Atmospheric
Measurement
Techniques

Supplement of

The NASA Carbon Airborne Flux Experiment (CARAFE): instrumentation and methodology

Glenn M. Wolfe et al.

Correspondence to: Glenn M. Wolfe (glenn.m.wolfe@nasa.gov)

S1 PPFD corrections

The LI-COR LI-190R PPFD sensor is designed to be stationary and level; thus, corrections are required to account for sensor tilt under flight conditions. The attitude of the aircraft with respect to the Earth's surface is characterized by two angles: pitch (P) and roll (R). The sensor is not completely level with respect to the aircraft axes, necessitating additional terms to account for offsets (e_P and e_R). The sensor can also experience perturbations to the "direct sun" component, which depends on the aircraft heading (H, with potential offset e_H) and attitude as well as the solar zenith (θ) and azimuth (ϕ) angles. For example, if the sun lies to the south and the aircraft is heading north with a positive pitch (nose pointed upwards), the sensor will be pointed more towards direct sunlight than if it were fully level. The sensor can also experience shading under high roll or pitch conditions. These effects are evident in the raw flight data (Fig. S3).

These considerations give rise to a trigonometric empirical correction for varying Earth-sensor and sun-sensor angles. The correction takes the following form:

$$PPFD_{corr} = \frac{PPFD_{raw}\cos\theta}{\cos[\theta + (P + e_P)\cos(H + e_H - \phi) + (R + e_R)\sin(H + e_H - \phi)]} \tag{S1}$$

The offset angles ($e_P = 6.1°$, $e_R = 2.6°$, $e_H = 1.4°$) are estimated by optimizing the correlation between corrected PPFD and cosθ for two clear-sky flights. Fig. S3 demonstrates the quality of this correction. In addition, PPFD data are discarded when roll angles exceed 5° to minimize sensor shading artifacts.

S2 Data gap filling

The alteration of a time series to eliminate gaps can markedly perturb wavelet-derived fluxes. Linear interpolation is adequate for gaps that are small relative to the integral time scale (of order 1 s), but it is a poorer approximation for larger gaps. It is also possible to "stitch" a time series together (literally removing the gaps). This method can create scale-dependent artifacts; in the Fourier analogy, this is similar to removing some portion of a sine wave and thereby creating a discontinuity. Figure S8 shows example errors introduced through these gap-filling methods for an artificially-gapped segment of potential temperature flux observations. Errors are particularly pronounced near gap edges but can propagate throughout the whole leg. Note that this is only an example, and the magnitude and breadth of errors will depend on the width and location of a gap as well as the nature of the underlying observations.

To reduce potential artifacts near gap edges, we have developed an empirical method that we refer to as "covariance filling." Covariance filling exploits the fact that we have two time series, x and w, that are expected to co-vary. For illustration, let us assume that only x contains gaps. The basic steps are as follows:

1. For each gap of width N, define a region of width $3N$ centered on the gap.
2. Calculate the covariance within the region using all valid data pairs.

1

$$cov(w, x) = \overline{(w - \bar{w})(x - \bar{x})} \tag{S2}$$

Overbars indicate regional means that are restricted to times when both x and w are valid.

3. Assume the gap contains the same covariance as the region. Then the predicted standard deviation of x within the gap, σ_x, is calculated from the identity

$$cov(w, x) = r_{wx}\sigma_w\sigma_x \tag{S3}$$

Here, σ_w refers to the standard deviation of w within the gap. The correlation coefficient r_{wx} is assumed equal to one, as we will directly scale w below.

4. Scale w within the gap to create a predicted time series of x and use this to fill the gap.

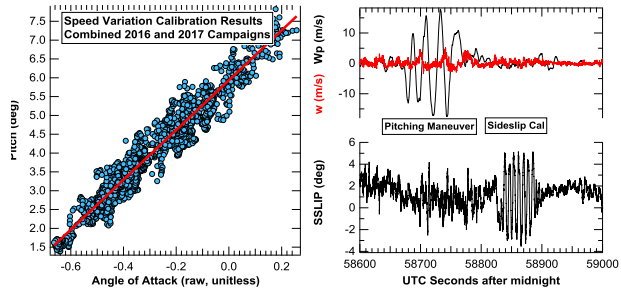$$x_g(t) = \frac{\sigma_x}{\sigma_w}\left(w_g(t) - \overline{w_g(t)}\right) \tag{S4}$$

As evident in Figure S8, covariance filling can reduce artifacts in derived fluxes compared to the interpolation and stitching methods. We find this to be true in general, based on a Monte Carlo experiment where 10,000 gaps of random size (1 to 600 points for 10 Hz data) and location were inserted into a potential temperature time series before wavelet flux calculation. Median maximum errors in this simulation are 30% for covariance filling, 45% for interpolation and 65% for stitching. Covariance filling does carry some limitations. Linear interpolation is used instead in the following cases: 1) a gap is small (less than 3 points), or 2) more than 50% of data pairs are missing in the region around a gap (due to adjacent gaps), or 3) overlapping gaps exist in both time series.

Table S1. Summary of flights for the 2016 and 2017 CARAFE missions.

Flight #	Flight Date	Flight Time (EDT)	Destination	Flux Legs	Tower Overpass[a]
1	7 Sep 2016	12:25 - 15:59	Pocomoke Forest (MD)	6	
2	9 Sep 2016	11:23 - 15:20	Pocomoke Forest (MD)	13	
3	12 Sep 2016	11:06 - 15:03	Choptank Ag (MD/DE)	9	USDA-Chop, US-StJ
4	14 Sep 2016	11:07 - 15:17	Pine Barrens (NJ)	12	US-Slt, US-Ced
5	16 Sep 2016	11:21 - 15:10	Pocomoke Forest (MD)	15	
6	22 Sep 2016	11:58 - 16:00	Prince Frederick (MD) Charles County (MD)	6 8	
7	23 Sep 2016	11:12 - 15:05	Pine Barrens (NJ)	12	US-Slt, US-Ced
8	24 Sep 2016	10:47 - 14:35	Great Dismal (VA/NC) Alligator River (NC)	7 6	US-NC4
9	26 Sep 2016	11:04 - 15:10	Atlantic Ocean Chesapeake Bay Pocomoke Forest (MD)	4 5 2	
1	3 May 2017	15:00 - 17:27	Pocomoke Forest (MD)	4	
2	4 May 2017	11:12 - 15:03	Choptank Ag (MD/DE) Pocomoke Forest (MD)	7 3	USDA-Chop, US-StJ
3	8 May 2017	10:42 - 12:23	Pocomoke Forest (MD)	7	
4	8 May 2017	14:02 - 17:40	Pocomoke Forest (MD)	16	
5	9 May 2017	11:01 - 15:05	Pine Barrens (NJ)	16	US-Slt, US-Ced
6	10 May 2017	11:56 - 14:32	Prince Frederick (MD) Charles County (MD)	7 8	
7	15 May 2017	12:59 - 15:49	Alligator River (NC)	6	US-NC4
8	16 May 2017	11:08 - 15:02	Great Dismal (VA/NC) Chesapeake Bay	4 6	
9	18 May 2017	11:05 - 15:05	Choptank Ag (MD/DE)	10	USDA-Chop, US-StJ
10	19 May 2017	10:04 - 13:29	Pocomoke Forest (MD)	13	
11	26 May 2017	08:51 - 12:55	Alligator River (NC)	11	US-NC4

[a]**Abbreviations: USDA-Chop = Choptank, MD; US-StJ = St. Jones Preserve, DE; US-Slt = Silas Little, NJ; US-Ced = Cedar Bridge, NJ; US-NC4 = Alligator River, NC. More information on towers available at the ameriflux.lbl.gov/ and https://www.ars.usda.gov/northeast-area/beltsville-md/beltsville-agricultural-research-center/hydrology-and-remote-sensing-laboratory/docs/research-sites/.**
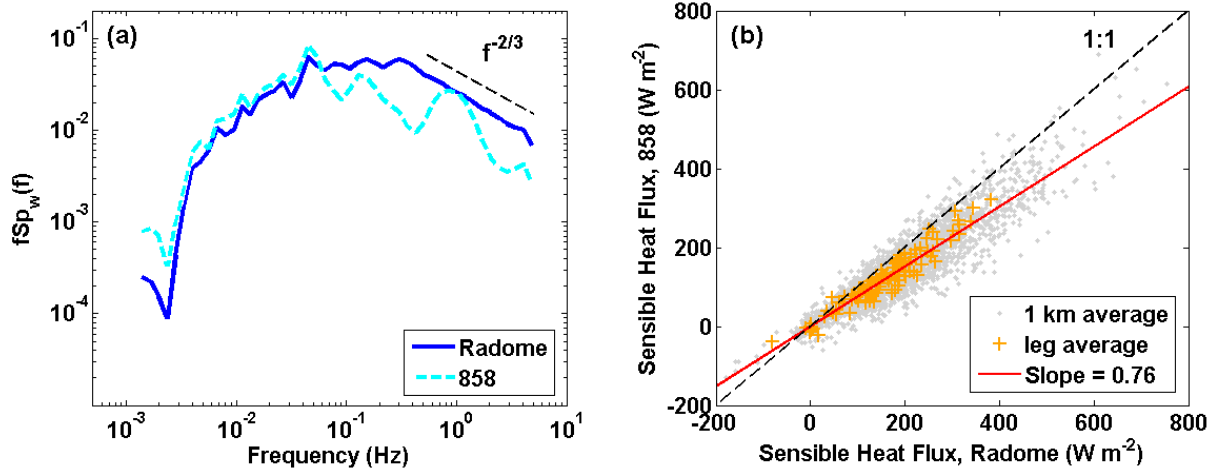
(b)

(c)

(a)

Fig S1. (a) Results for the angle of attack response during all speed variation maneuvers during both the 2016 and 2017 campaigns. (right) Derived vertical wind speed (red) measured during pitching maneuvers and sideslip calibrations showing that measured vertical wind speed is unaffected by changes in both pitch (b) and sideslip angle (c).

Figure S2. (a) Example frequency-multiplied vertical wind spectra for a single leg from the 2017 mission. Spectra are calculated from wavelet transforms for vertical wind data from the radome (solid blue) and 858 probe (dashed cyan) systems. The black dashed line denotes the -2/3 slope expected in the inertial subrange. (b) Comparison of sensible heat fluxes for all 2017 flights calculated using vertical winds from the radome or 858 probe. Gray dots represent 1 km average fluxes, orange crosses are averaged over whole legs. The red line represents the systematic error in 858 fluxes, based on a zero-intercept linear regression through all 2017 fluxes. The slope of this line is used to correct for the systematic bias in 2016 fluxes as described in Sect. 2.2. .
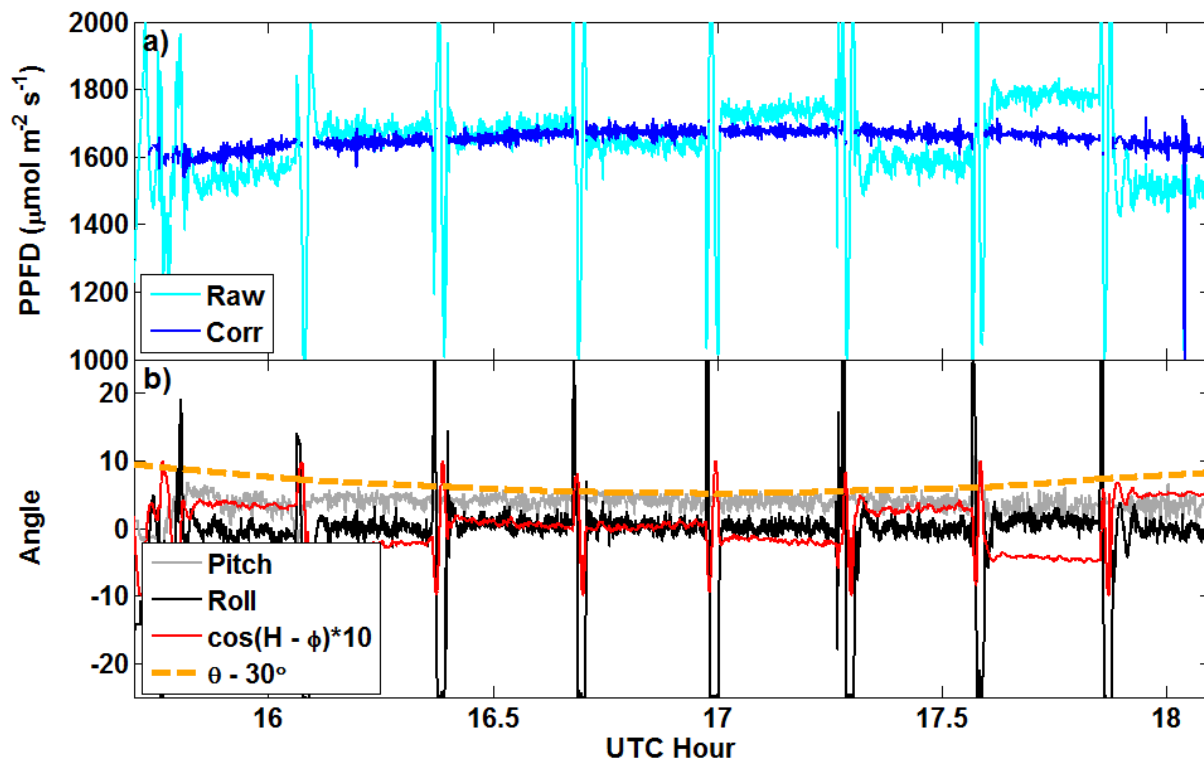
Figure S3. (a) 1 Hz PPFD observations from the 12 September 2016 flight under mostly clear-sky conditions. Raw data is shown in cyan, corrected/filtered data in blue. (b) Corresponding aircraft attitude (pitch: gray, roll: black) and solar zenith angle (orange dash) and cosine of difference of aircraft heading and solar azimuth (red). The latter is the relevant quantity for sun-sensor angle corrections.
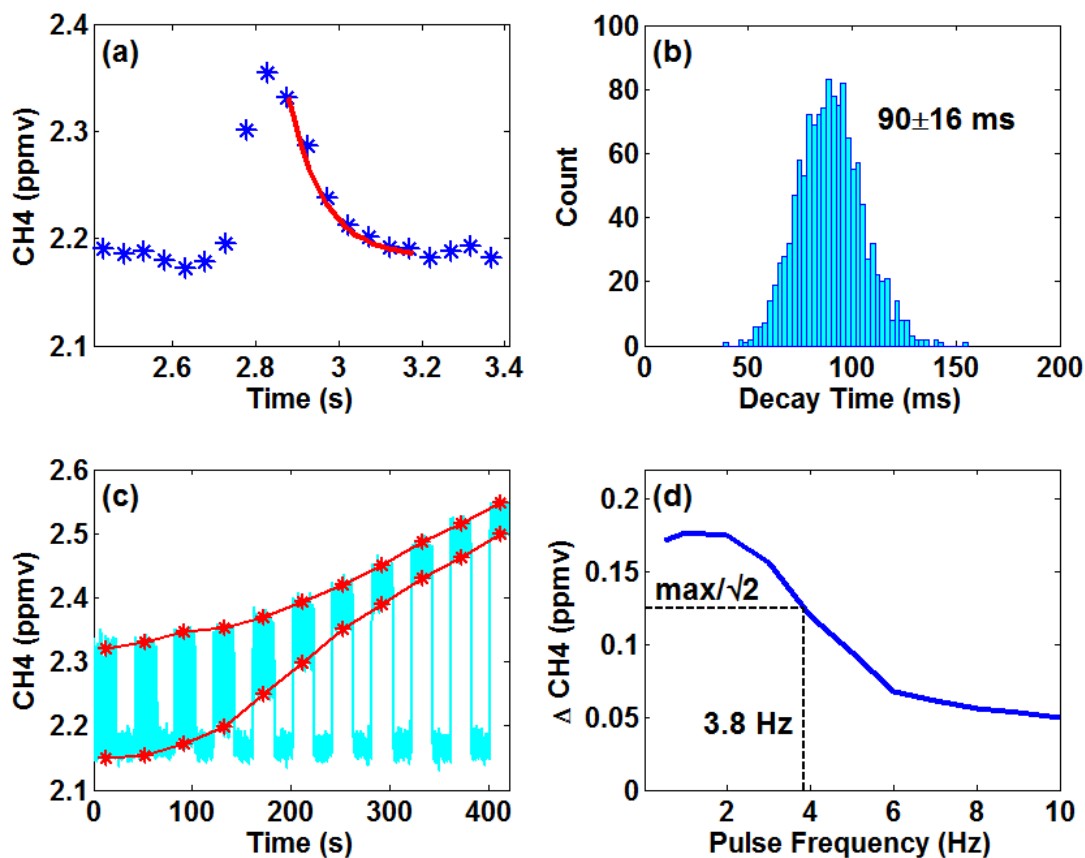
Figure S4. Laboratory tests to characterize the time response of the LGR/nXDS15i combination. In these experiments, a miniature solenoid valve (The Lee Company) was used to add 5 ms pulses of 100 ppmv CH_4 to room air while sampling at 20 Hz. Each pulse exhibits an exponential decay (a), and fitting the first 0.3 s of 1300 such decays gives a characteristic e-folding time of 90 ± 16 ms (b). We also performed a test analogous to that described in Aubinet et al. (2016), wherein the time between pulses is reduced (pulse frequency increased) and pulses "smear" together (c). The "cutoff frequency" is defined as the point where the difference between maximum and minimum concentrations decreases by $\sqrt{2}$ (d).
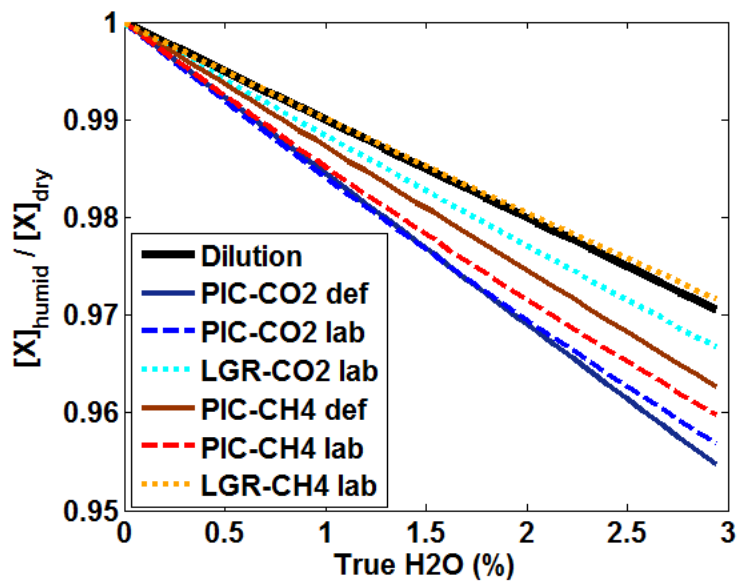
Figure S5. Empirical correction factors for conversion of ambient CO_2 (blues) and CH_4 (reds) observations from humid to dry mixing ratios. Correction factors are based on laboratory measurements of calibration gas with varying humidity levels controlled by a bubbler/dilution system. Correction factors inherently include both dilution and spectroscopic effects. Also shown for comparison is the correction factor for dilution only (solid black line) and the default correction factors for the Picarro G1301-m (solid blue and red lines). "True H2O" refers to the calibration-corrected Picarro water vapor mixing ratio.
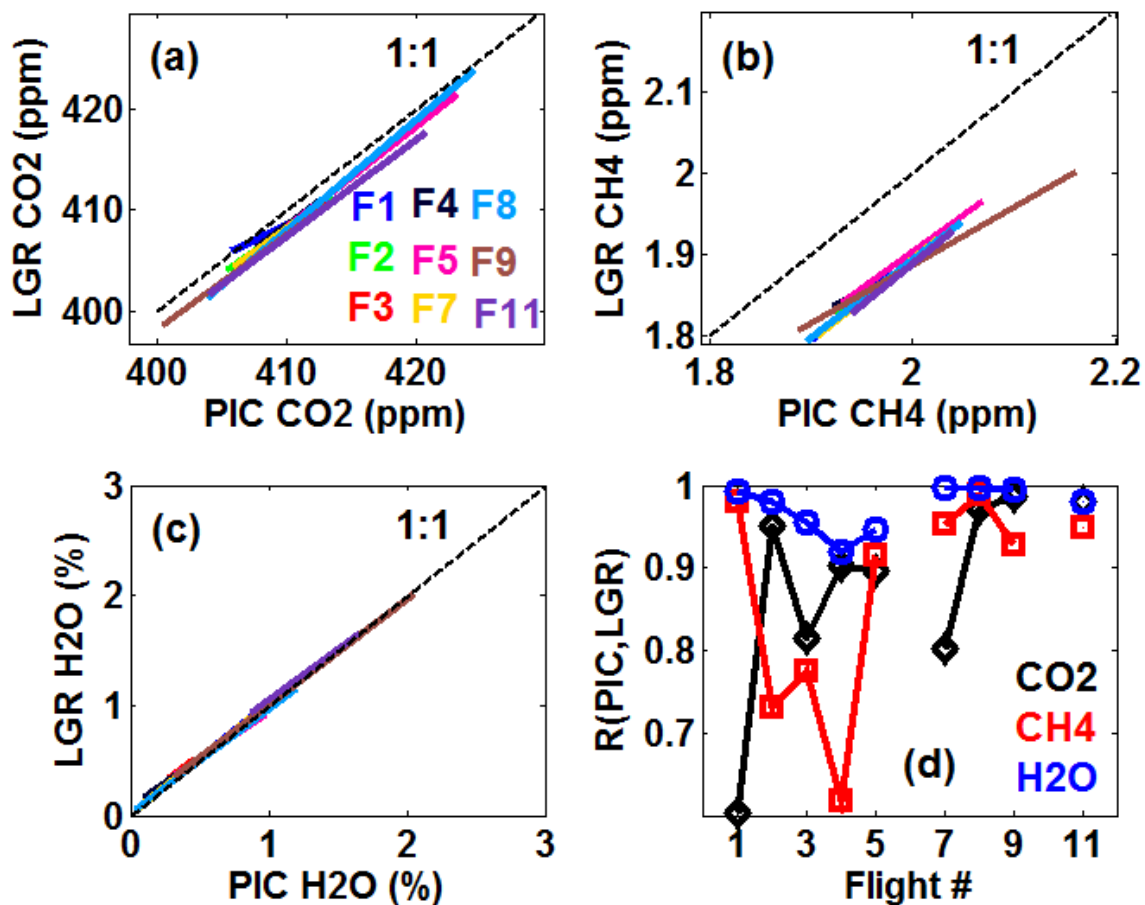
Figure S6. Least squares fit lines (a-c) and correlation coefficients (d) for Picarro and LGR dry mixing ratios of CO_2, CH_4 and H_2O obtained during flights in 2017. Fits are colored by flight and shown only over the range of mixing ratios observed on each flight. Dashed lines in a), b) and c) denote a 1:1 correlation. Fits are not available for flights 6 and 10 due to a malfunction of the Picarro instrument; in these cases, mission-averaged fit coefficients were used to correct the LGR data.
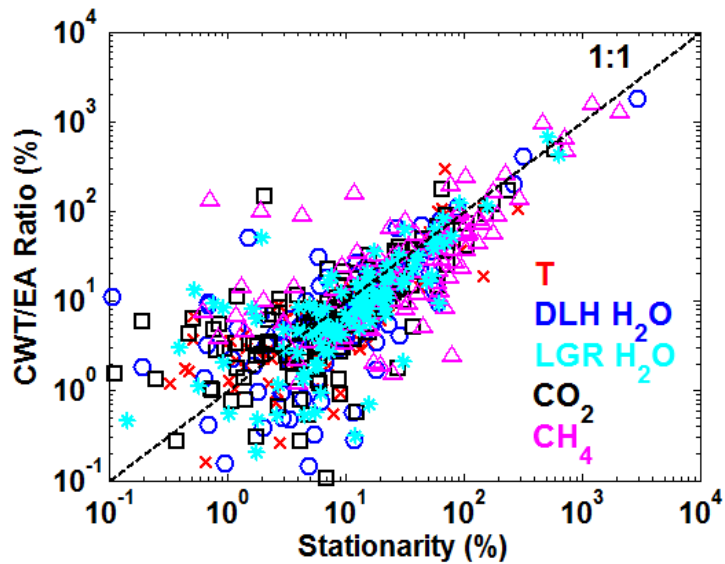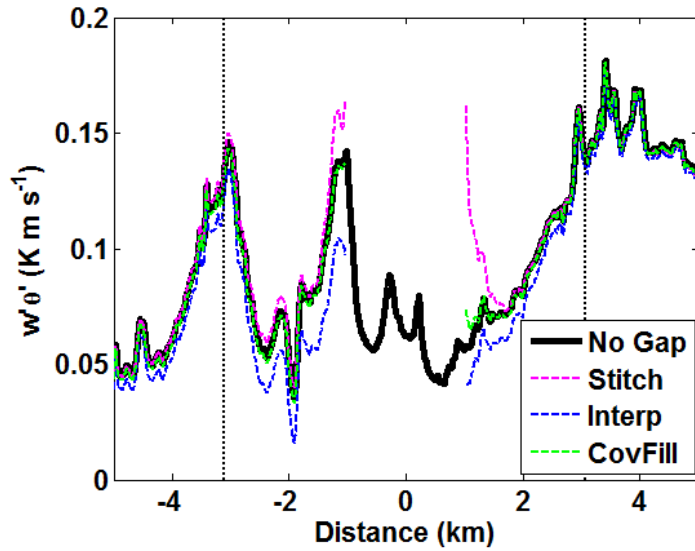
Figure S7. Comparison of stationarity and the ratio of CWT to EA fluxes for all scalars and flight legs of the 2017 mission. Stationarity is defined as described in the main text (Eq. (2)) but cast in percentage units. Analogously, the CWT/EA flux ratio is defined as $100 \times |1 - \langle F_{CWT} \rangle / F_{EA}|$.

Figure S8. Example of gap-filling effects on wavelet fluxes. The solid black line shows a portion of a sensible heat flux time series derived from wavelet transforms. A gap of random size (293 points) and location was inserted into the potential temperature time series, and wavelet fluxes were recomputed using one of three gap filling methods: stitching (magenta), linear interpolation (blue), or covariance filling (green). Perturbations near the gap edges are evident. Vertical dotted lines indicate the window within which fluxes would be discarded due to potential influence from gap filling. The window width is thrice the size of the gap and centered thereon.

5
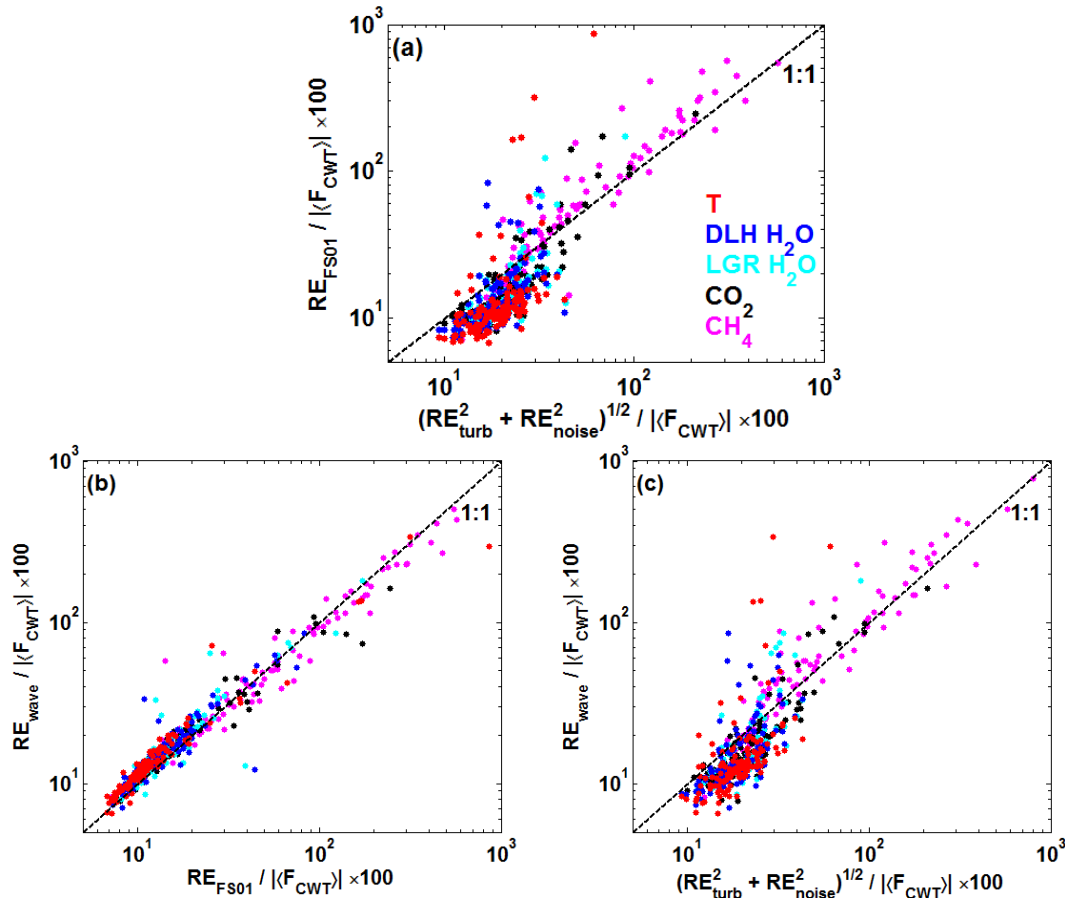
5   **Figure S9. Comparison of random error estimates: (a) ensemble empirical vs semi-theoretical, (b) wavelet-derived vs ensemble empirical, and (c) wavelet-derived vs semi-theoretical. Errors are shown for all legs and fluxes in the 2017 mission as percentage units relative to leg-averaged fluxes. Recall that RE_{turb} is a theoretical upper limit, hence we expect $\left(RE_{turb}^2 + RE_{noise}^2\right)^{1/2}$ to exceed RE_{FS01} and RE_{wave} in cases where turbulence is the dominant driver of total random error.**
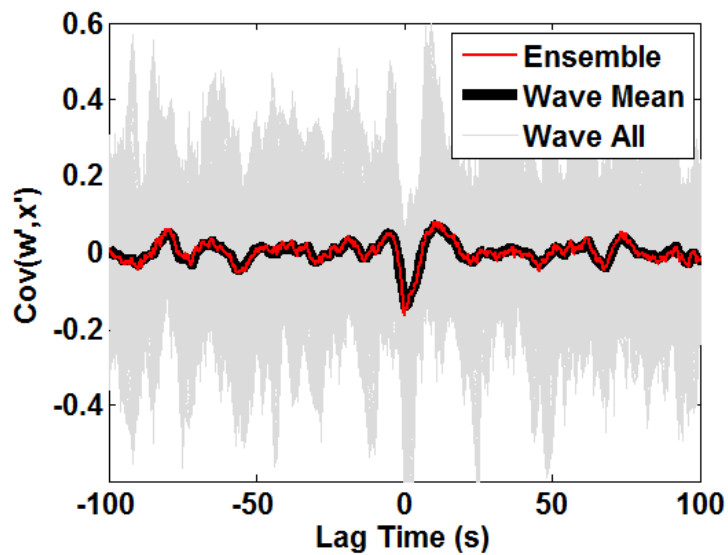
Figure S10. Example of wavelet lag-covariance calculation (Sect. 3.4.2) for a single leg of CO_2 fluxes. Gray lines represent 1 Hz wavelet cross-covariance functions of vertical wind and CO_2 for all 960 points in the time series. The mean of all wavelet cross-covariances (black) is nearly identical to the cross-covariance calculated on the ensemble dataset (red).
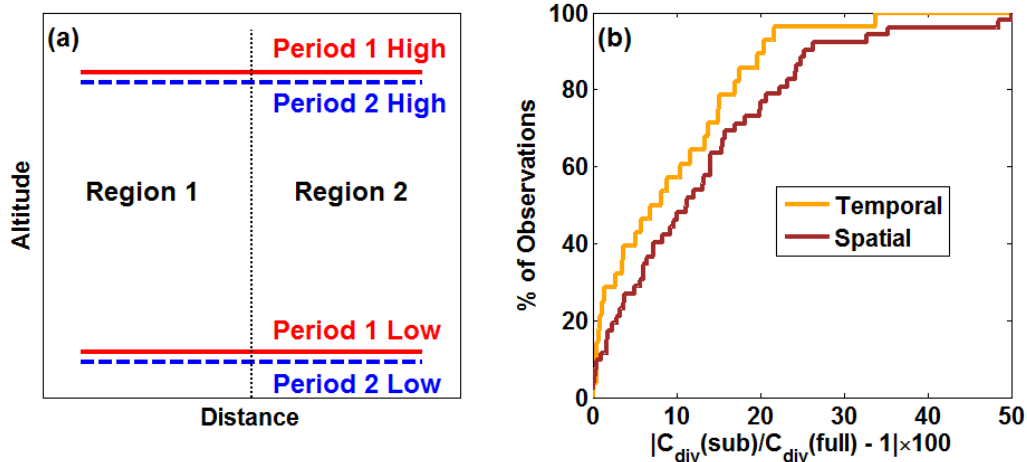
Figure S11. Illustration of tests for divergence correction variability in time and space. (a) Cartoon depicting the distribution of flux legs. For a given flight, two periods separated by 1.5 – 2.2 hours are chosen containing a high and low level leg over the same horizontal area. Fluxes of CO_2, temperature and H_2O (DLH and LGR) are used for these tests. For the temporal variability test, divergence corrections (C_{div}) for period 1 and period 2 ("sub-periods") are compared to C_{div} calculated using all four legs ("full"). For the spatial variability test, C_{div} for regions 1 and 2 ("sub-regions") are compared to C_{div} calculated using fluxes from both regions ("full"). The latter test is done separately for each period. **(b)** cumulative distribution of percent deviation of sub-period and sub-region C_{div} values from their "full" counterparts (essentially a measure of deviation relative to an "average" divergence correction).