Estimating Random Uncertainty in Airborne Flux Measurements over Alaskan Tundra: Update on the Flux Fragment Method

RONALD DOBOSY

Atmospheric Turbulence and Diffusion Division, NOAA/ARL, and Oak Ridge Associated Universities, Oak Ridge, Tennessee

DAVID SAYRES AND CLAIRE HEALY

Harvard University, Cambridge, Massachusetts

EDWARD DUMAS AND MARK HEUER

Atmospheric Turbulence and Diffusion Division, NOAA/ARL, and Oak Ridge Associated Universities, Oak Ridge, Tennessee

JOHN KOCHENDORFER AND BRUCE BAKER

Atmospheric Turbulence and Diffusion Division, NOAA/ARL, Oak Ridge, Tennessee

JAMES ANDERSON

Harvard University, Cambridge, Massachusetts

ABSTRACT

Airborne turbulence measurement gives a spatial distribution of air–surface fluxes that networks of fixed surface sites typically cannot capture. Much work has improved the accuracy of such measurements and the estimation of the uncertainty peculiar to streams of turbulence data measured from the air. A particularly significant challenge and opportunity is to distinguish fluxes from different surface types, especially those occurring in patches smaller than the necessary averaging length. The flux fragment method (FFM), a conditional-sampling variant of eddy covariance in the space–time domain, was presented in 2008. It was shown capable of segregating the mean flux density (CO_2, H_2O, sensible heat) in maize from that in soybeans over the patchwork farmlands of Illinois. This was, however, an ideal surface for the method, and the random-error estimate used a relatively rudimentary bootstrap resampling. The present paper describes an upgraded random-error estimate that accounts for the serial correlation of the time/space series and the heterogeneity of the signal. Results are presented from the Alaskan tundra. Though recognized as important, systematic error estimates are not covered in this paper. Some discussion is offered on the relation of the FFM to other approaches similarly motivated, particularly those using wavelets. Successful measurement of the variation of air–surface exchange over heterogeneous surfaces has value for developing and improving process models relating surface flux to remotely sensible quantities, such as the vegetative land-cover type and its condition.

1. Introduction

Surface heterogeneity is a prime reason to make flux measurements from aircraft, but it is also a significant source of uncertainty. One approach is to sample above the blending height (Wieringa 1986; Mahrt 2000) to obtain a measure of the bulk flux. If the intent is to develop and improve models relating surface flux to remotely sensible parameters, however, it is important to acquire the signal of the surface's influence on air–surface exchange. Depending on the parameters remotely measured, the surface may be characterized categorically (e.g., by surface class) or by an interval quantity such as a vegetation index or the land surface temperature. Of the several approaches that have been used, some involve integral transforms. A flux can be

Corresponding author: Ronald Dobosy, ron.dobosy@noaa.gov

determined from the Fourier cospectrum of vertical wind with a scalar (e.g., Karl et al. 2009) or from scale-dependent shape functions intrinsic to the two signals themselves, determined by Hilbert transform and used to compute the covariance (Barnhart et al. 2012). These two schemes apply more readily to measurements recorded over homogeneous surfaces or above the blending height. For small-scale heterogeneity, the continuous wavelet transform is well suited (Torrence and Compo 1998; Metzger et al. 2013; Karl et al. 2009; Mauder et al. 2008) and is becoming the standard method for airborne work.

For a number of reasons, however, time/space schemes that do not involve integral transforms are also attractive. Ogunjemiyo et al. (2003) computed the flux of CO_2 and latent and sensible heat over a grid of $(16 \, km)^2$. They flew the grid in parallel east–west flight tracks overlaid by similar north–south tracks, thus defining 2-km segments. Passing twice over the grid provided a 4-km average over each segment. In such an approach, the effective length scale of the detrending scheme (Ogunjemiyo et al. 1997) was necessarily independent of the 4-km averaging length. Land cover was classified in $(30 \, m)^2$ pixels by remote sensing. Footprints were determined for each twice-repeated 2-km segment, their area classified using the $(30 \, m)^2$ pixels. The contribution from each surface type to each segment's flux density was then determined as the probability-weighted average of the flux density from each surface class over each footprint.

As seen from the method of Ogunjemiyo et al. (2003), flux computation in the space–time domain is tolerant of gaps. Also, the scale of the base state need not be the same as the averaging length. This allows for conditional sampling from the departure data stream at will as long as there are enough samples under each condition to form a stable covariance. The flux fragment method (FFM) operates on this principle.

Kirby et al. (2008) introduced the FFM and applied it to an Illinois cornfield, probably the ideal application. Fundamental to the validity of the FFM, with its conditional sampling of fragments over long series of measurements, is a defensible estimate of the uncertainty, first random, then systematic. The uncertainty estimate used by Kirby et al. (2008) was rudimentary, containing poorly justified assumptions. This paper's focus is on improvements in the estimation of random uncertainty drawing on the rich field of time series analysis. Because of the recognized utility of wavelet approaches, there will also be a discussion at the end of the relation between these approaches and the FFM.

2. Background

a. Specific challenges of airborne flux measurement

Turbulent fluxes are determined from measurements as a covariance between fluctuations in turbulent wind and in whatever scalar is being transported, normally computed as a temporal or spatial average. Airborne turbulence measurement (Leise et al. 2013) is an order of magnitude more complex than turbulence measurement from fixed towers. Turbulent wind is especially fraught with complexity and uncertainty. Turbulent structures, complex on their own, are measured from samplers themselves moving and accelerating in 6 degrees of linear and rotational freedom. Such sensor movement must be accurately measured at a high sample rate.

Moreover, although distorted flow is anathema to wind sampling, strong and deliberate flow distortion is the very essence of being airborne. Sensors are therefore best placed as far forward of the engine(s) and fixed or rotary wings as possible. Care in theory and measurement can then account for what distortion remains (Holder et al. 2011; Avissar et al. 2009; Kalogiros and Wang 2002a,b; Crawford et al. 1996; Brown et al. 1983).

Finally, and very important, outside of a hurricane the usual airspeed of an aircraft is much greater than the wind speed to be measured. The wind velocity relative to Earth is the vector sum of the flow velocity relative to the sensor [(negative) airspeed] and the sensor velocity relative to Earth (ground speed). These two vectors nearly cancel, leaving the wind velocity to be the small difference between large numbers. The validity of Taylor's frozen-turbulence hypothesis is enhanced, but so is the measurement uncertainty. The near canceling of the strong airspeed and ground speed vectors does have one helpful consequence: Many calibration errors and instrument malfunctions become quickly obvious due to the narrow tolerances of the measurements.

b. Historical overview

Progress over the past 40 years has greatly reduced the size and cost of good-quality airborne turbulence measurement. High-precision, low-drift GPS/inertial navigation systems (INS) are now available commercially off the shelf (COTS) with their accuracy and sensitivity documented in the manufacturers' literature. The wind (gust) probes themselves are normally custom built, though incorporating COTS sensors. Considerable effort has been made to quantify and minimize the uncertainty using wind-tunnel tests of the probes (Dobosy et al. 2013; Metzger et al. 2012; van den Kroonenberg et al. 2008; Garman et al. 2006), airborne maneuvers

(Mallaun et al. 2015; Leise et al. 2013; Vellinga et al. 2013; Metzger et al. 2012; Lenschow et al. 2007; Bögel and Baumann 1991), and postprocessing quality control (Vickers and Mahrt 1997). Although the complexity of the problem almost surely imposes a degree of irreducible uncertainty beyond that of fixed sensors, such work continues to define and expand the scope of productive airborne measurement. Nevertheless, airborne turbulence measurement remains largely a research activity involving a small and slowly growing community of groups, each with its own dedicated or convertible aircraft.

Important scalar quantities and their turbulent fluctuations are measured by increasingly capable instruments for ever smaller aircraft. The Anderson Group at Harvard University has developed a suite of instruments using integrated cavity output spectroscopy (ICOS) to measure methane, carbon dioxide, nitrous oxide, and water vapor along with their isotopologues at a sufficient rate to measure eddy covariance from a light twin-engine aircraft (Witinski et al. 2011).

c. Treating surface heterogeneity in physical space

Two physical-space approaches (Fig. 1) have been developed to treat surface heterogeneity—which one applies depends on the scale of the heterogeneity. If it is larger than the pathlength average required to acceptably approximate the ensemble average, a running flux method (RFM) can be used, analogous to a running mean with or without overlap (Zulueta et al. 2011; Vellinga et al. 2010; LeMone et al. 2003; Ogunjemiyo et al. 2003; Mahrt 1998; Lenschow et al. 1994). For repeated heterogeneity on scales small compared to this length the FFM applies (Kirby et al. 2008).

The FFM is a conditional sampling technique, conditioned on the surface type of origin for each "sample." The sample is a fragment of short duration (e.g., 1 s) composed of high-frequency measurements summed to avoid the unproductive computation of too many footprints. These short clips are called flux fragments to emphasize the need to average over a large collection of them to incorporate the range of atmospheric conditions present in the ensemble-average exchange with the selected surface.

This paper provides a second-generation uncertainty analysis for the FFM. The method relies on the ability to establish a long enough path of statistical homogeneity to approximate ergodicity (Gluhovsky and Agee 1994). Application to the RFM was not attempted here because existing approaches to the RFM currently appear more fruitful.

3. Measurements: Type and location

a. Airborne measurements

The Anderson Group from Harvard collaborated with NOAA's Atmospheric Turbulence and Diffusion Division to produce an airborne system for eddy covariance flux measurement. Harvard's ICOS system and their expertise in interpreting its raw signals provided greenhouse-gas concentrations at a high rate (Witinski et al. 2011). The NOAA group's Best Aircraft Turbulence (BAT) probe measured the three components of wind along with the atmospheric temperature and pressure. Aurora Flight Sciences, Inc. provided a light twin-engine aircraft and the aeronautical engineering expertise to adapt it to carry the instruments. The system was named Flux Observations of Carbon from an Airborne Laboratory (FOCAL). The NOAA group also provided expertise in airborne flux analysis (Dobosy et al. 2013; Kirby et al. 2008; Crawford and Dobosy 1992). The first FOCAL research flights were over Alaska in 2013 (Sayres et al. 2017; Dumas et al. 2014).

In view of the potential significance of the large reservoir of carbon stored in permafrost regions, many surface measurements of CO_2 and CH_4 have been undertaken on the North Slope to date, most based on chamber or short flux tower techniques. These observations are currently the best available to determine the larger regional flux and any trends with time, but the upscaling in space is considerably aided by airborne measurements.

The value of airborne measurement of greenhouse-gas emissions from small aircraft at low altitude over the Arctic tundra has long been evident (Zulueta et al. 2011; Oechel et al. 2000, 1998; Brooks et al. 1997). The drivers are the difficulty in accessing large areas of tundra, the low relief of terrain, and the high amplitude of spatial heterogeneity in the fluxes. Improvements in the ability to measure methane as well as isotopologue ratios for methane and carbon dioxide have opened up new opportunities in such airborne measurements.

b. Surface measurements and general overview

Fixed measurements on the surface are synergistic and indispensable to an airborne campaign measuring air–surface exchange. The primary role of surface measurements is to provide continuous 24-h temporal coverage while the aircraft provides wide-ranging spatial coverage. The FOCAL campaign on Alaska's North Slope (Sayres et al. 2017) included a surface tower measuring fluxes of latent and sensible heat, CO_2, and CH_4, along with soil temperature and moisture. This was located in wet sedge south of the Deadhorse Airport, Prudhoe Bay, Alaska.
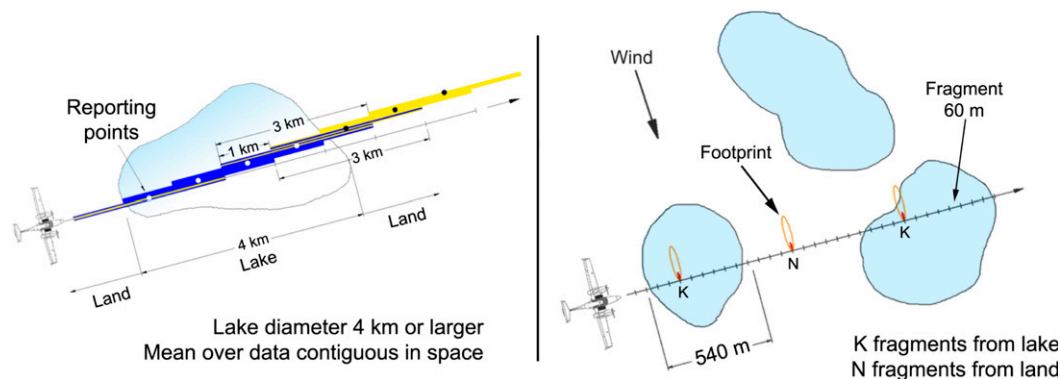
FIG. 1. Methods for computing air–surface exchange in time and space: (left) RFM and (right) FFM. The left panel depicts overlapping 3-km averages (yellow and blue segments). Each is composed of three parallel bars. All are blue over the water; all are yellow over the land. Over mixed water and land, the segments are blue–yellow–blue or yellow–blue–yellow depending on the proportions. Fragments in the right panel are accumulated separately for each major surface type and averaged separately to distinguish between the contributions from each surface type.

The airborne campaign included 11 successful flights of various types out of Deadhorse Airport. Fluxes of methane and latent heat were combined with footprint analysis (Kljun et al. 2004) to calculate and map the local and regional fluxes of H_2O and CH_4 for six of these flights. These can help to define the landscape-scale drivers of heterogeneity and variability. Three flights followed a straight transect past the tower (Fig. 2). They provide the examples used in this paper. Despite limitations inherent in the first field deployment of a new complex system, a large regional coverage was accomplished as shown by Sayres et al. (2017).

c. Land cover along the flight track

The surface under the tower-run flight track (Fig. 2) is largely a collection of tundra, small lakes, and river valleys, The more detailed classification of the North Slope Science Initiative (NSSI) shows the prominent individual surface classes to be wet sedge and lake followed by mesic (moderately wet) sedge and marsh. The distribution of these types along the flight track (Fig. 8; Table 1) also includes a significant fraction that does not fall into any one of these categories.

4. Flux fragment method: Necessary background

The uncertainty estimation process described in section 5 of this paper operates not on the individual data streams (turbulent wind w and a scalar) but on flux fragments (section 2c). Some necessary background on the definition of fragments and the implementation of the FFM for the example dataset is therefore given in this section. The basic data segment is a period of straight and level flight between turns and other maneuvers that can contaminate flux computations.

a. Selecting the averaging length

The eddy covariance procedure ideally uses a measured series that is long compared to the characteristic cross-correlation scale of w and the transported scalar, for example, trace-gas concentration. In practice one seeks a gap (if available) in the cospectrum of w with the scalar. In Fig. 3 mean spectra from flight 13.09:30 on 13 August (Sayres et al. 2017) are superposed through multiplication by a constant. Their maximum spectral energy density is at about 0.4 Hz for w and about 0.7 Hz for the water vapor mixing ratio. Methane's maximum energy density may be above the Nyquist frequency of the sample rate. The reduced spectral density of w above its maximum at 0.4 Hz, however, mitigates the associated loss of methane flux, estimated to be around 10% (Sayres et al. 2017).

The cospectra from 13 August (Fig. 4) use the series unmodified except for mean removal. For methane, the maximum cospectral density is difficult to identify, while for water, apart from the isolated spike, the cospectral maximum is at 0.5 Hz. A marked reduction in cospectral density occurs for both gases near 0.05 Hz, reflected in the ogive plots (integrated cospectra, lower panels) reaching near horizontal by about 0.02 Hz. Methane has a wide spread among the ogives for the individual segments, but they all tend to flatten between 0.05 and 0.02 Hz. On this basis the averaging length is chosen to be 50 s, or about 3 km at the $60 \, m \, s^{-1}$ nominal airspeed.

b. Determining the base state

The base state in principle describes the deterministic mesoscale environment embedding the random turbulence. If a cospectral gap separates the mesoscale from the turbulent scale (section 4a), then the base state for
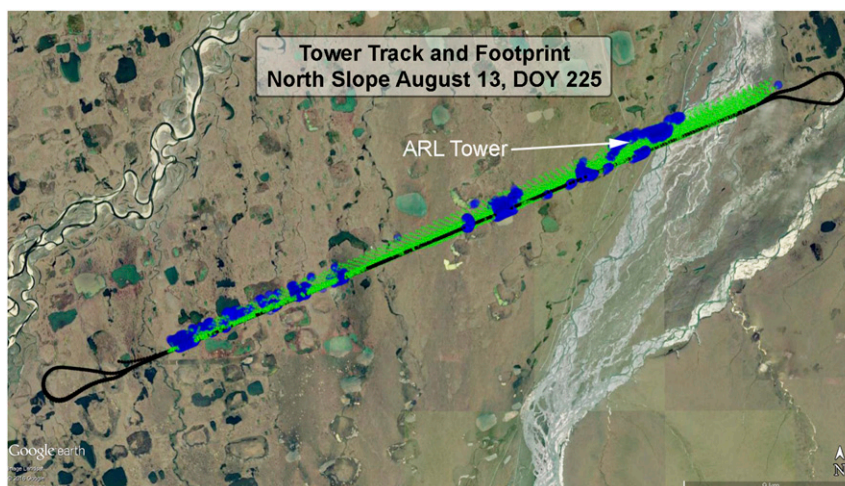
FIG. 2. Flight track past fixed surface tower with footprint spots from 13 Aug 2013 showing water (blue spots) and other surface types (green spots), of which wet sedge has the largest fraction and the strongest CH_4 emission. Prudhoe Bay is 20 km north of the ARL tower.

the RFM is simply the mean over the averaging length. The FFM, however, requires a longer track, sometimes much longer, to accumulate an averaging length's worth of fragments from a given surface. The best condition for confident specification of such a base state is a boundary layer in which the scale of turbulence is clearly separated from that of strong mesoscale flows, such as storms, frontal passage, land–sea breezes, or katabatic flows. The relatively flat North Slope, its heterogeneity on horizontal scales being typically only hundreds of meters, is among the better suited natural systems, although two of the six analyzed flights showed a mesoscale structure beyond the cospectral minimum.

In the absence of strong mesoscale patterns, the averaging length scale and larger can be assigned to the base state. In the space–time domain, this is usually done by suitable averaging or curve fitting. The flux fragment method, which requires an approach highly tolerant of

gaps, favors curve fitting. For example, the features larger than 3 km were found to be adequately fit by a fourth-order polynomial along each segment of the track on 13 August. Fourth-order polynomials fit to the vertical wind component (Fig. 5, upper) and the mixing ratio of methane (lower) are shown superimposed for comparison on 1-s nonoverlapping block averages from two segments of the time series from 13 August. The fit is not perfect. For example, the sharp drop-off in the mixing ratio of methane between the two segments in Fig. 5 is probably real, being somehow associated with the Sagavanirktok (Sag) River at the northeast end of the transect (Fig. 2).

Overfitting of the base state is problematic with the FFM. A large homogeneous region of the same sort of surface will skew the base state toward that surface type, degrading the signal from that surface in the turbulent fluxes. The base state needs to be defined on such scales

TABLE 1. Means and 95% confidence limits (low, high) for the histograms in Figs. 10–12. Bold digits indicate means significantly greater than italic digits.

| Gas flux			Methane (μg m^{-2}s^{-1})			Latent heat (W m^{-2})		
Date	Surface	Length (km)	Mean	Low	High	Mean	Low	High
13 Aug 2013	Wet sedge	53	**2.06**	1.35	2.78	**62.3**	54.7	69.9
	All	185	**1.42**	1.08	1.76	**52.7**	49.0	56.6
	Lake	12	*0.16*	−0.66	0.94	*30.0*	20.0	40.2
	Moderately wet sedge	3	−0.89	−3.02	0.98	*22.4*	−0.35	43.9
25 Aug 2013	Wet sedge	14	0.24	0.06	0.43	3.7	2.0	5.6
	All	55	0.35	0.24	0.47	5.7	4.4	7.0
27 Aug 2013	Wet sedge	25	0.23	0.12	0.34	*−0.32*	−1.03	0.37
	All	152	0.16	0.12	0.20	*0.42*	0.12	0.71
	Lake	10	0.08	−0.07	0.22	−0.09	−0.90	0.75
	Marsh	5	0.08	−0.15	0.29	**1.99**	0.73	3.34
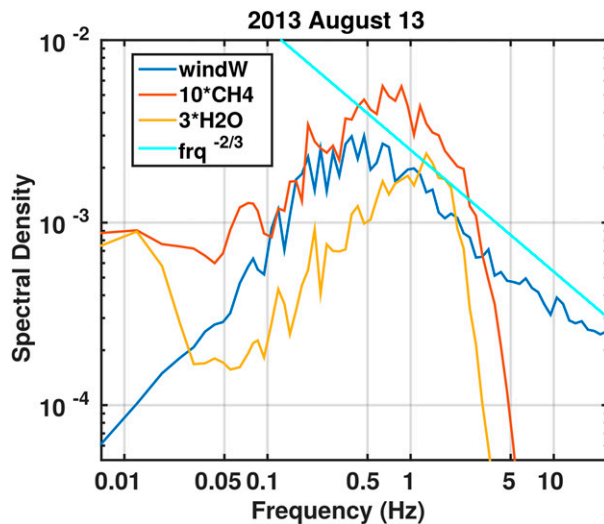
FIG. 3. Mean power spectra of vertical wind w and mixing ratios of methane and water over the eight straight segments of flight 13.09:30 on 13 Aug. Each segment's mean, but not its trend, is removed. Methane and water series are multiplied by 10 and 3, respectively, to overlay the curves. Spectral density is multiplied by the frequency, hence the spectral density in the inertial subrange is proportional to the $-2/3$ power of the frequency.

that the subfilter surface structure appears full force in the departure quantities.

Here we note, concerning wavelet techniques, that the scales in their domain extend to the scale of the averaging length from section 4a and larger. The vertical wind at these larger scales, however, is typically on the order of $0.1\,\mathrm{m\,s^{-1}}$. Note from Fig. 3 that the spectral density of w at $0.02\,\mathrm{Hz}$ is 13 dB below its peak at $0.4\,\mathrm{Hz}$, corresponding to a factor between 4 and 5 in amplitude. Also in Fig. 5 the fluctuation of the base-state w is no more than $\pm0.1\,\mathrm{m\,s^{-1}}$, about one-sixth of the range of w. This is on the lower edge of detection from the current airborne measurement. The airspeed and ground speed vectors, which nearly cancel, have a magnitude of about $60\,\mathrm{m\,s^{-1}}$. Accurately determining a wind component of $0.1\,\mathrm{m\,s^{-1}}$ requires much better than 99.4% accuracy in speeds and angles, an issue that must be addressed whether using the wavelet domain or the space–time domain.

c. Defining the flux fragments

With the base state defined and removed, the remaining path series is treated as a purely random process. The flux is a covariance between high-rate measurements of vertical wind w and a transported scalar, such as methane's mixing ratio with dry air r_{CH4}. Flux fragments are convenient-length partial sums (not averages) of the constituent cross products. For example, V is the airspeed, ρ is the (dry) air density, n is the sample rate, and $\delta t = n^{-1}$ is the sample interval. Primes

indicate a departure from the base state. A τ-second "native" fragment of methane flux is then (Webb et al. 1980)

$$f_{CH4} = \sum_{i=1}^{n\tau} (\rho_i w_i)' r'_{i,CH4} V_i \delta t. \quad (1)$$

Typically, $\tau = 1\mathrm{s}$ is long enough to significantly reduce the computational effort, particularly in computing footprints, but short enough to resolve the features of the surface. For the FOCAL system with 1-s fragments, $n = 50$. The native fragments have uniform duration but varying length depending on the airspeed:

$$L = \sum_{i=1}^{n\tau} V_i \delta t. \quad (2)$$

A subset S containing p fragments, contiguous or not, ordered or not, can be plucked and averaged to produce a flux estimate:

$$F_{CH4}(S) = \sum_{k=1}^{p} f_{CH4}(k) \left(\sum_{k=1}^{p} L_k \right)^{-1}, \quad (3)$$

where the denominator of (3) is the fragments' cumulative length. The subsets of primary interest are those determined by footprint analysis to all have come from instances of one given surface class. If enough such fragments are available from a given flight, then the air–surface flux can be estimated for that class of surface. Alternatively, the dependence of the flux on an interval quantity, such as the enhanced vegetation index (EVI), can be determined. The weighted-mean EVI is computed for each fragment's footprint and is mapped to the associated value of the flux by regression or by a machine-learning algorithm (e.g., Metzger et al. 2013).

5. Estimating the random uncertainty

The estimation of random uncertainty presented here is a second generation from that of Kirby et al. (2008). Although still based on bootstrap resampling (Efron and Tibshirani 1998), the present analysis additionally accounts for the serial correlation and heterogeneity of the series of flux fragments. It uses a method applied by Mudelsee (2010) to climate processes, but it is also applicable to turbulence.

The uncertainty analysis operates on the fragments, not the individual data streams (w or scalar). Although the native fragments from (1) are computed over uniform time intervals ($\tau = 1\mathrm{s}$), their length in air varies with the airspeed elevating the interfragment variance. This extra variance is removable by interpolating the fragments to uniform length in air. The interpolation
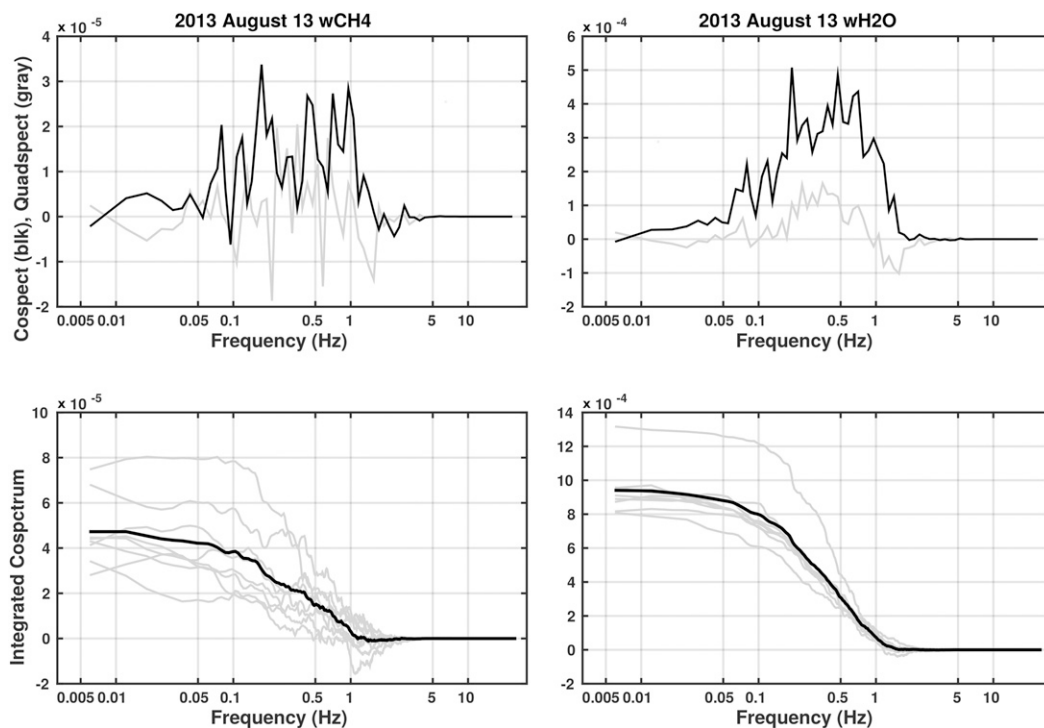
FIG. 4. Mean cospectra of w with mixing ratios of methane and water vapor from flight 13.09:30. Mean, but not trend, is removed from each of the eight individual segments of straight and level flight from which data were taken. Gray curves in the lower panels show the integrated cospectra from the eight segments individually.

operates on a cumulative sum over all fragments in a flight segment as a function of the pathlength, the cumulative sum of their individual lengths. A cubic-spline interpolation is evaluated at uniform intervals of pathlength, typically the distance traveled in 1 s at the mean airspeed over the flight. The first difference recovers the constant-length fragments. This step reduced the interfragment standard deviation by about 5%.

The next step is to extract and assemble the sub-samples for each surface of interest from the set of constant-length fragments. This must precede Mudel-see's processes or much of the desired subfilter signal is obliterated. Then for each surface type individually, consider a process represented by the random variable $X(S)$ dependent on the pathlength S and having both deterministic and stochastic components; that is, $X(S)$ is partially—but not fully—known. Such a process is observed as a serially correlated series. With geo-physical turbulence measurements as with climate, it is rare to have on hand more than one such realization comprising a sequence of measurements such as $x(i) = f_{CH4}(i)$, where the small x indicates a realization of $X(S)$. The random uncertainty can be estimated from this realization by bootstrap resampling from its statis-tical component once the deterministic component is

estimated and separated out. To do this Mudelsee de-composed $X(S)$ as

$$X(S) = X_{\text{trend}}(S) + X_{\text{out}}(S) + \sigma(S)X_{\text{acor}}(S), \quad (4)$$

where the trend $X_{\text{trend}}(S)$ and scaling function $\sigma(S)$ are deterministic and interpreted broadly. This trend is not to be confused with the base state defined in section 4b, since we are working here with the fragments, that is, the (nonlinear) covariances of departures from that base state. The outlier function $X_{\text{out}}(S)$ provides for a model of spikes or other valid but anomalous information. For example, the airplane may encounter the plume from human drilling operations or other nearby point sources of greenhouse gas (Brooks et al. 1997). The stochastic component $X_{\text{acor}}(S)$ is weakly homogeneous (i.e., having constant mean and variance along the track) and serially correlated, a space–domain reflection of the spectral properties evident in Fourier space. Weak homogeneity, as opposed to identical distribution, is sufficient, since moments higher than the variance of these fragments are not used. In general X_{acor} is not Gaussian. In the buoyancy-driven atmospheric mixed layer, X_{acor} can be strongly skewed and have high kurtosis (Lenschow et al. 1994). The bootstrap analysis used here is, however, nonparametric.

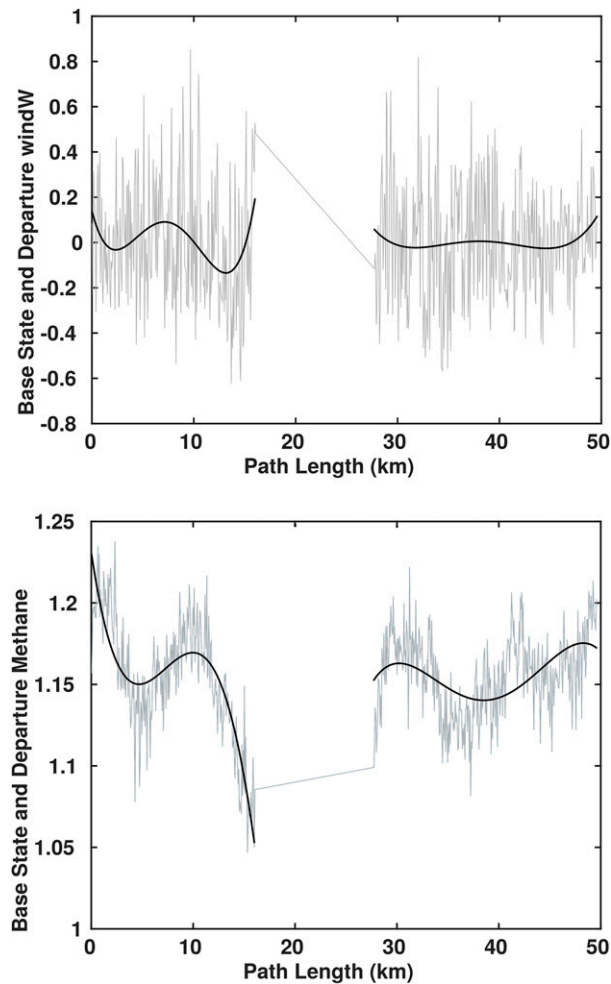FIG. 5. Fourth-order base state for vertical wind $(\mathrm{m\,s^{-1}})$ and mixing ratio of methane $(\mathrm{mgCH_4[kg(dry)]^{-1}})$.

Since the bootstrap process requires independence as well as weak homogeneity, the serial correlation of $X_{\mathrm{acor}}(S)$ is addressed by drawing on a parametric first-order autoregressive (Markov) model,[1]

$$
\begin{aligned}
X_{\mathrm{acor}}(1) &= \mathscr{E}(1), \\
X_{\mathrm{acor}}(i) &= a(i)X_{\mathrm{acor}}(i-1) + \mathscr{E}(i)[1-a(i)^2]^{1/2}, \\
& \qquad i = 2, \dots, n.
\end{aligned} \tag{5}
$$

Here $\mathscr{E}(i)$ has zero mean and unit variance for all i. For $i = 2, \dots, n$, $\mathscr{E}(i)$ is multiplied by $[1-a(i)^2]^{1/2}$ to retain the unit variance of $X_{\mathrm{acor}}(i)$ for all i. Parameter $a(i)$ defines the serial covariance for an unevenly spaced series,

---

[1] This implied smooth passage between continuous and discrete random variables is valid for a first-order Markov process but not always; see Mudelsee (2010).

such as encountered with the FFM. It decays exponentially with separation distance,

$$
a(i) = \exp\{-[S(i)-S(i-1)]/\lambda\}, \tag{6}
$$

where λ is the decorrelation length and $S(i)$ is the location of the center of fragment i along the path. The estimation of λ uses the full series regardless of surface type (as in Fig. 7).

The realized $x_{\mathrm{acor}}(i)$ term emerging from the detrending and scaling in (4) for each surface type should be at least weakly homogeneous (Gluhovsky and Agee 1994). With gaps in the data, however, accomplishing this can range from tedious to impossible depending on the sample size. Time constraints restricted the detrending in this case to a simple straight-line fit to each segment, or to the whole flight if the total subsample size from a flight was under 100 samples. A single variance was computed per segment (or per flight) after detrending. Visually, this appeared adequate for two of the three cases. The third case, 25 August, was markedly heterogeneous, showing a repeatable sharp peak at the tower in the RFM flux (Sayres et al. 2017). It was treated in the same way as the others, for want of a readily available alternative.

The process of isolating the (comparatively) stationary and serially decorrelated term $\varepsilon(i)$ of (5) for 13 August (Figs. 6 and 7) is shown as an example. The modified symbol ε indicates a realization from data. The constant-length fragments of Fig. 6 are only mildly autocorrelated at lag 1 but show considerable spread in their integrated autocovariance with increasing lag out to 50 s. With detrending and scaling, the residuals

$$
x_{\mathrm{acor}}(i) = [f_{\mathrm{CH4}}(i) - x_{\mathrm{trend}}(i)]\sigma_p^{-1} \tag{7}
$$

(Fig. 7, left) have a reduced spread evident among their lag-covariance curves. Decorrelating the residuals $x_{\mathrm{acor}}(i)$ to obtain the realized $\varepsilon(i)$ uses the first-order Markov model [(5)]:

$$
\varepsilon(i) = [x_{\mathrm{acor}}(i) - a(i)x_{\mathrm{acor}}(i)][1-a(i)^2]^{-1/2} \tag{8}
$$

using $a(i)$ from (6) with a decorrelation length $\lambda = 0.5$. The corresponding Markhov parameter is $a = 0.13$ for adjacent fragments, rapidly falling to zero with a widening gap. Mudelsee (2002) makes available a software package to estimate the value of λ. However, it assumes a Gaussian distribution for X_{acor}, while the skewness and kurtosis of x_{acor} are 0.73 and 7.63 (Gaussian has 0 and 3), respectively, for methane flux fragments and 3.17 and 25.5, respectively, for
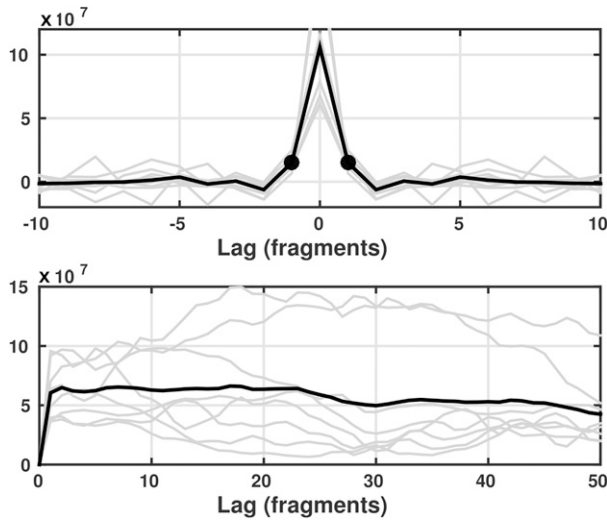
FIG. 6. Case 13 August 2013. (upper) Autocovariance plots of the methane-flux fragments $f_{CH4}(i)$ from the eight individual flight segments (gray) and their mean (black). The lag is in the fragments' length (61.01 m for 13 Aug) along the flight path. The black spots emphasize lag-1 autocovariance. (lower) Cumulative integrals of these autocovariances, reflective of the cospectral relation in Fourier space.

latent heat flux fragments. Therefore, the lag-1 covariance of zero evident in Fig. 7 (right) was achieved by trial and error. For this 13 August example, one unit of lag is 61.01 m, the length of a (constant length) fragment.

Serial correlation implies fewer degrees of freedom in a sample than there are members. The decorrelated series, in principle, is an independent sample having therefore the same number of degrees of freedom as members. The first-order Markov expression [(8)] provides the additional degrees of freedom at the expense

of increased variance at the high frequencies, verified by comparing the spectra of $x_{acor}(i)$ and $\varepsilon(i)$ (not shown).

The realized $\varepsilon(i)$, $i = 2, \ldots, n$ from (8) form the population of (ideally) independent and weakly homogeneous quantities from which N bootstrap resamples are drawn (here, $N = 2000$). Each resample has $n - 1$ members, n being the number of fragments representing the given surface class.

An ensemble of realizations of $X(S)$ from (4) is produced in reverse of the path to $\varepsilon(i)$. For ensemble member ν, $\nu = 1, \ldots, N$, an initial $x_{acor}^{\nu}(1)$ is randomly selected (with replacement) from the collection of realized x_{acor} from (7). Applied forward, (5) has the form of a low-pass causal filter working from an ensemble $\varepsilon^{\nu}(i)$, $\nu = 1, \ldots, N$ of independent realizations of $\mathscr{E}(S)$, defining $\varepsilon^{\nu}(1) \equiv x_{acor}^{\nu}(1)$. It produces a corresponding ensemble of $x_{acor}^{\nu}(i)$. These in turn are multiplied by the scale σ and added to $x_{trend}(i)$, both deterministic, hence constant with ν apart from their own uncertainty, which is beyond the present scope.

The new ensemble of (bootstrap) realizations of $f_{CH4}^{\nu}(i)$ can be used in any way the original realization can. Each such realization retains the original data stream's dependence on surface type and its heterogeneity of mean and variance over the flight. Such ensembles are developed for flux fragments of both methane and latent heat for the significant land classes from each of the three flights, a total of 20 ensembles. For the present each flux fragment series in each ensemble is averaged to produce a bulk flux (the distributions of these ensembles of bulk-flux realizations are displayed in Figs. 10–12 and were used to estimate 95% confidence intervals for the methane and latent heat fluxes.
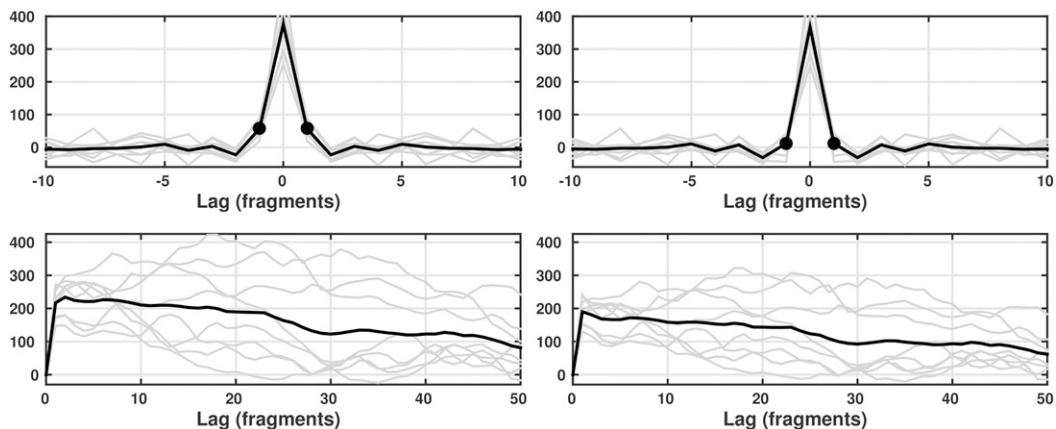

FIG. 7. As in Fig. 6, but showing (left) the residuals $x_{acor}(i)$ from (7) and (right) the random process $\varepsilon(i)$ from (8). Decorrelation resulting from application of the first-order Markov model is evident in the near-zero value of the lag-1 covariance (black spots) on the right.

6. Results and discussion

a. Intercomparison of airborne- and tower-based measurements

Intercomparison flights are a standard part of every campaign to verify the intercomparability of the systems and to provide the spatial/temporal link necessary to the synergy. A companion paper (Sayres et al. 2017) shows results from all three intercomparison flights of the 2013 FOCAL campaign. Flight 13.09:30 on 13 August is reproduced here to provide a context for the uncertainty estimates.

The 13 August case had strong insolation (for northern Alaska) shortly before the sharp drop in soil temperature (measured at the tower) and weaker turbulent flux, signaling autumn later in August. The tower was in wet sedge, and the portion of the flight track determined by the flux fragments' footprints to be over wet sedge is shown by the green spots in Fig. 8. Wet sedge along the track emitted methane at very nearly the same rate as at the tower (Fig. 9). This was greater than the mean methane flux over all fragments, though not significantly. In contrast, both the wet sedge and the general track emit significantly[2] more methane than either mesic sedge (msds) or open water (lakes). Indeed, the emission from the lakes and mesic sedge is not significantly different from zero.[3] The tendency of methane emission to be stronger from warmer and wetter soil is consistent with the diurnal pattern exhibited by the tower and in the greater emission from wet sedge than from mesic sedge.

The latent heat flux evidently responded sharply to the rise in solar radiation after 0700 local time (LT, UTC −10 h). As with methane, the sedge fraction of the flight track emits latent heat more intensely than either the open water or the track overall. The open water, however, emits significantly more than zero latent heat. Also, the latent heat flux from the sedge along the flight track is reported significantly lower than that from the tower. The influence of the airborne instrument cannot be ruled out because it was not optimized for measuring water vapor. Future deployments will include such
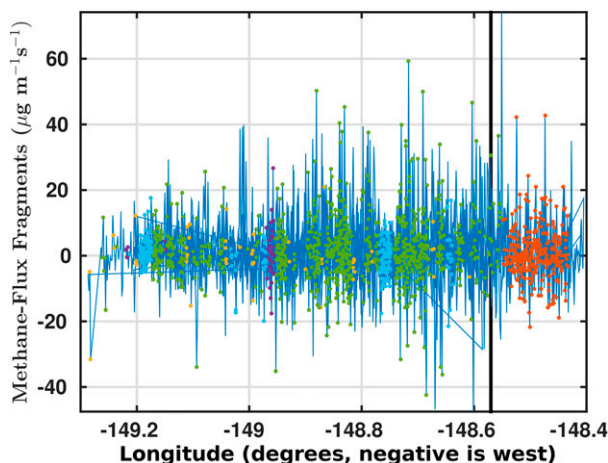

FIG. 8. Fragment values for 13 Aug 2013, with the eight passes on the flight line depicted in Fig. 2. Specific surface classes are indicated: wet sedge, green; mesic (moderately wet) sedge, yellow; lakes, light blue; marsh, purple; and Sagavanirktok River, red. All fragments are connected regardless of surface type (dark blue lines). Location of the tower is shown (heavy black line).

optimization. Additional findings from the expedition are given by Sayres et al. (2017).

b. Summary of uncertainty estimates

The uncertainty assessments show physically plausible results on all three Alaska tower flights. The mean and 95% confidence limits for each flight, surface, and gas appear in Table 1. The 13 August case, with its stronger insolation (0930 LT) and higher soil temperature has been discussed. The 25 August case had a relatively short sample (two passes) taken over reduced soil temperature and late in the day (1800 LT). Unlike the other two flights, it had stronger flux from the total path than from wet sedge alone for both gases, probably reflecting the strong peak at the tower noted in section 5. On 27 August, also in late season and late day (1900 LT), the marsh class had stronger latent heat flux than any other surface, significantly so from the wet sedge and total path, and almost significantly so from the lakes.

For visual representation of Table 1, the uncertainty distribution of methane or latent heat flux from wet sedge is compared in Figs. 10–12 with that from another surface class during the same flight. No class represented by less than 3 km of cumulative pathlength was included. Each histogram contains 2000 of these means over a pathlength series reconstructed from bootstrap resampling as described in section 5.

Source series with fewer samples produce broader and less well-determined confidence intervals (i.e., less confidence). Bootstrap estimates of confidence intervals, based on the lower-probability regions (tails)

---

[2] We are requiring here, mostly for convenience, that the two uncertain quantities' 95% confidence intervals not overlap. This is a strong criterion. If two uncorrelated surfaces actually emitted the same mean flux density, then this observed outcome would be below the 2.5th percentile for wet sedge *and* above the 97.5th percentile for the lake or the mesic sedge, a probability less than 0.0625%.

[3] The confidence interval has its ordinary meaning when comparing an uncertain quantity to a known quantity.
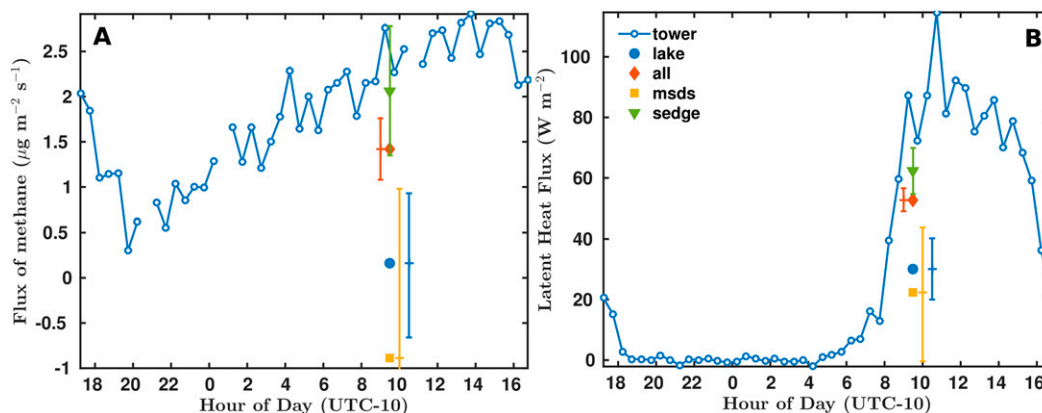
FIG. 9. Comparing (a) methane and (b) latent heat flux measured at the tower with the same measured from the air. MSDS is mesic (moderately wet) sedge with dwarf shrub. Both panels include all fragments; 95% confidence intervals are indicated for the airborne measurements

of a distribution, always carry greater uncertainty than estimates of central measures (e.g., mean): the smaller the original sample, the greater the uncertainty. The ways to estimate and optimize the statistical power for smaller samples have not yet been explored.

c. Reynolds criteria

The fragments defined in section 4 are products of terms of two detrended data streams. The function of the detrending is to distinguish between the ogive-determined scales of turbulent vertical flux and larger scales to be treated four-dimensionally. The detrended (departure) quantities all sum to zero over each segment of a multisegment flight; hence, the Reynolds criteria (Monin and Yaglom 1971) are valid over each entire flight segment.

The conditionally sampled subsets represent individual ensembles within the stream of turbulence. In this paper, each ensemble corresponds to a surface class. Combining and averaging enough fragments to characterize the distribution of ambient atmospheric states that are broadly homogeneous in the absence of strong mesoscale features provides an estimate of the flux from the given surface type. Shorter samples carry lower confidence, as noted in the previous section. Deriving the base state from the full series allows the subsets to be readily combined to obtain the bulk mean flux density over a segment. The subsamples then represent a partition of this segment into meaningful components.

d. Flux fragments and wavelets

The FFM is presented as a useful alternative approach to wavelet analysis, which is becoming the standard for turbulence measurements from aircraft over heterogeneous surfaces. Wavelet transforms, a middle method

between the space–time and frequency domains, can provide information on both location and scale, thus balancing the assets and liabilities of the two extremes. The work of Farge (1992) and Torrence and Compo (1998) with continuous wavelet transforms have produced a practical method that suits turbulence well. Wavelet transforms can describe turbulence in limited atmospheric subregions like space/time approaches while integrating over many samples like Fourier methods.

A recognized advantage of the wavelet technique is the facility with which it treats scales beyond those determined to belong to the turbulence. This avoids systematic error in estimating air–surface exchange due to otherwise unaccounted transport on scales larger than those of turbulence. For example, Foken et al. (2006) examined ogive plots for fixed sensors and found an energy flux beyond the normal averaging scale of 1800 s and having sign and magnitude sufficient to account for the long-recognized imbalance in the measured surface energy budget. Mauder et al. (2007) applied wavelet analysis to the 115-km Candle Lake flight line from the BOREAS study (Sellers et al. 1997) and found significant mesoscale patterns from land–water and other contrasts along the track. Mahrt (2010) noted the problem of omitting stationary eddies, such as rolls aligned with the mean wind (LeMone 1973) or secondary circulations anchored to a particular surface feature and not advected past fixed instruments.

While acknowledging the importance of mesoscale structures when these are present, the current focus of the FFM is on vertical turbulent transport. In this role, an important asset of the FFM is its gap tolerance. The benefit is greater than it may appear. Like the FFM, wavelets can provide short segments for conditional
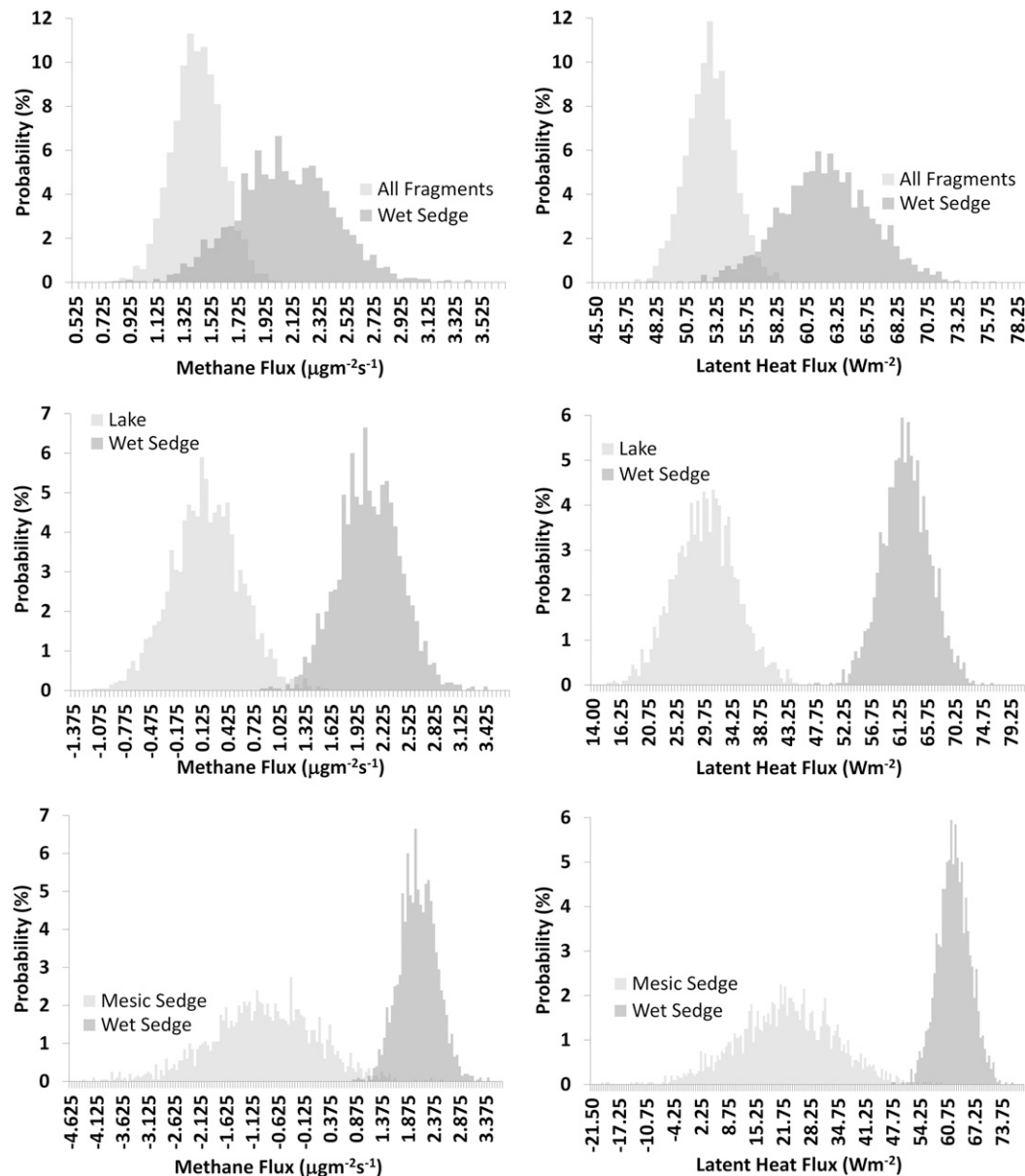
FIG. 10. Case 0930 LT 13 Aug 2013. Distributions over bootstrap samples pairing wet sedge (dark gray) with the other primary surface classes present. Tower is in wet sedge. Mesic sedge is midway between wet and dry.

sampling. Furthermore, these include a multiscale characterization of the environment called the cone of influence (COI) (Torrence and Compo 1998). This additional benefit, however, requires a gap-free data stream over the width of the COI, 3 times the largest wavelet scale for the commonly used Morlet wavelet. The overhead cost to a flight segment beyond the working section is therefore 3 times the maximum scale of the turbulence, more if larger mesoscales are included. For the FFM, the shortest flight leg need not be

longer than 1 or 2 times the maximum scale of turbulence to properly partition the flow between deterministic and turbulent. There is no overhead, and segments can be repeatedly sampled, or they can form a multisegment track when required by terrain, population, or airspace.

A wavelet-derived flux estimate is locally self-sufficient. Over a surface characterized by finescale patches of differing surface type, however, the larger-scale wavelets must increasingly contaminate the signal
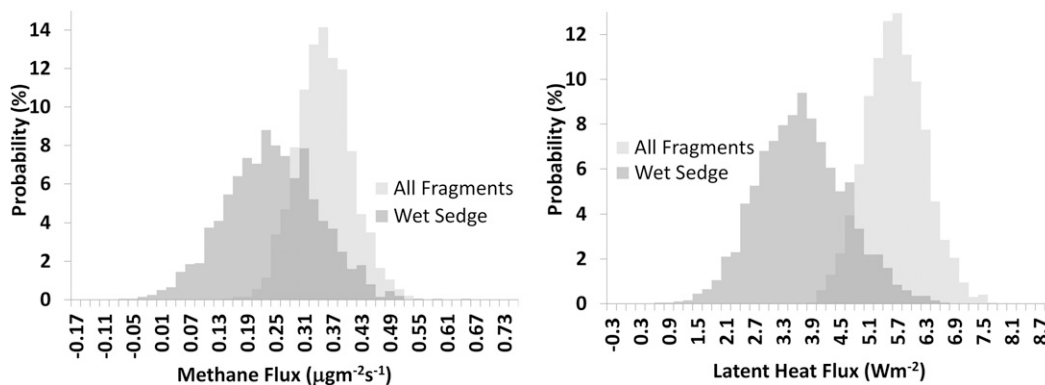
FIG. 11. Case 1800 LT 25 Aug 2013. As in Fig. 10, but this flight was short, and wet sedge was the only significant individual contributor. See Sayres et al. (2017) for details.

from the central surface type by involving samples gathered over adjacent surfaces with different characteristics. A flux fragment, in contrast, is a snap sample from a single source comprising all scales of turbulence. It, however, requires additional fragments to form an average. Typically, these additional fragments come from separate instances of similar surfaces (some more similar than others). This issue motivates the uncertainty estimation in the current paper.

The wavelet scheme is well suited to the Candle Lake run studied by Mauder et al. (2007). Its run was long compared to the intended maximum wavelet scale. Its mesoscale patterns under strong springtime land–water and other contrasts were well defined and largely two-dimensional in the vertical plane along the flight track (Sun et al. 1997). Otherwise, there was low-relief terrain and sparse population.

The FFM is most attractive when the surface offers a patchwork of strongly contrasting land classes in an area requiring shorter, more complex flight tracks. For this, the Illinois study of Kirby et al. (2008) in June was ideal. Maize has strong CO_2 uptake relative to soybeans, and the area is likely not a source of strong mesoscale patterns. The flight segments, however, must be at least as long as the scale of the base state, and longer flight legs are more efficient. Also, if larger mesoscale patterns are significant, they can and still must be treated in another way.

A lot depends on the design of the measurement program to fit the questions being asked. Airborne systems best function where the flow is 1) horizontally homogeneous, 2) essentially two-dimensional, or 3) separable into segments having such properties. Airborne campaigns are generally not sufficient in themselves. They are better supplemented to the extent possible by scenario-specific modeling, multisite fixed

measurement, remote sensing, and extensive use of operational observations and models.

7. Conclusions

The analysis of random uncertainty developed in this paper for the flux fragment method (FFM) in the space–time domain supersedes Kirby et al. (2008), in that it incorporates the effect of the serial correlation and heterogeneity in the data stream. This aligns the uncertainty analysis more closely than before with the character of the dataset being analyzed. The deterministic aspects of the analysis were also more clearly delineated, although uncertainty was not estimated for them.

The influence of the serial correlation was found to be small: the decorrelation length was half the width of a fragment. The resulting first-order Markov parameter a was 13% for adjacent fragments falling exponentially as the gap between them increased. This result could be bolstered by an explanation from turbulence theory. A number of expediencies remain in the present method, some of which were indicated in the text. The uncertainty in this analysis was wholly assigned to the stochastic (random) component. The trends and variances (x_{trend}, σ), though also computed from the data, were not subjected to uncertainty analysis. The uncertainty, both random and systematic, of the fragments' attribution to surface type (Kljun et al. 2004) likewise remains to be explored.

A valuable next step is to analyze these FOCAL data using the wavelet technique to develop a comparison between the two methods. The FFM has not yet used an interval quantity such as a vegetation index to characterize the source of the observed flux. Doing so will benefit from the upgrade of the footprint model to two dimensions (Kljun et al. 2015). The synergy between the
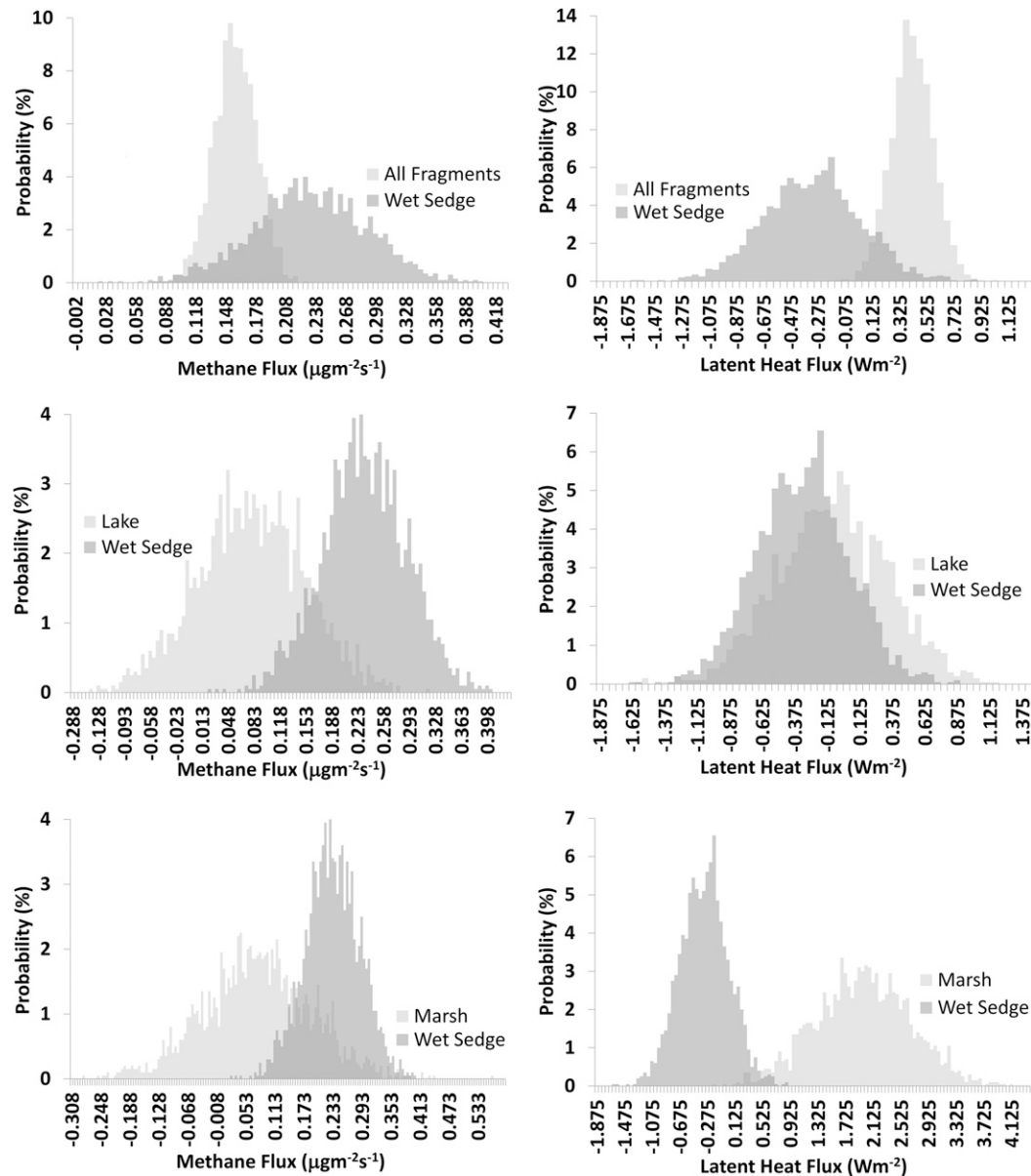
FIG. 12. Case 1900 LT 27 Aug 2013. As in Fig. 10, but the path included significant marsh instead of mesic sedge.

FFM, which encompasses features shorter than the averaging length, and the RFM, which treats the flux variation on longer scales, needs further development.

In the space–time domain, such as with the FFM and RFM, explicit treatment of the mesoscale and the turbulence should use a filter that clearly defines which components of the flow are to be treated deterministically and which stochastically. A filter such as that used by Foken et al. (2006) appears attractive. The first FOCAL campaign was focused on the air–surface exchange itself. The flights were designed to maximize the received signal from the surface and minimize mesoscale influences. The design specified low flight altitude over terrain as smooth as possible and having well-defined surface classes. The goal was to provide a spatial complement to individual surface sites that are rooted to one spot but can support a full set of observations of air, water, and soil.

The FFM/RFM is found to provide a viable scheme in the space–time domain for analysis of air–surface exchange, especially for airborne studies spatially constrained by population, terrain, or airspace limits. It is a development of the well-established space/time statistical treatment of turbulent flow. The statistical theory and practice of data-stream analysis provides a rich source for potential development and improvement.

REFERENCES

Avissar, R., and Coauthors, 2009: The Duke University helicopter observation platform. *Bull. Amer. Meteor. Soc.*, **90**, 939–954, doi:10.1175/2008BAMS2628.1.

Barnhart, B., W. Eichinger, and J. Prueger, 2012: A new eddy-covariance method using empirical mode decomposition. *Bound.-Layer Meteor.*, **145**, 369–382, doi:10.1007/s10546-012-9741-6.

Bögel, W., and R. Baumann, 1991: Test and calibration of the DLR Falcon wind measuring system by maneuvers. *J. Atmos. Oceanic Technol.*, **8**, 5–18, doi:10.1175/1520-0426(1991)008<0005:TACOTD>2.0.CO;2.

Brooks, S. B., T. L. Crawford, and W. C. Oechel, 1997: Measurement of carbon dioxide emissions plumes from Prudhoe Bay, Alaska oil fields. *J. Atmos. Chem.*, **27**, 197–207, doi:10.1023/A:1005890318796.

Brown, E. N., C. A. Friehe, and D. H. Lenschow, 1983: The use of pressure fluctuations on the nose of an aircraft for measuring air motion. *J. Climate Appl. Meteor.*, **22**, 171–180, doi:10.1175/1520-0450(1983)022<0171:TUOPFO>2.0.CO;2.

Crawford, T. L., and R. J. Dobosy, 1992: A sensitive fast-response probe to measure turbulence and heat flux from any airplane. *Bound.-Layer Meteor.*, **59**, 257–278, doi:10.1007/BF00119816.

——, ——, and E. J. Dumas, 1996: Aircraft wind measurement considering lift-induced upwash. *Bound.-Layer Meteor.*, **80**, 79–94, doi:10.1007/BF00119012.

Dobosy, R., and Coauthors, 2013: Calibration and quality assurance of an airborne turbulence probe in an aeronautical wind tunnel. *J. Atmos. Oceanic Technol.*, **30**, 182–196, doi:10.1175/JTECH-D-11-00206.1.

Dumas, E. J., and Coauthors, 2014: Airborne measurements of CO_2 and CH_4 fluxes over the Alaskan North Slope using the Flux Observations of Carbon from an Airborne Laboratory (FOCAL) system. NOAA Tech. Memo. ARL-267, 46 pp. [Available online at http://www.arl.noaa.gov/documents/reports/ARL%20TM-267.pdf.]

Efron, B., and R. Tibshirani, 1998: *An Introduction to the Bootstrap*. CRC Press, 436 pp.

Farge, M., 1992: Wavelet transforms and their applications to turbulence. *Annu. Rev. Fluid Mech.*, **24**, 395–458, doi:10.1146/annurev.fl.24.010192.002143.

Foken, T., F. Wimmer, M. Mauder, C. Thomas, and C. Liebethal, 2006: Some aspects of the energy balance closure problem. *Atmos. Chem. Phys.*, **6**, 4395–4402, doi:10.5194/acp-6-4395-2006.

Garman, K. E., and Coauthors, 2006: An airborne and wind tunnel evaluation of a wind turbulence measurement system for aircraft-based flux measurements. *J. Atmos. Oceanic Technol.*, **23**, 1696–1708, doi:10.1175/JTECH1940.1.

Gluhovsky, A., and E. Agee, 1994: A definitive approach to turbulence statistical studies in planetary boundary layers. *J. Atmos. Sci.*, **51**, 1682–1690, doi:10.1175/1520-0469(1994)051<1682:ADATTS>2.0.CO;2.

Holder, H. E., M. A. Bolch, and R. Avissar, 2011: Processing turbulence data collected on board the helicopter observation platform (HOP) with the empirical mode decomposition (EMD) method. *J. Atmos. Oceanic Technol.*, **28**, 671–683, doi:10.1175/2011JTECHA1410.1.

Kalogiros, J. A., and Q. Wang, 2002a: Aerodynamic effects on wind turbulence measurements with research aircraft. *J. Atmos. Oceanic Technol.*, **19**, 1567–1576, doi:10.1175/1520-0426(2002)019<1567:AEOWTM>2.0.CO;2.

——, and ——, 2002b: Calibration of a radome-differential GPS system on a Twin Otter research aircraft for turbulence measurements. *J. Atmos. Oceanic Technol.*, **19**, 159–171, doi:10.1175/1520-0426(2002)019<0159:COARDG>2.0.CO;2.

Karl, T., E. Apel, A. Hodzic, D. Riemer, D. Blake, and C. Wiedinmyer, 2009: Emissions of volatile organic compounds inferred from airborne flux measurements over a megacity. *Atmos. Chem. Phys.*, **9**, 271–285, doi:10.5194/acp-9-271-2009.

Kirby, S., R. Dobosy, D. Williamson, and E. Dumas, 2008: An aircraft-based analysis method for discerning individual fluxes in a heterogeneous agricultural landscape. *Agric. For. Meteor.*, **148**, 481–489, doi:10.1016/j.agrformet.2007.10.011.

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

——, ——, ——, and ——, 2015: A simple two-dimensional parameterisation for Flux Footprint Prediction (FFP). *Geosci. Model Dev.*, **8**, 3695–3713, doi:10.5194/gmd-8-3695-2015.

Leise, J. A., J. M. Masters, and R. J. Dobosy, 2013: Wind measurement from aircraft. NOAA Tech. Memo. OAR ARL-266, 209 pp., doi:10.7289/V5/TM-OAR-ARL-266.

LeMone, M. A., 1973: The structure and dynamics of horizontal roll vortices in the planetary boundary layer. *J. Atmos. Sci.*, **30**, 1077–1091, doi:10.1175/1520-0469(1973)030<1077:TSADOH>2.0.CO;2.

——, 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. *J. Hydrometeor.*, **4**, 179–195, doi:10.1175/1525-7541(2003)4<179:CTAIFC>2.0.CO;2.

Lenschow, D. H., J. Mann, and L. Kristensen, 1994: How long is long enough when measuring fluxes and other turbulence statistics? *J. Atmos. Oceanic Technol.*, **11**, 661–673, doi:10.1175/1520-0426(1994)011<0661:HLILEW>2.0.CO;2.

——, V. Savic-Jovcic, and B. Stevens, 2007: Divergence and vorticity from aircraft air motion measurements. *J. Atmos. Oceanic Technol.*, **24**, 2062–2072, doi:10.1175/2007JTECHA940.1.

Mahrt, L., 1998: Flux sampling errors for aircraft and towers. *J. Atmos. Oceanic Technol.*, **15**, 416–429, doi:10.1175/1520-0426(1998)015<0416:FSEFAA>2.0.CO;2.

——, 2000: Surface heterogeneity and vertical structure of the boundary layer. *Bound.-Layer Meteor.*, **96**, 33–62, doi:10.1023/A:1002482332477.

——, 2010: Computing turbulent fluxes near the surface: Needed improvements. *Agric. For. Meteor.*, **150**, 501–509, doi:10.1016/j.agrformet.2010.01.015.

Mallaun, C., A. Giez, and R. Baumann, 2015: Calibration of 3-D wind measurements on a single-engine research aircraft. *Atmos. Meas. Tech.*, **8**, 3177–3196, doi:10.5194/amt-8-3177-2015.

Mauder, M., R. L. Desjardins, and I. MacPherson, 2007: Scale analysis of airborne flux measurements over heterogeneous terrain in a boreal ecosystem. *J. Geophys. Res.*, **112**, D13112, doi:10.1029/2006JD008133.

——, ——, and ——, 2008: Creating surface flux maps from airborne measurements: Application to the Mackenzie Area

GEWEX Study MAGS 1999. *Bound.-Layer Meteor.*, **129**, 431–450, doi:10.1007/s10546-008-9326-6.

Metzger, S., W. Junkermann, M. Mauder, F. Beyrich, K. Butterbach-Bahl, H. Schmid, and T. Foken, 2012: Eddy-covariance flux measurements with a weight-shift microlight aircraft. *Atmos. Meas. Tech.*, **5**, 1699–1717, doi:10.5194/amt-5-1699-2012.

——, and Coauthors, 2013: Spatially explicit regionalization of airborne flux measurements using environmental response functions. *Biogeosciences*, **10**, 2193–2217, doi:10.5194/bg-10-2193-2013.

Monin, A., and A. Yaglom, 1971: *Statistical Fluid Mechanics*. Vol. 1. *Mechanics of Turbulence*, MIT Press, 782 pp.

Mudelsee, M., 2002: TAUEST: A computer program for estimating persistence in unevenly spaced weather/climate time series. *Comput. Geosci.*, **28**, 69–72, doi:10.1016/S0098-3004(01)00041-3.

——, 2010: *Climate Time Series Analysis*. Springer, 474 pp.

Oechel, W. C., G. L. Vourlitis, S. Brooks, T. L. Crawford, and E. Dumas, 1998: Intercomparison among chamber, tower, and aircraft net CO_2 and energy fluxes measured during the Arctic System Science Land-Atmosphere-Ice Interactions (ARCSS-LAII) Flux Study. *J. Geophys. Res.*, **103**, 28 993–29 003, doi:10.1029/1998JD200015.

——, and Coauthors, 2000: A scaling approach for quantifying the net CO_2 flux of the Kuparuk River Basin, Alaska. *Global Change Biol.*, **6**, 160–173, doi:10.1046/j.1365-2486.2000.06018.x.

Ogunjemiyo, S. O., P. Schuepp, J. MacPherson, and R. Desjardins, 1997: Analysis of flux maps versus surface characteristics from Twin Otter grid flights in BOREAS 1994. *J. Geophys. Res.*, **102**, 29 135–29 145, doi:10.1029/97JD01108.

——, S. K. Kaharabata, P. H. Schuepp, I. J. MacPherson, R. L. Desjardins, and D. A. Roberts, 2003: Methods of estimating CO_2, latent heat and sensible heat fluxes from estimates of land cover fractions in the flux footprint. *Agric. For. Meteor.*, **117**, 125–144, doi:10.1016/S0168-1923(03)00061-3.

Sayres, D., and Coauthors, 2017: Arctic regional methane fluxes by ecotope as derived using eddy covariance from a low flying aircraft. *Atmos. Chem. Phys.*, **17**, 8619–8633, doi:10.5194/acp-17-8619-2017.

Sellers, P. J., and Coauthors, 1997: BOREAS in 1997: Experiment overview, scientific results, and future directions. *J. Geophys. Res.*, **102**, 28 731–28 769, doi:10.1029/97JD03300.

Sun, J., and Coauthors, 1997: Lake-induced atmospheric circulations during BOREAS. *J. Geophys. Res.*, **102**, 29 155–29 166, doi:10.1029/97JD01114.

Torrence, C., and G. P. Compo, 1998: A practical guide to wavelet analysis. *Bull. Amer. Meteor. Soc.*, **79**, 61–78, doi:10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2.

van den Kroonenberg, A., T. Martin, M. Buschmann, J. Bange, and P. Vörsmann, 2008: Measuring the wind vector using the autonomous mini aerial vehicle M^2AV. *J. Atmos. Oceanic Technol.*, **25**, 1969–1982, doi:10.1175/2008JTECHA1114.1.

Vellinga, O. S., B. Gioli, J. A. Elbers, A. A. M. Holtslag, P. Kabat, and R. W. A. Hutjes, 2010: Regional carbon dioxide and energy fluxes from airborne observations using flight-path segmentation based on landscape characteristics. *Biogeosciences*, **7**, 1307–1321, doi:10.5194/bg-7-1307-2010.

——, 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. *J. Atmos. Oceanic Technol.*, **30**, 161–181, doi:10.1175/JTECH-D-11-00138.1.

Vickers, D., and L. Mahrt, 1997: Quality control and flux sampling problems for tower and aircraft data. *J. Atmos. Oceanic Technol.*, **14**, 512–526, doi:10.1175/1520-0426(1997)014<0512:QCAFSP>2.0.CO;2.

Webb, E. K., G. I. Pearman, and R. Leuning, 1980: Correction of flux measurements for density effects due to heat and water vapour transfer. *Quart. J. Roy. Meteor. Soc.*, **106**, 85–100, doi:10.1002/qj.49710644707.

Wieringa, J., 1986: Roughness-dependent geographical interpolation of surface wind speed averages. *Quart. J. Roy. Meteor. Soc.*, **112**, 867–889, doi:10.1002/qj.49711247316.

Witinski, M., D. Sayres, and J. Anderson, 2011: High precision methane isotopologue ratio measurements at ambient mixing ratios using integrated cavity output spectroscopy. *Appl. Phys.*, **102B**, 375–380, doi:10.1007/s00340-010-3957-2.

Zulueta, R. C., W. C. Oechel, H. W. Loescher, W. T. Lawrence, and K. T. Paw U, 2011: Aircraft-derived regional scale CO_2 fluxes from vegetated drained thaw-lake basins and interstitial tundra on the Arctic Coastal Plain of Alaska. *Global Change Biol.*, **17**, 2781–2802, doi:10.1111/j.1365-2486.2011.02433.x.