same instrument provides simultaneous measurements of nitrous oxide and water vapor along with methane. This technique ~~provided extremely high signal to noise observations~~ provides an extremely high signal-to-noise ratio as well as a robust measurement in flight and has been the basis for several ICOS flight instruments built by this group (Witinski et al., 2011; Sayres et al., 2009; Engel et al., 2006; Paul et al., 2001). Periodic calibration in flight using calibrated gas cylinders ~~were used to track~~ tracks the drift over the course of the flight and from flight to flight.

There were several other small instruments that augmented FOCAL's capabilities: a radar altimeter, for height above ground which is essential for accurate footprint calculations, and a ~~visible~~ visible-light camera, which provided a visual record of the terrain directly under the aircraft ~~used~~ to check the accuracy of the remotely sensed products ~~that are~~ used for primary landscape classification.

Aurora Flight Sciences' version of the DA-42, named the Centaur, is a twin-engine aircraft ~~, and has~~ having several characteristics that make it an ideal platform for the work discussed here. ~~Due to the twin engine configuration ,~~ The Centaur's twin-engine configuration leaves the entire center fuselage ~~is~~ available for instrumentation and sampling. The ~~Centaur~~ aircraft is electrically and structurally well-adapted for carrying a sophisticated scientific payload, having ample spare power from its two alternators and ideally located hard points for the probe and the spectroscopic equipment. ~~Finally, once fixed costs (e.g. aircraft access, instrument integration and certification) have been accounted for, the operating cost of the Centaur are just 1500 per day and 600 per flight hour - a substantial savings compared to many other scientific platforms.~~

2.2 Turbulence measurements

Eddy covariance is a direct way to determine the exchange of mass (e.g., trace gases), momentum, and energy between the atmosphere and the surface. ~~Since the observed flux~~ In principle for a gas, the covariance between the turbulent fluctuating gas concentration and the turbulent vertical wind component determines the flux. Since the flux thus obtained is assumed to represent the exchange at the surface, the airplane is flown as low as is safely possible, typically below 30 m (Mahrt, 1998).

Flux measurements from fixed surface sites, important complements to the airborne measurements, provide extended temporal coverage at selected locations as well as validation of the airborne flux measurements. ~~The covariance of the~~

The mass flux of a minor gas constituent in air, such as methane, is calculated following Webb et al. (1980); Gu et al. (2012) . Let ρ_a be the the partial density of air apart from water vapor and w be the vertical wind velocity. Then $\rho_a w$ is the dry-air ~~mixing ratio of these gases (Webb et al., 1980; Gu et al., 2012) with the turbulent vertical wind component determines the flux as shown in Eq. (2) where the bar represents the average which defines~~ mass flux, which is expanded into base state and turbulent departure with the base state ~~, ρ_a is the dry air density, w' is the departure of the wind from the base state, and c' is the analogous departure of the methane mixing ratio of dry air.~~ represented by an overbar and the departure by a prime:

$$\rho_a w = \overline{\rho_a w} + (\rho_a w)'. \tag{1}$$

Since dry air is not exchanged with the surface, $\overline{\rho_a w} = 0$. The flux of a gas is then the covariance of the turbulent dry-air mixing ratio c' with the turbulent dry-air mass flux $(\rho_a w)'$:

$$F = \overline{(\rho_a w)' c'}. \tag{2}$$

Unlike from a stationary tower, measuring the turbulent vertical wind component from an airplane requires finding the small (vector) sum of the airspeed and the ground speed, two large, nearly canceling vectors. Since both vectors fluctuate rapidly and independently, many independent measurements must be made with precise synchrony at high accuracy and sample rate. Since turbulent fluctuations can be less than 0.1~~m/s~~ $\mathrm{m \cdot s^{-1}}$, the two large velocities must each be accurate within 0.1~~m/s~~ $\mathrm{m \cdot s^{-1}}$. Four samples define the minimum effectively resolvable turbulent eddy size, about 5 m at 50 samples per second and 60~~m/s.~~

~~The Centaur uses a small Inertial Navigation System integrated with a GPS (GPS/INS) to report its ground-speed vector over the surface as well as its roll, pitch, and heading, all at 20 Hz. The low-frequency component of the Centaur's velocity is filtered to 1 Hz and extrapolated to the probe's location to mix with the high-frequency component measured directly at the probe. The airspeed vector (airflow relative to the probe) is determined from the distribution of induced pressure over the BAT probe's hemispherical surface. One static pressure and three differential pressures are taken over nine ports. From these four pressure measurements plus temperature come the five relative-flow parameters: ambient pressure, ambient temperature, and three components of the airflow relative to the probe. The dominant airflow component is along the airplane's longitudinal axis, approximately equal to the airplane's true airspeed~~ $\mathrm{m \cdot s^{-1}}$.

2.3 Methane Flux Measurements

2.3.1 Running Flux Method

The running flux method (RFM) ~~(e.g. LeMone et al., 2003)~~ is commonly used to analyze airborne fluxes (e.g. LeMone et al., 2003) . The RFM calculates the mean flux over a contiguous integration length (e.g., 3 km). As opposed to a stationary tower, which averages in time, the aircraft is moving over the landscape, so that fluxes are more appropriately averages over distance. Here

we use the same notation as Crawford et al. (1993)

$$F = \frac{\sum_{k=1}^{N} (\rho_d w)'_k c'_k V_k}{\sum_{k=1}^{N} V_k} \tag{3}$$

where ρ_a, w', and c' are defined as in Eq. (2) and V is the airspeed of the aircraft. The sum is over N consecutive samples and the denominator is the ~~spacial~~ spatial averaging length. For the analysis presented here we use a 3 km window that is moved by 1 km increments so that, unlike the normal practice with tower data, there is overlap between adjacent calculated fluxes to provide ~~smoothed interpolation~~somewhat finer spatial localization. The RFM quantitatively describes the relation between measured flux and underlying surface features of scales comparable to the averaging length or larger. This method works well as shown by LeMone et al. (2003) who found a 4 km moving average on the US Great Plains to be an appropriate compromise between uncertainty in flux estimation and resolution of landscape-scale heterogeneity. In the Arctic in 2013, the much smaller mixed layer depth gave rise to smaller turbulence scales. Ogive analysis of the frequency distribution showed 3 km to suffice as the integration distance (Berger et al., 2001). However, heterogeneity in the resulting flux estimates was large~~and repeated~~ . Repeated flight segments gave variable results likely due to changes in winds and sampling footprints and to the integration lengths being longer than the scale of the underlying surface features~~being smaller than the integration length~~. Nevertheless, there was good agreement between methane fluxes calculated by the RFM using 3 km integration centered near the tower location and fluxes computed directly from the tower measurements (see sec. 3.1). ~~However, given the small scale heterogeneity of surface features on~~ Using the RFM over the small-scale heterogeneity of the North Slope~~using the RMF~~ 's surface features, however, limits the ability to ~~separate out flux contribution~~ isolate the flux contributions from individual surface ~~types~~classes.

2.3.2 Flux Fragment Method

The ~~flux fragment method~~ Flux Fragment Method (FFM) ~~(Kirby et al., 2008)~~ was conceived to assess the homogeneity in properties of a remotely determined land class over multiple instances occurring in patches on the landscape. Often such patches are too small for a traditional RFM (Kirby et al., 2008) . The FFM uses a conditional sampling scheme whereby ~~flux~~ the flux, of methane for example, is compiled from many τ-second 'fragments' ~~of flux of a quantity, such as methane ,~~ f_i of methane flux along a transect~~given by~~ , each given by

$$f_i = \delta t \sum_{k=1}^{n\tau} [(\rho_d w)'_k c'_k V_k]_i \tag{4}$$

$$L_i = \delta t \sum_{k=1}^{n\tau} [V_k]_i. \tag{5}$$

~~where~~ Here n is the number of samples per second, ~~δt~~ δt is the sample interval, and everything else is defined as in Eq. (3) except that instead of summing over a large distance, such as 3 km, the sum is only over a few samples. Note, however, that the departure quantities used to form the fragments are relative to the same base state as in Eq. (3), a base-state of 3-km scale or more, determined by ogive analysis (Foken, 2008) to be an upper limit for the turbulence present at the time of measurement.

The fragments therefore contain information on all scales from the Nyquist wavelength of the sample rate up to the 3-km scale of the spectral gap determined from the ogive analysis. Yet, the air packets quantified by the fragments are also short enough to have likely interacted with a single class of surface. In the case of the data presented here the fragments are ~~1 second~~ 1-s sums ($\tau = 1$~~second~~ s) of approximately 60 m length. The fragments, labeled f_i, do not constitute a Reynolds average individually~~meaning~~. That is, an individual fragment~~is not long enough to average over all the frequencies that are important for eddy transport. Instead, they~~, though containing all turbulent scales, is only a short grab sample. Fragments provide a meaningful flux estimate only in aggregate. ~~Fragments~~ They can be grouped, for example, by surface class, determined from footprint estimation (Fig. 3). Fluxes are calculated only for those surface-class groups whose total length is greater than 3 km. The sum over each group divided by the ~~total~~ cumulative length of all fragments in the group provides the mean flux from the associated surface class as given by

$$F_S = \frac{\sum_{i \in S} f_i}{\sum_{i \in S} L_i} \tag{6}$$

The FFM is most appropriate in a region that is heterogeneous on small scales (100 m to 3 km), but relatively homogeneous on large scales such that many instances of the surface class, or other classification used to group the fragments, are sampled during the flight (See Kirby et al. (2008) for the full description of the method). Initial assessments of the data presented here indicate that the FFM is well suited for application to the North Slope, where Arctic tundra is interspersed with thermokarst lakes, bogs, fens and bare ground. First, land-cover data ~~is~~ are classified using a current land-cover image at 100 m resolution or better (e.g. LandSat). We use this to establish transects at altitudes typically 10 m to 30 m above ground; low as safely possible. These are flown repeatedly and coordinated with eddy-covariance towers for validation and temporal continuity. The base state is then defined, representing in principle the deterministic (non-turbulent) mesoscale component of the flow. Flux fragments are calculated using 1 s sums of squares and cross products of departures from the base state. Finally, a footprint model is applied to estimate the level of influence of each surface type on each fragment~~which allows the fragments to be grouped by type having a predetermined probability of coming from a single surface type (see~~. See sec. 3.2 for examples of how FFM is used to interpret these data~~).~~.

For the questions to be addressed in this paper the footprint model provides a measure of a fragment's membership in the fuzzy set (Nguyen and Walker, 2000) associated with each surface type, treated as a categorical variable. Fragments having a sufficient level of membership for a particular surface class are assigned to that class. A membership level above 0.5 restricts all fragments to a maximum of one class. Fragments can thus be grouped into sets in which all members have a measure greater than the prespecified level that they came from the same surface type.

We use the parameterization scheme described in Kljun et al. (2004) which uses a backward Lagrangian model (Kljun et al., 2002) for a range of heights, stability measures and other turbulence quantities that are measured from the aircraft. The required turbulence quantities are computed from averages taken over the length of each flight leg, where the flight leg is defined as the straight segment, between turns, over which the collected data are used. The more recent two-dimensional version (Kljun et al., 2015) was not considered necessary because of the footprint's current restricted use as a membership criterion to assign a selected subset of fragments to the surface categories. The flux estimate for each land surface type is

the sum of the fragments in the associated group divided by their accumulated length. The number of fragments necessary to provide a robust result can be determined by bootstrap resampling (Kirby et al., 2008). For the data presented here 3 km or ∼50 fragments suffice.

The questions to be answered by the FFM, using a fuzzy-logic approach (Nguyen and Walker, 2000) to assign surface classes to fragments and then to conditionally sample them based on those classes include:

1. What is the mean flux over all measured instances of each surface class?
2. What surface classes dominate the methane emission, and by how much?
3. How much does the flux over each class vary? Is there a spatial pattern to the variation? The variability will come both from the prevailing atmospheric environment and the heterogeneity of the emission within the same class.
4. How representative is a particular instance of all similar instances over the landscape?

2.4 Land Surface Classification

The land surface on the North Slope can be divided into different classes based on dominant plant species, topography, soil content, and soil moisture. The North Slope Science Initiative (NSSI) has identified 24 classes ~~based on~~ using Landsat Thematic Mapper (TM) 30 meter resolution land cover maps in conjunction with field surveys (Initiative, 2013). These classifications ~~, assigned based on remotely sensed data, are good~~ are plausible proxies for properties that have been shown to be primary drivers of methane production and emission~~such as~~ , including water table height, soil ~~content~~temperature, and emission pathways such as sedge roots. The areas ~~flown over by FOCAL were dominated by a mixture~~ sampled by FOCAL (Fig. 2) were covered by patches of wet sedge, mesic sedge - dwarf shrub, fresh water marsh, tussock tundra, and open water~~consisting of both lakes and rivers (Fig. 2)~~. Open water is visible from the air, and includes lakes of various sizes and origin ~~and riversbut~~ along with rivers. Coastal waters, however, are excluded for this analysis~~excludes coastal waters~~. By definition in the tussock-tundra land class, shrubs more than 20 cm tall occupy less than 25% of the surface, and tussocks occupy more than 35%. The sites are cold, poorly drained and underlain by ~~mesic to wet, silty to sandy~~ moderately moist (mesic) to wet mineral soils with silty to sandy texture and a shallow surface organic layer surrounding the tussocks. Wet sedge sites are defined as those with sedge species accounting for more than 25% of the cover and Sphagnum for less than 25%. Soils range from acidic to non-acidic, are saturated during the summer, and typically have an organic layer over silt or sand. Mesic sedge - dwarf shrub has shrubs less than 25 cm tall covering more than 25% of the area, and sedge cover is also more than 25%. Soil surface is generally mesic, but sometimes wet and is calcareous to acidic. The fresh water marshes (FWM) are semi-permanently flooded, but some have seasonal flooding, and the water depth typically exceeds 10 cm. Soils are muck or mineral, and the water can be nutrient-rich.

We use land types ~~, as~~ defined by a remote measurement, as opposed to soil properties such as moisture, organic carbon content, temperature, etc. ~~as~~ because the remotely based definition is ~~most~~ more appropriate to comparing to larger regional scale models and satellites. ~~It should be noted therefore that~~ Thus the land type here is usually a proxy for general classifications of areas with different soil moisture and other properties which are likely the primary drivers of differences in methane emis-

sions~~, though certain~~ . Certain plants such as sedge, however, have been shown to act as conduits ~~for direct~~ directly facilitating methane release from the soil to the atmosphere through the plants' vascular system (Olefeldt et al., 2013).

2.5 Tower Measurements

Starting a few weeks before the flight campaign and throughout the month of August, a small portable flux tower was ~~setup~~ installed at 70.08545° North latitude, 148.57016° West longitude, just south of Prudhoe Bay off the Dalton Highway. During that time the tower recorded CO_2 flux, CH_4 flux, latent heat flux, sensible heat flux, air temperature, and incoming radiation. Soil temperature probes were ~~also~~ used to record soil temperature at 2-cm, 5-cm, 10-cm, and 20-cm depth at three different locations around the tower. The tower was situated in an area dominated by sedge grass, and the surrounding area's water table was frequently near the surface such that the surroundings were puddled and muddy, especially in late August 2013. On the NSSI map the area is labeled as wet sedge. ~~Since low light and convective conditions~~ Low light and limited convective mixing are common on the North Slope of Alaska, ~~data collected during still conditions~~ and data collected in very weak wind do not provide reliable ~~eddy covariance~~ eddy-covariance flux measurements. Consequently~~data where~~ , data were removed from the final set when the standard deviation of the vertical wind speed was less than 0.1~~were removed from the final data set~~ $\mathrm{m \cdot s^{-1}}$.

3 Results and discussion

3.1 Comparison between aircraft and tower fluxes

On ~~three separate flights FOCAL flew along a transect near the flux tower affording the opportunity for~~ August 13, 25, and 27 the FOCAL aircraft flew repeated passes over a constant northeast/southwest track near the tower affording direct comparison between eddy-covariance methane flux measured ~~by an eddy covariance tower and by eddy covariance from a moving aircraft. It also allowed for comparisons of methane flux from the ecotope of the tower , in this case wet sedge, with methane flux from sedge averaged over the 50 km transect of the aircraft to look at spatial heterogeneity in the flux from this land type. We present these comparisons with varying levels of footprint overlap and show quantitatively how proper overlap affects the agreement between the tower and aircraft data. On August 13, 25, and 27 the aircraft flew multiple transects in a nominally East-West direction near the tower~~ from the tower and from the moving aircraft in both RFM and FFM modes (Fig. 2). ~~Based~~ The flight track was displaced north or south depending on the forecast wind direction ~~, the tracks were slightly south of the tower~~ so that the aircraft footprint could pass over the tower footprint. For the northerly winds on August 13 and 25~~and north~~ , the flight track was displaced south of the tower~~on August 27. Figure 4 shows 30-minute mean methane fluxes as measured by the flux towerfrom August 12 to August 29. Fluxes~~ . For the easterly winds of August 27 the track passed north of the tower.

Two factors, diurnal and seasonal, influenced the fluxes at the tower site ~~roughly correlate with soil temperature with the first half of Augustshowing most 30-minute mean methane fluxes ranging from 1 to 2.5 , when~~ (Fig. 4). The flight 13.09:30 on August 13 (DOY-225) was in the daytime earlier in August, when the turbulence was stronger and the soil temperatures at ~~a depth of 10 cm~~ 10-cm depth were 10-14 °C~~, and the second half of August showing most~~ . The 30-minute-mean methane

fluxes at the tower ranged from 1 to 2.5 $\mu g \cdot m^{-2} \cdot s^{-1}$. The flights 25.18:00 and 27.19:00 on August 25 and 27 (DOY-237 and -239) were in the evening and later in August with weaker turbulence and lower soil temperatures of 3-6 ^{o}C at 10 cm depth. Most 30-minute mean methane fluxes ~~ranging~~ ranged from 0.5 to 1.3 $\mu g \cdot m^{-2} \cdot s^{-1}$~~, when soil temperatures at a depth of 10 cm were 3-6~~. The observed ~~temperature dependence~~ variation with soil temperature is consistent with previous studies (e.g. Yvon-Durocher et al., 2014). Aircraft methane fluxes ~~for comparison~~ were compared with the tower ~~were calculated by finding the closest point along each leg of the flight to the flux tower and taking a 3-km RFM centered at that point. The mean of the RFM fluxes for each flight are plotted over the 30-minute fluxes from the towerin Fig. 4 (orange circles). Also shown are the fluxes associated with four different land types sampled during the flights using the FFM to calculate flux. The 95confidence intervals, using bootstrapping, for fluxes derived using RFM and FFM are shown in the three insets.~~ in two modes: as local RFM, the mean over all transects of a flight of the 3-km flux blocks downwind of and centered nearest to the tower, and as FFM, the mean of the fragments from wet sedge gathered from the whole 50-km transect and the whole flight.

Agreement between the aircraft and tower ~~on August 13 and 25 are~~ by local RFM (orange circle), near the tower but not differentiated by surface class, is within the confidence intervals of the data ~~, though the aircraft measures significantly less methane flux~~ from 13.09:30 and 25.18:00. For 27.19:00 the aircraft measured significantly lower methane flux by local RFM than the tower~~on August 27. On August 13 the methane flux~~. By FFM from wet sedge (red line), ~~as derived using FFM,~~ the same surface class as the tower but not local to it, the methane flux from 13.09:30 agrees very well with the magnitude of the flux measured ~~by the tower, which sits on a sedge site~~on August 13 at the tower. However, ~~on August 25 the flux from sedge is~~ for 25.18:00 the FFM flux from wet sedge is significantly lower than the ~~tower measurements and on August 27 is also lower, though~~ August 25 tower measurement. It is likewise for 27.19:00, though the FFM flux over wet sedge is closer to the ~~tower flux than~~ corresponding tower flux on August 27 than is the flux calculated by ~~RFM near the tower~~the local RFM.

The ~~differences between the flux measured by the tower and aircraft can be explained by looking at the aircraft footprints for the three flights. Figure 5 shows footprints for selected fragments from each day. On both August 13 and 25 the dominant wind direction was from the North, and the parts of the footprint with the highest probability of influence on the flux fragments (maroon points in Fig.5) are over sedge. On August~~ results from the three near-tower flights represent three different situations. On August 27 ~~the aircraft flew North~~ (flight 27.19:00), the footprint of the airborne measurement (Fig. 5) differed from that of the tower~~as the winds were primarily from the East. While the aircraft footprints did overlap the tower footprint ,~~. On flights 13.09:30 and 25.18:00 the footprint analyses (red to maroon contours in Fig. 5) indicate the highest probability of influence ~~for the fragments (>80) was from lakes North-West of the tower~~on the RFM flux (3-km length centered nearest the tower) are over sedge. On 27.19:00 lakes make up more than half of the RFM flux footprint. Lakes have been shown to be sporadic hot spots of methane ebullition, but at least at the time of flight these lakes showed very low methane ~~emissions. It is clear that even with flight plans that take into account forecast wind direction, careful attention to footprints must be taken when comparing different measurements. Most importantly though, is that when there is reasonable overlap between tower and aircraft footprints, the~~ emission. On August 27, the sedge, which makes up more than twice as much of the transect as the lakes, is visible to the FFM, but not to the local RFM applied here. Also, the turbulence on August 27.19:00 was weak, with $\sigma_w \sim 0.15\ m \cdot s^{-1}$. This is a case where some signal may have been lost due to insufficient sample rate for the altitude, or

perhaps because the measurement was made above the shallow layer of "constant" flux. This is a tradeoff that plagues evening and morning flights. Notable about flight 27.19:00 is its demonstration of the need for, and difficulty of obtaining, matching footprints when comparing flux measurements from ~~the aircraft agree with the tower measurements adding another level of validation to the aircraft data~~different instruments.

~~Though the tower site was situated on wet sedge, the agreement between averaging sedge sites along the approximately 50 km transect of the aircraft and just looking at the sedge in proximity to the tower varied considerably throughout the course of the mission.During the period of consistent warmth in the beginning part of August, methane emissions from sedge averaged over the whole domain of~~ On August 25, the local RFM produced a good match with the tower, in contrast to the (distributed) FFM. Plotting the entire set of RFM fluxes from 25.18:00 yielded a surprise (Fig. 6), where the tower appears to be in a local hot spot. Plots of methane flux against the height of airborne measurement and the strength of turbulence (σ_w) suggested no simple dependence on these. This flight dramatically shows the hazards inherent in relying on point measurements, which are potentially in nonrepresentative locations, to estimate the area-wide flux. Also note in the middle lower panel of Fig. 4 that the flux of 1 $\mu g \cdot m^{-2} \cdot s^{-1}$ at the tower, though isolated in space, was not isolated in time.

On August 13 everything matched. For flight 13.09:30 the wind was light and the mixing strong ($\sigma_w \sim 0.45\,m \cdot s^{-1}$). The warm soil produced a strong methane flux, and the methane flux measured at the tower matches the local RFM flux near the tower as well as the FFM flux from the distributed patches of wet sedge. Importantly, both the summer daytime (13.09:30) and autumn evening (25.18:00) flights showed that when there is reasonable overlap between the tower and aircraft footprints, the flux measurements from the ~~aircraft flight track agreed with the tower measurements (Fig. 4, inset (a)) . During the latter half of August, as both the soil and air temperatures cooled, likely due to decreased insolation, the tower consistently measured almost twice as much emission as the aircraft (Fig.4, insets (b) and (c)). Using the RFM to calculate 3 km mean fluxes along the flight transect, which is dominated by sedge,~~ aircraft agree with those from the ~~tower location stands out as a hot spot for methane emissions as shown in Fig. **??**. Emissions from the tower region are twice as large as those from the rest of the flight track consistent with the tower measuring almost twice as much as sedge. Though not shown, latent heat flux showed a similar pattern to that of methane. While we have soil temperature and moisture measurements only near the tower, it is possible that the tower was in a locally wet or warm spot leading to larger methane emissions than westward of the tower. This underscores the importance of coordinating local measurements, such as with a tower , with larger scale measurements from an aircraft that can show the regional heterogeneity of methane fluxes associated with land classes that from remote measurements look the same~~tower adding another level of validation to the aircraft data.

3.2 Regional methane fluxes

During August ~~,~~2013 FOCAL measured methane flux from a variety of ecotopes across the North Slope. There are six flights used in this analysis; four in the day time and two in the evening (1800 - 1900 local time) which were covered individually in the last section. Keeping that discussion in mind, these data are comparable as a set. Based on the tower data, which exhibit strong and regular diurnal cycles of carbon dioxide and latent heat (not shown), methane has a generally weak diurnal cycle. The sharp feature in the tower trace on August 13 (DOY 225) very likely has a diurnal component, but its shape suggests more

than just solar input. This discussion, therefore, will focus the seasonal change and the methane-emission characteristics of the various surface classes (Figs. 4 and 7).

In order to distinguish the contribution to the total methane flux from individual land types and to assess the variability across ecotopes, the data are filtered to only include flux fragments where at least 85% of the crosswind-integrated probability density comes from a single land class. Increasing this ~~metric, increases the causal relationship~~ threshold increases the link between the calculated flux and a single land class, but reduces the number of footprints available for the analysis thus loosening the confidence interval. Varying ~~this filter~~ the threshold between 80% and 95% produces only a small effect on the quantification of flux from each land class. We find that 85% is a good compromise between singling out individual land classes while still retaining a ~~large~~ sufficient dataset. For the flight speed of the Centaur at low altitude and wind conditions during the flights, the footprint length for each 60-m fragment varied between 0.5 and 5 km, though the part of the footprint whose probability ~~density~~ of contributing more than 90% of the flux was ~~only~~ between 100 and 800 m long. The above filter eliminates about half of the flux fragments from each flight. Of those, we limit the land classes to those where the total number of flux fragments is more than 50 fragments or an equivalent distance of 3 km. The flux fragments are summed and then divided by the total integration length for each land cover type (Fig. 7).

Land cover type varies over the North Slope, so different flights sampled different types of ~~terrain~~ land cover (see Table 1 and Fig. ~~2), with wet sedge being the most dominant and thus~~ 2). Wet sedge was the most prevalent and thus was sampled on each flight~~.~~, except for the morning flight of August 28. Other land cover classes such as bare ground, dwarf shrub, and low-tall willow were also observed but ~~not with sufficient frequency~~ in insufficient quantity to calculate a statistically significant flux. Prevalent near the tower site, which was sampled on August 13, 25, 27, were wet sedge, mesic ~~sedge-dwarf~~ sedge - dwarf shrub, some lakes, the Sagavanirktok (Sag) ~~river~~River, and fresh water marsh. Soil temperatures in mid-August varied by 1.5 °C with a mean soil temperature of 8 °C at 5 cm depth. By the end of August soil temperatures had dropped to a mean of 3 °C. ~~Sedge~~ Wet sedge showed the strongest correlation with soil temperature, with fluxes falling from 2.1 $\mu g \cdot m^{-2} \cdot s^{-1}$ on August 13th to less than 0.5 $\mu g \cdot m^{-2} \cdot s^{-1}$ by the end of August. This relationship held true for emissions from the Sag ~~river~~ River with emissions falling from 1 $\mu g \cdot m^{-2} \cdot s^{-1}$ to near 0. Wet sedge, followed by the Sag River, had the largest observed flux of any of the land classes sampled during the first half of August. The other land classes have smaller, more variable fluxes on most flights so that surface class alone does not distinguish them. Most likely the true variability, contributing to the large confidence intervals, is caused by heterogeneity within the surface class in sub-surface soil temperature and water table height. However, within that we can still derive a mean flux based on a large regional sample. Once the soil cools, wet sedge shows reduced, though still positive, flux of methane consistent with the other surface classes measured such as mesic sedge and lakes. The Sag River shows close to zero methane flux. Lakes showed no trend~~though it~~. It should be noted, however, that the number of lakes sampled on August 13 was small and the flux variable as indicated by the large 95% confidence interval. While data ~~on August 13 was sparse for~~ from the other land classes sampled ~~, during the latter half of August~~ on August 13 were sparse, emissions from fresh water marsh and tussock tundra during the latter half of August were similar to those from lakes and ~~sedge~~the two sedge classes.

Airborne measurements made during August, 2013 are consistent with findings from other studies. Olefeldt et al. (2013) reported sites dominated by sedge and wet soils having methane emissions ranging from 0.46 to 1.6 $\mu g \cdot m^{-2} \cdot s^{-1}$ with a median value of 0.87 $\mu g \cdot m^{-2} \cdot s^{-1}$ across multiple permafrost sites. Other studies at single locations fall into this same range. For example, Harazono et al. (2006) measured methane fluxes from a wet sedge site in Happy Valley, AK during August of 1995 ranging from 0.38 to 1.5 $\mu g \cdot m^{-2} \cdot s^{-1}$ and Sturtevant and Oechel (2013) measured wet sedge near Barrow with emissions of 0.39 ± 0.03 $\mu g \cdot m^{-2} \cdot s^{-1}$ with short periods of higher emissions up to 1.1 $\mu g \cdot m^{-2} \cdot s^{-1}$. Emissions from mesic-sedge sites near the Sag ~~river~~River, though south of the areas measured by FOCAL, showed fluxes of 0.35 to 0.58 $\mu g \cdot m^{-2} \cdot s^{-1}$ in the first half of August falling to 0.12 to 0.23 $\mu g \cdot m^{-2} \cdot s^{-1}$ in the second half of August (Harazono et al., 2006).

Emissions from lakes tend to be more variable than the land classes~~with measured~~. Measured emissions from individual lakes ~~ranging~~ ranged from 0.25 to 6.3 $\mu g \cdot m^{-2} \cdot s^{-1}$ across various thermokarst and other lakes in the North Slope (Walter et al., 2007a; Sepulveda-Jauregui et al., 2015). These fluxes are reported as means over a year, so emission rates during short periods of time may be lower or higher for an individual lake. While FOCAL did not sample the same lakes as in the aforementioned studies, during the flights near the tower multiple passes over the same lakes allow emissions from individual lakes to be measured. On August 13, five lakes were sampled with sufficient frequency to produce a statistically significant flux~~ranging~~. The flux for individual lakes ranged from 0 to 2.6 $\mu g \cdot m^{-2} \cdot s^{-1}$ ~~. On August 25 a single lake was measured with an emissions rate of 0.72~~ with a mean for all lakes sampled of 0.18 $\mu g \cdot m^{-2} \cdot s^{-1}$. On August 27 four lakes were measured with emissions ranging from 0.09 to 0.18 $\mu g \cdot m^{-2} \cdot s^{-1}$. The mean methane flux from lakes over the period of the flights shows little flux, except for the lakes sampled on the morning flight of August 28. These are in a different area 250 km west of the tower. Those lakes show an aggregate mean of 0.36 $\mu g \cdot m^{-2} \cdot s^{-1}$, the only flux measured from lakes that was statistically significantly positive (Fig. 7). These data are consistent with the rates measured by the above studies.

4 Conclusions

The FOCAL campaign during the summer of 2013 showed how methane fluxes could be successfully measured over large regions using airborne eddy covariance measurements from a small, low-flying aircraft. Comparing the airborne measurements to those of a tower showed that the data were quantitatively comparable when there was good overlap between the tower footprint and aircraft footprint. However, along the flight track local conditions dominated the flux especially in the transition season from summer to fall in late August. Comparing wet sedge at the tower site with wet sedge west of the tower showed a factor of two difference in methane emissions during the ~~later~~ latter half of August which underscores the importance of regional measurements as fluxes can have large dependence on spatial heterogeneity even over relatively short distances. During the middle of the summer fluxes from wet sedge were more homogeneous across the area flown.

Measurements of methane flux over the North Slope of Alaska in August showed a strong correlation with soil temperature consistent with previous studies. Wet sedge showed the highest persistent methane emissions with mean fluxes about 2 $\mu g \cdot m^{-2} \cdot s^{-1}$ in the first half of August falling to 0.2 $\mu g \cdot m^{-2} \cdot s^{-1}$ in the latter half of August. Emissions from the Sag ~~river~~ River showed a similar trend, while other land surface classes were not sampled enough during the first half of August to

provide a statistically significant sample. Individual lakes sampled near the tower showed a large range of emissions varying from near 0 to 2.6 $\mu g \cdot m^{-2} \cdot s^{-1}$ consistent with the range of lake emissions reported in the literature.

Aircraft measurements of surface flux can play an important role in bridging the gap between ground-based measurements and regional measurements based on inversion modeling or downwind-upwind differences. While airborne campaigns are generally more costly than ground based measurements, ~~using small aircraft~~ these costs can be minimized ~~and for~~ by using small aircraft. For areas that are logistically challenging to access, such as the North Slope, airborne eddy covariance presents the easiest and least expensive way to directly measure surface fluxes regionally with large coverage.

5 Data Availability

All data ~~is publically~~ are publicly archived at the NSF ACADIS website (~~http~~https://~~articdata~~arcticdata.io) under citation: David Sayres. 2014. Collaborative Research: Multi-Regional Scale Aircraft Observations of Methane and Carbon Dioxide Isotopic Fluxes in the Arctic. NSF Arctic Data Center. urn:uuid:58bddf69-74fe-4a20-958e-4cd23bb6941f.

Acknowledgements. This work was supported by NSF Grant 1203583 and data is archived at the ACADIS website. The authors wish to gratefully acknowledge the efforts and exceptional flying of our pilot, Bernard 'Bernie' Charlemagne, without whom these data could not have been collected.

References

Avissar, R., Holder, H. E., Abehserra, N., Bolch, M. A., Canning, P., Magalhaes, J., Walko, R. L., Novick, K., Katul, G., Prince, K., et al.: The Duke University Helicopter Observation Platform, Bulletin of the American Meteorological Society, 90, 939–954, 2009.

Bange, J., Spiess, T., and van den Kroonenberg, A.: Characteristics of the early-morning shallow convective boundary layer from Helipod Flights during STINHO-2, Theoretical and Applied Climatology, 90, 113–126, http://dx.doi.org/10.1007/s00704-006-0272-2, 10.1007/s00704-006-0272-2, 2007.

Bastviken, D., Cole, J., Pace, M., and Tranvik, L.: Methane emissions from lakes: Dependence of lake characteristics, two regional assessments, and a global estimate, Global Biogeochemical Cycles, 18, GB4009, doi:10.1029/2004GB002238, 2004.

Berger, B. W., Davis, K. J., Yi, C. X., Bakwin, P. S., and Zhao, C. L.: Long-term carbon dioxide fluxes from a very tall tower in a northern forest: Flux measurement methodology, Journal of Atmospheric and Oceanic Technology, 18, 529–542, doi:10.1175/1520-0426(2001)018<0529:LTCDFF>2.0.CO;2, 2001.

Boucher, O., Friedlingstein, P., Collins, B., and Shine, K. P.: The indirect global warming potential and global temperature change potential due to methane oxidation, Environmental Research Letters, 4, 044 007, doi:10.1088/1748-9326/4/4/044007, 2009.

Casper, P., Maberly, S. C., Hall, G. H., and Finlay, B. J.: Fluxes of methane and carbon dioxide from a small productive lake to the atmosphere, Biogeochemistry, 49, 1–19, doi:10.1023/A:1006269900174, 2000.

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., Bruhwiler, L. M. P., and Wofsy, S. C.: Methane emissions from Alaska in 2012 from CARVE airborne observations, Proceedings of the National Academy of Sciences of the United States of America, 111, 16 694–16 699, doi:10.1073/pnas.1412953111, 2014.

Crawford, T. L. and Dobosy, R. J.: A Sensitive Fast-response Probe To Measure Turbulence and Heat-flux From Any Airplane, Boundary-layer Meteorology, 59, 257–278, doi:10.1007/BF00119816, 1992.

Crawford, T. L. and Dobosy, R. J.: Pieces to a puzzle: Air-surface exchange and climate, GPS World, 8, 32–39, 1997.

Crawford, T. L., McMillen, R. T., Dobosy, R. J., and MacPherson, I.: Correcting airborne flux measurements for aircraft speed variation, Boundary-Layer Meteorology, 66, 237–245, 1993.

Crawford, T. L., Dobosy, R. J., and Dumas, E. J.: Aircraft wind measurement considering lift-induced upwash, Boundary-layer Meteorology, 80, 79–94, doi:10.1007/BF00119012, 1996.

Damm, E., Helmke, E., Thoms, S., Schauer, U., Nothig, E., Bakker, K., and Kiene, R. P.: Methane production in aerobic oligotrophic surface water in the central Arctic Ocean, Biogeosciences, 7, 1099–1108, 2010.

Dobosy, R., Dumas, E., Senn, D., B. Baker, a. D. S., Witinski, M., C.E. Healy, a. J. M., and Anderson, J.: Calibration and quality assurance of an airborne turbulence probe in an aeronautical wind tunnel, Journal of Atmospheric and Oceanic Technology, 30, 182–196, 2013.

Dobosy, R. J., Sayres, D. S., Healy, C., Dumas, E., Heuer, M., Kochendorfer, J., Baker, B., and Anderson, J. G.: Estimating airborne-flux uncertainty over Alaskan tundra: Update on the Flux-Fragment Method, J.Atmos. Ocean. Tech., JTECH-D-16-0187, in review, 2017.

Engel, G. S., Drisdell, W. S., Keutsch, F. N., Moyer, E. J., and Anderson, J. G.: Ultrasensitive near-infrared integrated cavity output spectroscopy technique for detection of CO at 1.57 mu m: new sensitivity limits for absorption measurements in passive optical cavities, Applied Optics, 45, 9221–9229, 2006.

Foken, T.: The energy balance closure problem - An overview, Ecol. Appl., 18, 1351–1367, 2008.

Gioli, B., Miglietta, F., De Martino, B., Hutjes, R. W. A., Dolman, H. A. J., Lindroth, A., Schumacher, M., Sanz, M. J., Manca, G., Peressotti, A., and Dumas, E. J.: Comparison between tower and aircraft-based eddy covariance fluxes in five European regions, Agricultural and Forest Meteorology, 127, 1–16, 2004.

Gioli, B., Miglietta, F., Vaccari, F. P., Zaldei, A., and De Martino, B.: The Sky Arrow ERA, an innovative airborne platform to monitor mass, momentum and energy exchange of ecosystems, Annals of Geophysics, 49, 109–116, 2006.

Gu, L., Massman, W. J., Leuning, R., Pallardy, S. G., Meyers, T., Hanson, P. J., Riggs, J. S., Hosman, K. P., and Yang, B.: The fundamental equation of eddy covariance and its application in flux measurements, Agricultural and Forest Meteorology, 152, 135–148, 2012.

Harazono, Y., Mano, M., Miyata, A., Yoshimoto, M., Zulueta, R., Vourlitis, G., Kwon, H., and Oechel, W. C.: Temporal and spatial differences of methane flux at arctic tundra in Alaska, Mem. Natl Inst. Polar Res., Spec. Issue, 59, 79–95, 2006.

Hutjes, R. W. A., Vellinga, O. S., Gioli, B., and Miglietta, F.: Dis-aggregation of airborne flux measurements using footprint analysis, Agricultural and Forest Meteorology, 150, 966–983, 2010.

Initiative, N. S. S.: North Slope Science Initiative Landcover Mapping Summary Report, Electronic, http://catalog.northslope.org/catalogs/6979-2013-nssi-landcover-for-north-slope-of-alaska, 2013.

Kinnard, C., Zdanowicz, C. M., Fisher, D. A., Isaksson, E., de Vernal, A., and Thompson, L. G.: Reconstructed changes in Arctic sea ice over the past 1,450 years, Nature, 479, 509–U231, doi:10.1038/nature10581, 2011.

Kirby, S., Dobosy, R., Williamson, D., and Dumas, E.: An aircraft-based data analysis method for discerning individual fluxes in a heterogeneous agricultural landscape, Agricultural And Forest Meteorology, 148, 481–489, 2008.

Kirschke, S., Bousquet, P., Ciais, P., Saunois, M., Canadell, J. G., Dlugokencky, E. J., Bergamaschi, P., Bergmann, D., Blake, D. R., Bruhwiler, L., Cameron-Smith, P., Castaldi, S., Chevallier, F., Feng, L., Fraser, A., Heimann, M., Hodson, E. L., Houweling, S., Josse, B., Fraser, P. J., Krummel, P. B., Lamarque, J.-F., Langenfelds, R. L., Le Quere, C., Naik, V., O'Doherty, S., Palmer, P. I., Pison, I., Plummer, D., Poulter, B., Prinn, R. G., Rigby, M., Ringeval, B., Santini, M., Schmidt, M., Shindell, D. T., Simpson, I. J., Spahni, R., Steele, L. P., Strode, S. A., Sudo, K., Szopa, S., van der Werf, G. R., Voulgarakis, A., van Weele, M., Weiss, R. F., Williams, J. E., and Zeng, G.: Three decades of global methane sources and sinks, Nature Geoscience, 6, 813–823, doi:10.1038/NGEO1955, 2013.

Kljun, N., Rotach, M., and Schmid, H.: A three-dimensional backward lagrangian footprint model for a wide range of boundary-layer stratifications, BOUNDARY-LAYER METEOROLOGY, 103, 205–226, doi:10.1023/A:1014556300021, 2002.

Kljun, N., Calanca, P., Rotach, M., and Schmid, H.: A simple parameterisation for flux footprint predictions, BOUNDARY-LAYER METEOROLOGY, 112, 503–523, doi:10.1023/B:BOUN.0000030653.71031.96, 2004.

Kljun, N., Calanca, P., Rotach, M., and Schmid, H.: A simple two-dimensional parameterization for Flux Footprint Prediction (FFP), Geoscientific Model Development, 8, Geoscientific Model Development, 2015.

Koven, C. D., Ringeval, B., Friedlingstein, P., Ciais, P., Cadule, P., Khvorostyanov, D., Krinner, G., and Tarnocai, C.: Permafrost carbon-climate feedbacks accelerate global warming, Proceedings of the National Academy of Sciences of the United States of America, 108, 14 769–14 774, doi:10.1073/pnas.1103910108, 2011.

LeMone, M. A., Grossman, R. L., Chen, F., Ikeda, K., and Yates, D.: Choosing the averaging interval for comparison of observed and modeled fluxes along aircraft transects over a heterogeneous surface, Journal of Hydrometeorology, 4, 179–195, doi:10.1175/1525-7541(2003)4<179:CTAIFC>2.0.CO;2, 2003.

Mahrt, L.: Flux sampling errors for aircraft and towers, Journal Of Atmospheric And Oceanic Technology, 15, 416–429, 1998.

Metzger, S., Junkermann, W., Mauder, M., Butterbach-Bahl, K., Trancón y Widemann, B., Neidl, F., Schäfer, K., Wieneke, S., Zheng, X., Schmid, H., et al.: Spatially explicit regionalization of airborne flux measurements using environmental response functions, Biogeosciences, 10, 2193–2217, 2013.

Nguyen, H. and Walker, E. A.: A First Course In Fuzzy Logic, Chapman & Hall/CRC, doi:ISBN 0-8493-1659-6, 2000.

Oechel, W. C., Vourlitis, G. L., Brooks, S., Crawford, T. L., and Dumas, E.: Intercomparison among chamber, tower, and aircraft net CO2 and energy fluxes measured during the Arctic System Science Land-Atmosphere-Ice Interactions (ARCSS-LAII) Flux Study, Journal of Geophysical Research: Atmospheres, 103, 28 993–29 003, 1998.

Oechel, W. C., Vourlitis, G. L., Verfaillie, J., Crawford, T., Brooks, S., Dumas, E., Hope, A., Stow, D., Boynton, B., Nosov, V., et al.: A scaling approach for quantifying the net CO2 flux of the Kuparuk River Basin, Alaska, Global Change Biology, 6, 160–173, 2000.

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, doi:10.1111/gcb.12071, 2013.

Paul, J., Lapson, L., and Anderson, J.: Ultrasensitive absorption spectroscopy with a high-finesse optical cavity and off-axis alignment, Applied Optics, 40, 4904 – 4910, 2001.

Reagan, M. T., Moridis, G. J., Elliott, S. M., and Maltrud, M.: Contribution of oceanic gas hydrate dissociation to the formation of Arctic Ocean methane plumes, Journal of Geophysical Research-oceans, 116, C09 014, doi:10.1029/2011JC007189, 2011.

Sayres, D. S., Moyer, E. J., Hanisco, T. F., Clair, J. M., Keutsch, F. N., O'Brien, A., Allen, N. T., Lapson, L., Demusz, J. N., Rivero, M., Martin, T., Greenberg, M., Tuozzolo, C., Engel, G. S., Kroll, J. H., Paul, J. B., and Anderson, J. G.: A new cavity based absorption instrument for detection of water isotopologues in the upper troposphere and lower stratosphere, Review Of Scientific Instruments, 80, 2009.

Schneider, J., Grosse, G., and Wagner, D.: Land cover classification of tundra environments in the Arctic Lena Delta based on Landsat 7 ETM+ data and its application for upscaling of methane emissions,, Remote Sensing of Environment, 113, 380–391, 2009.

Schuur, E. A. G., McGuire, A. D., Schaedel, C., Grosse, G., Harden, J. W., Hayes, D. J., Hugelius, G., Koven, C. D., Kuhry, P., Lawrence, D. M., Natali, S. M., Olefeldt, D., Romanovsky, V. E., Schaefer, K., Turetsky, M. R., Treat, C. C., and Vonk, J. E.: Climate change and the permafrost carbon feedback, Nature, 520, 171–179, doi:10.1038/nature14338, 2015.

Sepulveda-Jauregui, A., Anthony, K. M. W., Martinez-Cruz, K., Greene, S., and Thalasso, F.: Methane and carbon dioxide emissions from 40 lakes along a north-south latitudinal transect in Alaska, Biogeosciences, 12, 3197–3223, doi:10.5194/bg-12-3197-2015, 2015.

Shakhova, N., Semiletov, I., Leifer, I., Sergienko, V., Salyuk, A., Kosmach, D., Chernykh, D., Stubbs, C., Nicolsky, D., Tumskoy, V., and Gustafsson, O.: Ebullition and storm-induced methane release from the East Siberian Arctic Shelf, Nature Geoscience, 7, 64–70, doi:10.1038/NGEO2007, 2014.

Sturtevant, C. S. and Oechel, W. C.: Spatial variation in landscape-level CO2 and CH4 fluxes from arctic coastal tundra: influence from vegetation, wetness, and the thaw lake cycle, Global Change Biology, 19, 2853–2866, doi:10.1111/gcb.12247, 2013.

Sturtevant, C. S., Oechel, W. C., Zona, D., Kim, Y., and Emerson, C. E.: Soil moisture control over autumn season methane flux, Arctic Coastal Plain of Alaska, Biogeosciences, 9, 1423–1440, doi:10.5194/bg-9-1423-2012, 2012.

Tarnocai, C., Canadell, J. G., Schuur, E. A. G., Kuhry, P., Mazhitova, G., and Zimov, S.: Soil organic carbon pools in the northern circumpolar permafrost region, Global Biogeochemical Cycles, 23, GB2023, doi:10.1029/2008GB003327, 2009.

Vellinga, O. S., Gioli, B., Elbers, J. A., Holtslag, A. A. M., Kabat, P., and Hutjes, R. W. A.: Regional carbon dioxide and energy fluxes from airborne observations using flight-path segmentation based on landscape characteristics, Biogeosciences, 7, 1307–1321, http://www.biogeosciences.net/7/1307/2010/, 2010.

Figure 1. Picture of the FOCAL system flying near the NOAA/ATDD flux tower in North Slope, AK.

Vellinga, O. S., Dobosy, R. J., Dumas, E. J., Gioli, B., Elgers, J. A., and Hutjes, R. W. A.: Calibration and quality assurance of flux observations from a small research aircraft., Journal of Atmospheric and Oceanic Technology, 30, 161–181, 2013.

Walter, K. M., Edwards, M. E., Grosse, G., Zimov, S. A., and Chapin, F. S.: Thermokarst lakes as a source of atmospheric CH4 during the last deglaciation, Science, 318, 633–636, 2007a.

Walter, K. M., Smith, L. C., and Chapin, F. S.: Methane bubbling from northern lakes: present and future contributions to the global methane budget, Philosophical Transactions of the Royal Society A-mathematical Physical and Engineering Sciences, 365, 1657–1676, 2007b.

Webb, E. K., Pearman, G. I., and Leuning, R.: Correction of Flux Measurements For Density Effects Due To Heat and Water-vapor Transfer, Quarterly Journal of the Royal Meteorological Society, 106, 85–100, doi:10.1002/qj.49710644707, 1980.

Whiticar, M. and Schaefer, H.: Constraining past global tropospheric methane budgets with carbon and hydrogen isotope ratios in ice, Philosophical Transactions of the Royal Society A-mathematical Physical and Engineering Sciences, 365, 1793–1828, doi:10.1098/rsta.2007.2048, 2007.

Witinski, M. F., Sayres, D. S., and Anderson, J. G.: High precision methane isotopologue ratio measurements at ambient mixing ratios using integrated cavity output spectroscopy, Applied Physics B-lasers and Optics, 102, 375–380, doi:10.1007/s00340-010-3957-2, 2011.

Wofsy, S. C.: HIAPER Pole-to-Pole Observations (HIPPO): fine-grained, global-scale measurements of climatically important atmospheric gases and aerosols, Philosophical Transactions of the Royal Society A-mathematical Physical and Engineering Sciences, 369, 2073–2086, doi:10.1098/rsta.2010.0313, 2011.

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, doi:10.1038/nature13164, 2014.

Zachos, J. C., Dickens, G. R., and Zeebe, R. E.: An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics, Nature, 451, 279–283, doi:10.1038/nature06588, 2008.

Zulueta, R. C., Oechel, W. C., Loescher, H. W., Lawrence, W. T., and Paw U, K. T.: Aircraft-derived regional scale CO2 fluxes from vegetated drained thaw-lake basins and interstitial tundra on the Arctic Coastal Plain of Alaska, Global Change Biology, 17, 2781–2802, doi:10.1111/j.1365-2486.2011.02433.x, 2011.

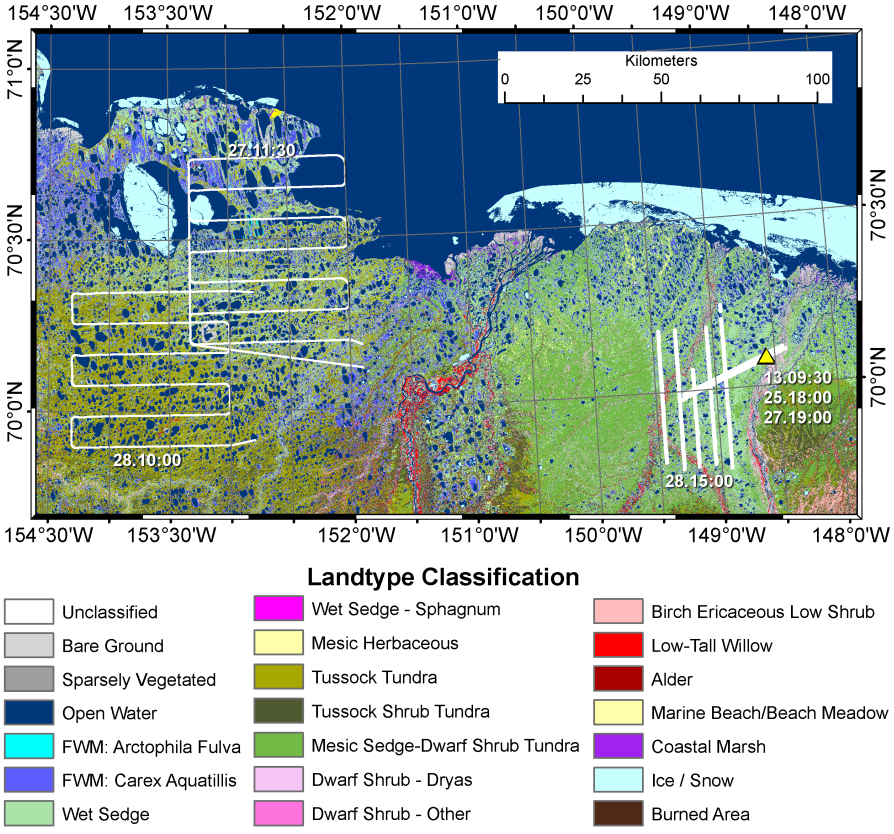


Figure 2. Six flight tracks flown by FOCAL during August ~~,~~ 2013 are shown in white. ~~Dates~~ Flights are given in the figure as DD.HH:MM, where DD is the ~~local day of the month of~~ date in August ~~,~~ and HH:MM is the time (UTC-10hr) of the middle of the flight rounded to nearest half-hour. Flight tracks ~~shown~~ are shown only for ~~portions of~~ the ~~flight below~~ portions flown within 25 m ~~above~~ of the ground. ~~Flight tracks are displayed over a false color.~~ The underlying chart gives the NSSI-assigned land cover ~~map~~ produced from LandSat ~~30 m~~ 30-m Thematic ~~mapper~~ Mapper data~~with land types assigned by the NSSI as described in the text~~. ~~Yellow~~ The yellow triangle ~~gives position of~~ locates the NOAA/ATDD flux tower.

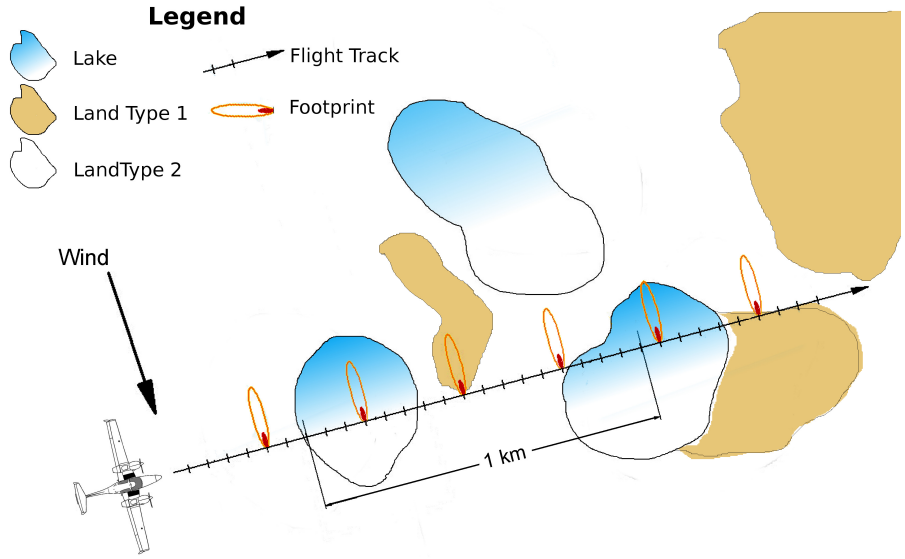


Figure 3. The Flux Fragment Method (FFM) ~~relies on dividing up~~ divides the covariance measurements into small fragments whose footprints can be attributed to different landscape features or classes. In the figure the landscape has been divided into lakes and two types of land, for example wet sedge and fresh water marsh. Footprints are calculated for each fragment and footprints that lie mostly (>85%) on a single land type are ~~labeled as~~ assigned to that land type. All footprints for a ~~single~~ given land type can then be summed and divided by the cumulative path length in air.

Table 1. Flights used in the analysis along with location, time of day, mean air temperature, and surface land classes.

Flight Date[1] DD.HH:MM	Location	Start Time UTC-10	End Time UTC-10	Temperature[2] (°C)	Dominant Land Types[3]
13.09:30	Tower	08:19	10:22	16	Sedge, Mesic Sedge, Lakes, Sag ~~river~~River, FWM
~~25.19~~25.18:00	Tower	17:43	19:49	5	Sedge, Mesic Sedge, Lakes, Sag ~~river~~River, FWM
27.11:30	Western Grid	09:40	13:00	6	Sedge, FWM, Lakes, Tussock tundra
~~27.18:30~~ 27.19:00	Tower	16:46	20:02	10	Sedge, Mesic Sedge, Lakes, Sag ~~river~~River, FWM
28.10:00	Western Grid	08:39	11:39	11	Tussock tundra, Lakes, Lake margins are FWM and Sedge
28.15:00	Eastern Grid	13:59	15:44	16	Sedge, Mesic Sedge, Lakes, Kuparuk River, FWM

[1] All flights are during August 2013. DD is the local date of the flight and HH:MM is the middle time of the flight rounded to the nearest half-hour. [2] Temperature calculated as mean temperature recorded by instrument during flight time and below 100 m. [3] Land classes are given in order of largest percent coverage first.

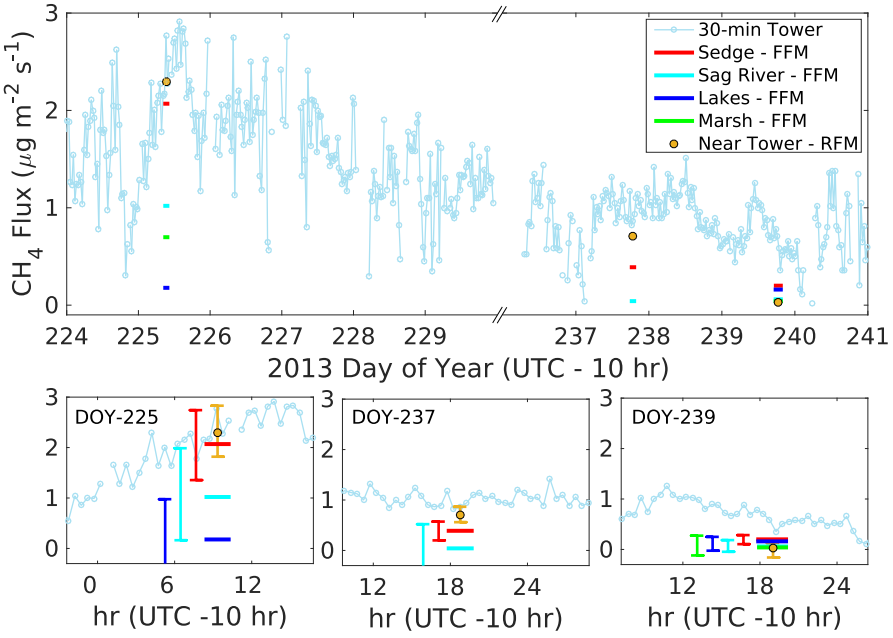


Figure 4. ~~Comparison of methane~~ Methane flux measured ~~by~~ from the flux tower compared with fluxes measured by the FOCAL system. Tower methane fluxes (top plot) are 30-minute means plotted versus day of year. ~~Three~~ The flights (~~Aug. 13~~13.09:30, ~~25~~25.18:00, 27.19:00 on DOY 225, 237, and ~~27~~239, respectively) each made repeated ~~flight transects~~ passes near the tower. ~~A running~~ The orange circle gives the mean ~~flux, using~~ over these passes of the RFM-determined 3-km flux centered nearest ~~3 km of flight track to~~ the tower~~for each leg, was calculated and the mean of these fluxes is plotted for each day as an orange circle~~. Fluxes ~~for wet sedge, marsh, lakes, and the Sag river were calculated using~~ by FFM ~~using data from~~ were aggregated by surface class over the whole flight~~and are plotted for each day, color coded according to the legend, with the~~ . The length of the line along the time axis ~~representing~~ represents the ~~time~~ period over which the data were taken, typically 1.5 hr. ~~Bottom plots~~ Lower panels show details for each flight day, labeled by day of year (DOY), with vertical bars showing the 95% confidence interval based on bootstrap analysis. Bars are offset along the x-axis for clarity.

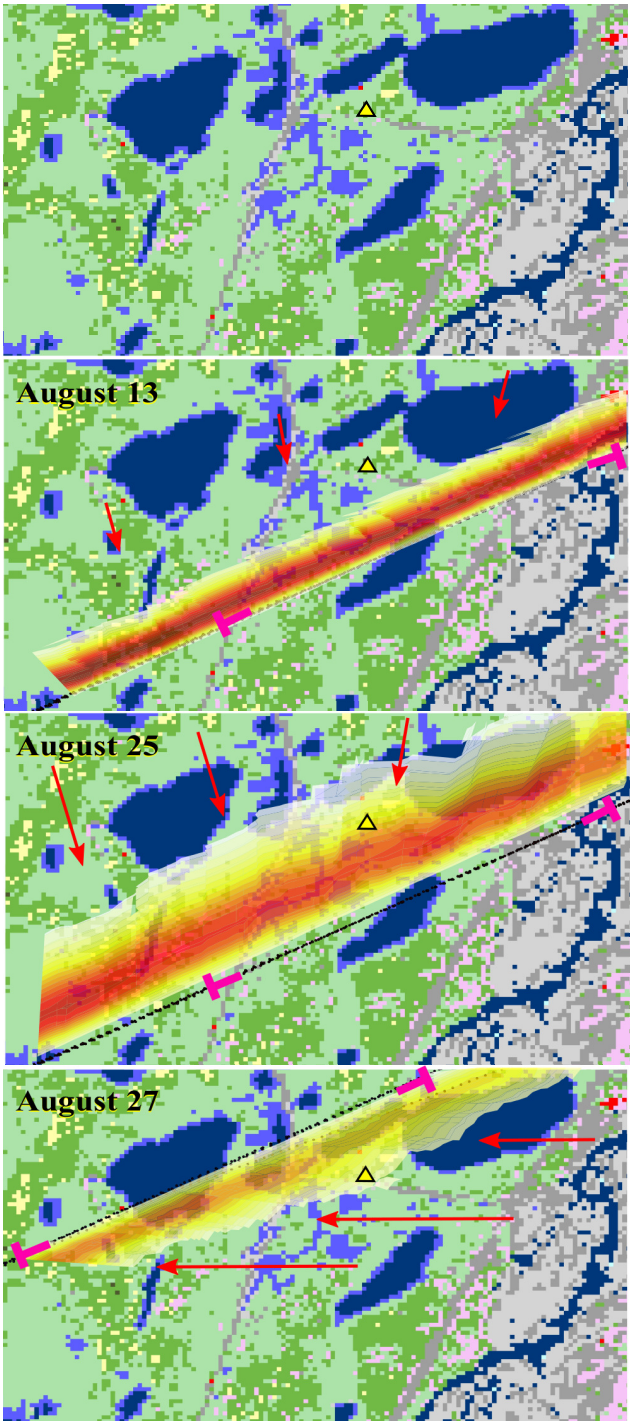


Figure 5. ~~Map of area surrounding~~ Flux footprints near the ~~flux~~ tower (yellow triangle) ~~with false color map representing different~~ for the three tower flights (13.09:30, 25.18:00, and 27.19:00). They are laid over the NSSI-classified land ~~classes defined as in~~ cover map (see Fig. 2). ~~Bottom three plots show three days when data was taken near~~ The top panel facilitates identifying the ~~tower~~surface classes under each footprint. The flight track~~for each flight~~ , always passing downwind of the tower, is shown as black points, ~~where~~ each ~~point is~~ giving the start position of a flux fragment. ~~Colored ribbon shows the flux footprints along the flight track.~~ The darker and redder ~~color of the~~ ribbon color represents ~~larger~~ greater probability of contribution to the total flux as described in the text. Red arrows indicate the mean direction of the wind. The part of the flight track used in the near tower RFM calculation is located between the magenta brackets.

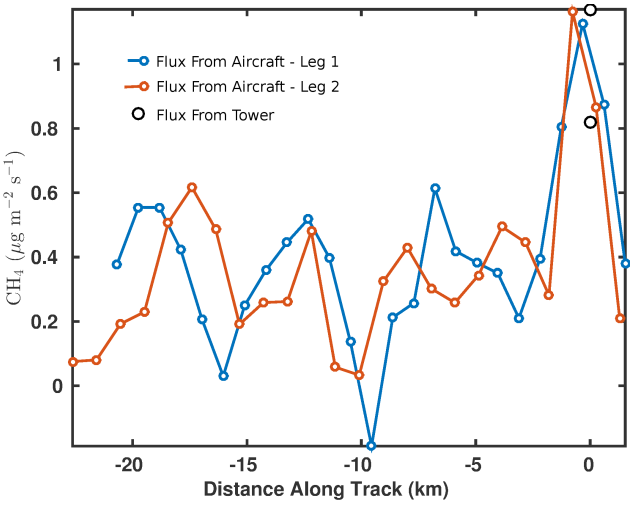


Figure 6. Plot of methane flux derived using RFM versus distance from flux tower for two flight legs on August 25. Positive (negative) distance is East (West) of the tower position. The East to West transect (blue) was flown 30 minutes after the West to East transect (orange). Black circles are the methane flux measured by the tower at the nearest time to when the aircraft passed the tower.

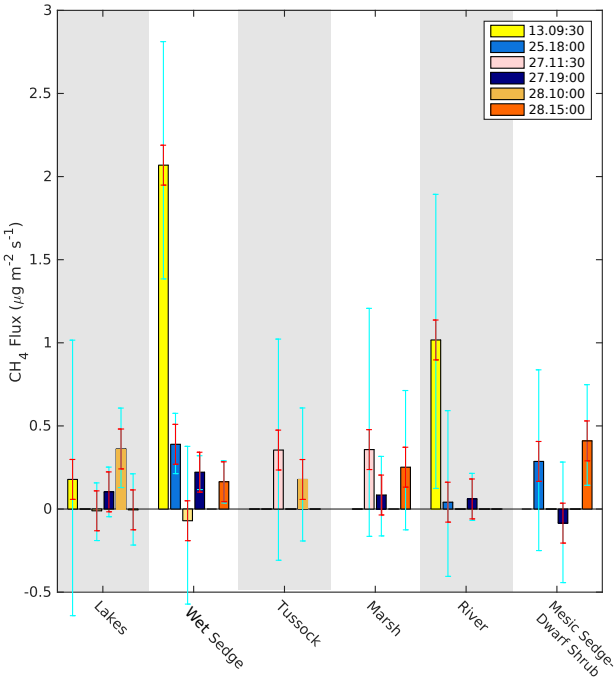


Figure 7. Mean methane fluxes by land surface class derived using the FFM for each of six flights as given in the legend. Dates of flights are given as day of month in August followed by the time of the middle of the flight. Bars give the instrument uncertainty (red) and the 95% confidence interval as calculated using bootstrapping (blue).