Contents

List of Figures

List of Tables

S1 OTM 33A Analysis Program

S1.1 Wind and Mixing Ratio Binning

The OTM 33A analysis program, written in MATLAB (2015), bins methane enhancements by 10° wind bins in order to generate a Gaussian distribution and a peak concentration. To do this, the average value of the lowest 5% of observed methane mixing ratios are subtracted from the methane mixing ratio time series to create a series of methane enhancements above background. Then, the methane enhancements are binned by 10° wind bins to find the total enhancement for that wind direction. The number of data points that occur in each wind bin are also counted during this time. Bins that have too few measurements (or values below zero) are excluded from further analysis. The total methane enhancement counted in each bin is divided by the number of data points in each wind bin (Figure 2(a)) to yield the methane enhancement per 10° wind bin (Figure 2(b)). The R^2 value of the Gaussian fit is calculated as the sum of squares of the residuals divided by the sum of squares of the data following Eq. S1. Plot created using IGOR (2018).

$$R^2 = \left(\frac{\Sigma(model\, y - mean(y))^2}{\Sigma(y - mean(y))^2} \right) \tag{S1}$$

S1.2 OTM 33A Data Quality Indicators

The OTM 33A analysis program includes many data quality indicators (DQI) or "flags" that help discern if an OTM 33A flux estimate is a valid approximation. Category 1 measurements have fewer than 5 flags, Category 2 measurements have fewer than 10 flags, and Category 3 measurements have 10 or more flags. Even with these analysis flags, in-field observations of possible measurement issues and manual checks of the Gaussian fit are required.

DQI NAME	LIST OF EXCEEDANCE LEVELS	FLAG POINT VALUE
COUNT	< 15 minutes of measurement	3
WIND DIRECTION	$(\theta_y + \theta_{vertical\,wind})/2 > 30$	1
	$\frac{(\theta_y - \theta_{vertical\,wind})}{(\theta_y + \theta_{vertical\,wind})/2} > 0.5$	1
TURBULENT INTENSITY	TI > 0.22	1
WIND SPEED	$\overline{U} < 1.5$ m/s	1
	$\overline{U} < 1$ m/s	5
WIND VARIANCE	Wind Variance > 2.5	5
	Wind Variance > 5	10
BINNED DATA	Highest Concentration Bin ≠ 180 ± 30˚	1
	Highest Concentration Bin ≠ 180 ± 60˚	3
GAUSSIAN FIT	R < 0.95	5
	R < 0.90	10
METHANE LEVEL	Background Methane < 1.7 ppm	1
	CH4 Enhancement < 0.15 ppm & Dist. < 50 m	10
	CH4 Enhancement < 0.15 ppm & Dist. 50 – 100 m	5
	CH4 Enhancement < 0.15 ppm & Dist. 100 – 150 m	3
	CH4 Enhancement < 0.15 ppm & Dist. > 150 m	1
	CH4 Enhancement < 0.10 ppm & Dist. < 50 m	10
	CH4 Enhancement < 0.10 ppm & Dist. 50 – 100 m	5
	CH4 Enhancement < 0.10 ppm & Dist. 100 – 150 m	3
	CH4 Enhancement < 0.10 ppm & Dist. > 150 m	1

Figure S1. Table and caption after Snare (2015) Table 2.2.7. DQI categories and values for analysis flags. Variables include the standard deviation of the 2D wind speed Θ, mean wind speed ($\bar{U}$), turbulent intensity (TI), correlation coefficient (R), source distance (Dist.) and methane enhancement.

S1.3 Accepted OTM 33A Flux Measurements

The accepted OTM 33A flux measurements and associated measurement variables are included as an Excel file.

S2 Ordinary Least Squares Regression

Residuals and leverage analysis for the OLS regression are included below. The heteroskedasticity of the residuals for both fits and point leverages well below one further support the validity of an OLS model for the test release data.

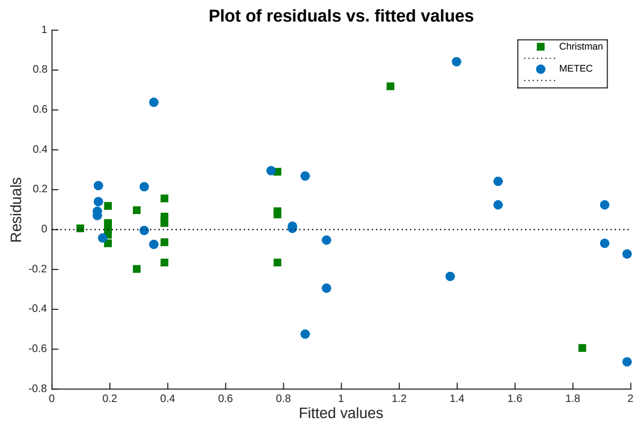

Figure S2. Residuals of ordinary least square regression with intercept set to (0,0) for both test release experiments.

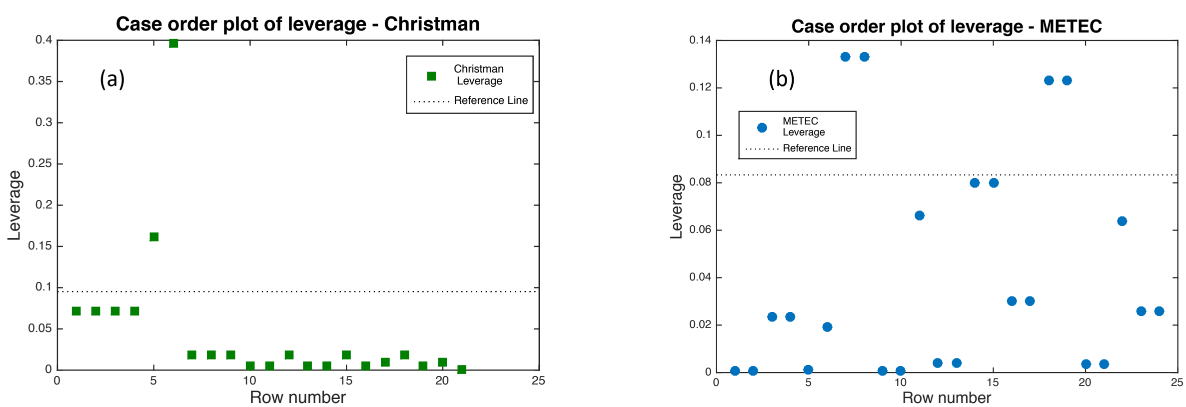

Figure S3. Leverage analysis of ordinary least squares regression models for CF-TR (a) and METEC-TR (b).

S3 Percent Error - Grouped Measurements

When the combined test release dataset (N=45) is fit to a Gaussian, the 1σ error is $\pm 34.5\%$. The Gaussian fit suggests a low bias of -15%. The goodness of fit parameter R is calculated using Equation S1.

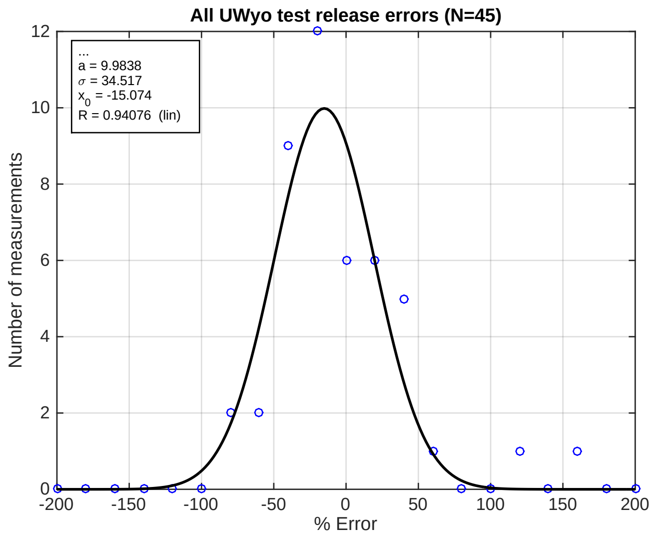

Figure S4. Percent errors of both METEC-TR and CF-TR binned by 20% error bins. The 2σ error is $\pm 69\%$.

S4 Bland-Altman Analysis

The Bland-Altman analysis used in Section 4.3 of the main text requires that the method difference (known release - OTM 33A flux) follow a normal distribution. Normality was determined following output from a normality test package developed by Öner et al.. The results for normality are summarized below.

Test Release	Test Name	Test Statistic	p-value	Normality (1:Normal,0:Not Normal)
CF-TR	KS Limiting Form	0.8551	0.4576	1
CF-TR	KS Stephens Modification	0.8878	0.0547	1
CF-TR	KS Marsaglia Method	0.8551	0.4075	1
CF-TR	KS Lilliefors Modification	0.1866	0.0545	1
METEC-TR	KS Limiting Form	0.6719	0.7574	1
METEC-TR	KS Stephens Modification	0.6943	0.1500	1
METEC-TR	KS Marsaglia Method	0.6719	0.7071	1
METEC-TR	KS Lilliefors Modification	0.1371	0.2000	1

Table S1. Results from six normality tests for the Bland-Altman Analysis.

S5 Probability Density Functions for Bootstrapped Data

Box plots of bootstrapped PDF for the METEC and CF test releases.

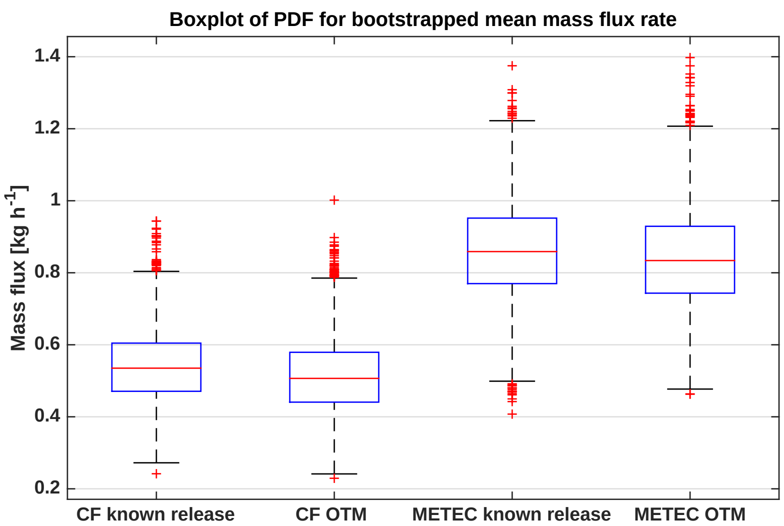

Figure S5. Box plot of PDFs from Section 4.3 of bootstrapped mean mass emission flux for OTM measurements and the known tracer release. The median is marked by the red line, 25 and 75th percentiles are represented by the box edges. Whiskers extend to twice the 90th percentile, and outliers are red markers.

S6 OLS fitting of Arkansas data

Data from the Bell et al. (2017) study which compared onsite flux estimates (SOE) and OTM 33A flux estimates for the same facilities using an ordinary least squares regression in addition to the variance weighted least squares model. The VWLS (slope = 0.41) is slightly greater than that of OLS (slope = 0.39). The lower CI of both regressions overlap and are identical if rounded to 2 significant figures (0.24). The upper CI of VWLS (0.92) is closer to 1:1 than upper CI of OLS (0.78).

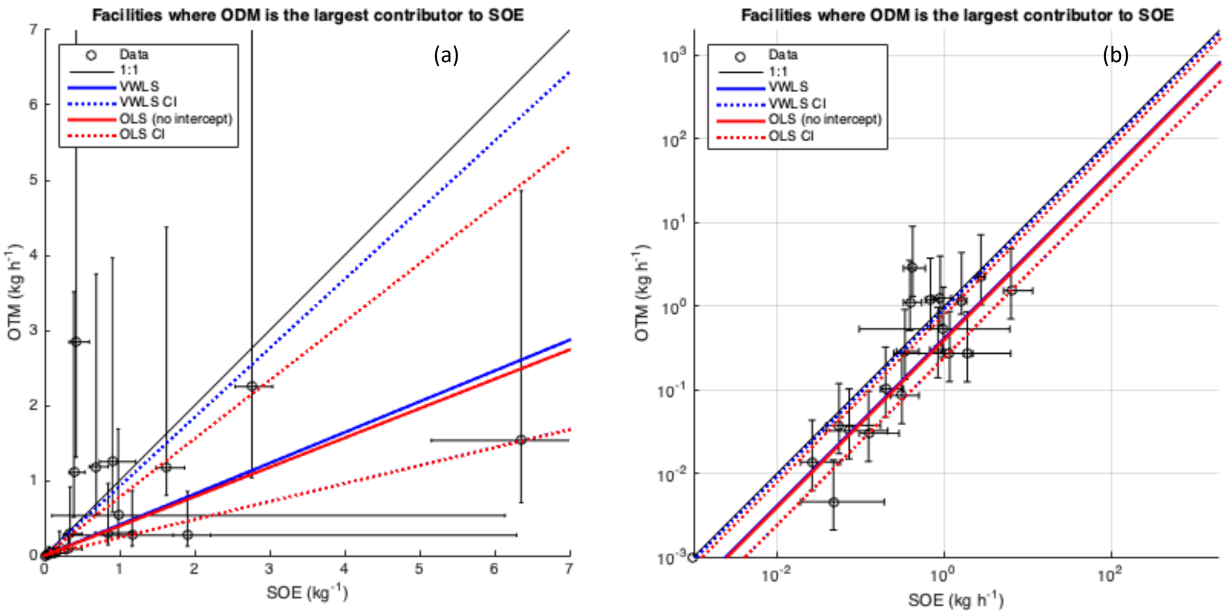

Figure S6. Correlation plots of OTM 33A measurements and onsite measurements (SOE) for 20 sites. (a) OLS and VLWS fits, linear scale. (b) OLS and VWLS fits, log scale (Bell et al., 2017).

References

Bell, C., Vaughn, T., Zimmerle, D., Herndon, S., Yacovitch, T., Heath, G., Pétron, G., Edie, R., Field, R., Murphy, S., Robertson, A., and Soltis, J.: Comparison of methane emission estimates from multiple measurement techniques at natural gas production pads, Elementa: Science of the Anthropocene, 5, 1–14, 2017.

IGOR: Version 7.08, WaveMetrics Inc., Lake Oswego, Oregon, 2018.

MATLAB: version 8.5.0 (R2015a), The MathWorks Inc., Natick, Massachusetts, 2015.

Öner, M., Applied, İ. D. K. J. o. M., and 2017: Jmasm 49: A compilation of some popular goodness of fit tests for normal distribution: Their algorithms and matlab codes (matlab), digitalcommons.wayne.edu.

Snare, D. A.: A comparison of ground-based and aircraft-based methane emission flux estimates in a western oil and natural gas production basin, M.S. Thesis, University of Wyoming, Laramie, Wyoming, United States of America, 2015.