Supplement of Atmos. Meas. Tech., 16, 5771–5785, 2023
https://doi.org/10.5194/amt-16-5771-2023-supplement

Atmospheric Measurement Techniques
Open Access
EGU

Supplement of

Methane point source quantification using MethaneAIR: a new airborne imaging spectrometer

Apisada Chulakadabba et al.

Correspondence to: Apisada Chulakadabba (achulakadabba@seas.harvard.edu)

S1 Additional Information

S1.1 MethaneAIR and MethaneSAT comparison

Table S1: Specification comparison for MethaneAIR and MethaneSAT

Specification	**MethaneAIR**	**MethaneSAT**
O_2 passband (nm)	1237–1319	1249–1305
O_2 dispersion (nm/pixel)	0.08	0.06
O_2 spectral FWHM (nm)	0.23	0.18
CH_4 passband (nm)	1592–1678*	1598–1683
CH_4 dispersion (nm/pixel)	0.10	0.08
CH_4 spectral FWHM (nm)	0.28	0.24
Field of view (°)	23.7	21.3
Cross-track pixel† (m)	~ 5 at 12 km	~ 108
Along-track pixel† (m)	~ 25	~ 400
Point spread function (pixels)	2.5	1.8
Single pixel SNR‡	~ 95	~ 190

*MethaneAIR uses an InGaAs detector with reduced QE beyond ~ 1660 nm
†Distance between pixel centers at nadir
‡CH_4 band with nominal radiance of 1.4×10^{13} photons cm^{-2} nm^{-1} sec^{-1} sr^{-1}

S2 Calculations

S2.1 mIME Effective Wind Speeds

The effective wind speeds used in the mIME calculation came from the relationship proposed by Varon et al. (2018), $U_{eff} = \alpha \cdot log(10\,m\,wind) + 0.6$ where α is between 0.9-1.1. We took advantage of the LES runs by introducing the LES-specific effective wind $U_{adaptive,\,eff}$ for each time step. $U_{adaptive,\,eff}$ is defined as $\frac{Q \cdot L}{IME}$. Because we know all the terms in this equation from the LES, we can calculate the true $U_{adaptive,\,eff}$ that gives the best estimate of Q.

By calculating the $U_{adaptive,\,eff}$ for all the time steps of interest, we end up with multiple pairs of $U_{adaptive,\,eff}$ and U, which can be fitted into a linear regression function of the form $U_{eff} = a \cdot log(10\,m\,wind) + b$ with unique a and b coefficients.

The LES-specific relationship can introduce overfitting results when the data points are limited, and wind speeds are low. If we overfit the relationship, we cannot justify the predicted effective wind speeds and may come up with biased emission estimates. Since the wind speeds during the controlled release experiments were lower than 5 m/s for most of the time, we decided to use the coefficients in (Varon et al., 2018) for the sake of simplicity.

Though the coefficients (Varon et al., 2018) derived from idealized LES with varying meteorological conditions are the most comprehensive IME coefficients available, we should be aware of ambient weather and turbulence conditions that may affect the relationship's universality.

S2.2 Divergence Integral

The divergence integral method is based on the integral form of the continuity equation, which states that the net production of CH_4 inside an enclosing surface, $\int_V P$, through an enclosing surface is given by

$$\int_V P = \int_V \vec{\nabla} \cdot c\vec{v} + \int_V \frac{dm}{dt} = 0 \tag{S1}$$

where c is the methane concentration, v is the wind speed, and $\langle \frac{\partial m}{\partial t} \rangle$ is the change in mass of methane with time within volume V. $\int_V P = 0$ since there is no *in situ* production. The flux divergence can be decomposed into two terms:

$$\nabla \cdot cv = v \cdot \nabla c + c \cdot \nabla v \tag{S2}$$

Conley et al. (2017) demonstrate that for wind speeds over 1 m/s and fluxes above 200 kg/hr, the wind divergence term (2nd term in Eq. S2) near a source is 1% or less of the concentration gradient term (1st term in Eq. S2); therefore, we can neglect the 2nd term for our application. To isolate the XCH_4 enhancement relative

to the surroundings, we subtract the mean column XCH_4 along the rectangle from the divergence term, which does not alter ∇c, because $\nabla\langle c\rangle = 0$

$$c' = c - \langle c\rangle \tag{S3}$$

$$\nabla c = \nabla(\langle c\rangle + c') = \nabla c' \tag{S4}$$

where c' is the XCH_4 enhancement and $\langle c\rangle$ is the mean XCH_4 around the rectangle. The c' values in Eq. S3 equation can be both positive and negative. When there is a positive flux through the surface, the positive c' values outweigh the negative c' values. There are some instances where the flux through the surface is negative (i.e., Figure 2d, 420m), and those values are averaged into the total flux as well. In most cases, these reflect the influence of boundary layer eddies.

For analysis of the total column data from MethaneAIR in the vicinity of an emission source, we assume that methane concentrations above the planetary boundary layer are equal to unperturbed values outside the bounding surface. Then, the horizontal gradients of XCH_4 arise entirely in the boundary layer, and we may take v as the pressure-weighted mean boundary layer wind speed and direction.

We then use Gauss's theorem to relate the volume integral to a surface integral around the cuboid:

$$\begin{aligned} 0 &= \int_V v\cdot\nabla c' dV + \langle\frac{\partial m}{\partial t}\rangle \\ &= -\Phi_{surf} + \oint_{\partial V=S} v\cdot\hat{n}c' dS + \langle\frac{\partial m}{\partial t}\rangle \\ &= -\Phi_{surf} + \oint_{\partial V=S} v\cdot\hat{n}((XCH_4 - \langle XCH_4\rangle_{rect.}) * n_{column} * M_{CH4})dS + \langle\frac{\partial m}{\partial t}\rangle \\ \Phi_{surf} &= \oint_{\partial V=S} v\cdot\hat{n}((XCH_4 - \langle XCH_4\rangle_{rect.}) * n_{column} * M_{CH4})dS \end{aligned} \tag{S5}$$

where Φ_{surf} is the flux into the volume from the land surface, S is the surface enclosing the cuboid, $\hat{n}$ is an outward pointing unit vector normal to the surface, n_{column} is the moles of air in the column based on the surface pressure from HRRR and M_{CH4} is the molar mass of methane. For a volume the size of a single plume, the term $\langle\frac{\partial m}{\partial t}\rangle$ should be insignificant, assuming that the plume is in a steady state and no mass is building up in the volume. We assume that the vertical flux through the top of the cuboid at the planetary boundary layer (PBL) height is zero, consistent with neglecting the divergence of the horizontal wind. Thus, the horizontal flux through the surface is balanced by and therefore equal to the flux through the bottom of the cuboid from the surface source, Φ_{surf} (kg/hr), we wish to measure. Based on Eq. S5, the flux from the source is computed from observed XCH_4 as follows:

$$\Phi_{surf} = \sum^{\text{around rect}} (XCH_{4_i} - \langle XCH_4\rangle_{rect.})\cdot n_{\text{column}}\cdot M_{CH_4}\cdot v\cdot\hat{n}\cdot\Delta l \tag{S6}$$

where Δl is the distance between successive XCH_4 measurements. We average the fluxes calculated over a sequence of rectangles extending downwind of the source, spanning several eddy scales. The box grows by 1 pixel in each step except for the upwind direction, where it moves over by 1/4 pixel. The upwind boundary was selected to be close to the upwind of the source using the moment of inertia of the plume (see Fig. 2)

S2.3 The Ratio Method

We developed a new method based on WRF-LES-HRRR outputs, assuming that the LES plumes represented the actual plumes in both areas and concentration distributions within the plume. At a given location and time, the probabilistic distribution of methane plumes should be similar at any emission rate. The emission rates should linearly depend on the magnitude of concentrations, which can be written as

$$\frac{Q_{unknown}}{Q_{known}} = \frac{IME_{unknown}}{IME_{known}}.$$

IME_{know} refers to the values from the LES simulations with a nominal emission rate (usually 988 kg/hr). Thus, the emission rates of the observed plumes are a function of the nominal emission rate of the simulated plumes multiplied by the ratio of the total masses of the observed plumes to the simulated plumes. The relationship can be expressed as

$$Q_{unknown} = Q_{known}\cdot\frac{IME_{unknown}}{IME_{known}}. \tag{S7}$$

This ratio method was first used in (Irakulis-Loitxate et al., 2021). It is applicable to very large sources with plumes large enough to be well represented by the 111 m × 111 m resolution. This is the limitation of the method that it only works when simulated winds accurately represent the actual winds.

Since LES cannot capture the shape of a snapshot of the observed plume due to large stochastic variations of the atmosphere, we use a range of snaps across time to represent the observed plumes for these two LES methods. We took the standard deviation of the emission rates calculated from 7 snapshots across a 30 minutes interval (15 minutes before and 15 minutes after the retrieval), 5 minutes apart from one another, to avoid autocorrelation effects. For the MethaneAIR data from RF04 and RF05 flights, only a few plumes were large enough to be assessed with the ratio and this method is not used in our post-unblinding analysis.

S2.4 Gridding by Nearest Neighbor and Physics-Based Methods

The nearest neighbor method for gridding used the R function binMean from the oce package to place the georeferenced XCH_4 data onto a predefined 10m × 10m grid. The retrieved XCH_4 values have a standard deviation of about 80 ppb. Since the the point spread function is about 12 m wide across the track, and 35 m along track, the grid of native pixels spatially oversamples the image on the ground. Therefore, we spatially filtered the retrieved XCH_4 data to create the gridded product in two steps. First, the holes in the initial gridded field were filled using a 3x3 averaging matrix. Then we further spatially filtered this field by a Gaussian filter with full width at half maximum of 1.7 pixels before inputting the XCH_4 values to our point source quantification methods. The resulting oversampled gridded product has a standard deviation of 40 ppb and a spatial correlation length of about 20 m, comparable to the length scale of the point spread function. The nearest-neighbor gridding method was used for the initial (blinded) input from Harvard to Stanford.

The improved estimates after unblinding used the "physics-based" gridding method of Sun et al. (2018).

S2.5 York Regression

When errors in the observables are present on both the x-axis and the y-axis, it is recommended to use York regression and other equations derived from York to determine the slope, y-intercept, and errors of the least-square estimation solutions. The full equations and derivations can be found on York (1968) and York et al. (2004)

S2.6 Mean Percentage Error

$$MPE = \frac{100\%}{n} \sum_{t=1}^{n} \frac{a_t - f_t}{a_t} \tag{S8}$$

where a_t is the actual value of the quantity being forecast, f_t is the forecast, and n is the number of different times for which the variable is forecast.

S3 Additional Results

S3.1 Blinded Volume-Controlled Release Experiments

The release rate was measured by a Sierra 640i QuadraTherm meter. Stanford performed an extensive set of meter intercomparison tests to characterize uncertainty in the QuadraTherm 640i meters (Rutherford et al. (2023)). Based on these intercomparison tests, a bias of up to 6% can be observed between QuadraTherm 640i meters. Error in metered controlled release volumes is characterized using a Monte Carlo analysis that also accounts for uncertainty in gas composition.

On seven occasions during the missions, the Stanford team informed the MethaneAIR team of the wind direction at the site since winds were light and variable, as reported by nearby ASOS systems. The MethaneAIR team considered reorienting the flight track to capture the plumes due to the wind conditions, but no changes in the flight track were required. On two occasions, Stanford confirmed that local winds were not zero. Although null winds were not desired for this first test of MethaneAIR, since the winds were not zero, the decision was made to complete the surveys as planned.

For the analysis, the MethaneAIR team used the York fit, which accounts for uncertainties in both the X and the Y coordinates, instead of the commonly used OLS, which assumes that uncertainties are significant only in the Y coordinate. In this experiment, the uncertainties in the metered flow rates are not negligible, and the error estimates are available from the Stanford team. Hence, the York fit was selected to approximate the desired Maximum Likelihood Estimator for the relationship between the Harvard values and the Stanford emission rates.

The MethaneAIR team conducted tests on the OLS to verify whether the OLS is an appropriate regression choice. The slope of the OLS line relating the Harvard mIME emissions to the Stanford single-blind volume-controlled release rates is 0.78 for the original submission. Based on the heteroskedasticity test, this OLS fit violates the assumptions for linear regression (p-value 0.006). The violation indicates a significant relationship between the residuals and the predictor (Stanford emission rates). The Non-Constant Error Variance test (ncvTest) gave a p-value of 0.14, suggesting violation at the 84% level. These tests support the application of the York fit for the 2021 results.

Based on the 2021 experiment (N = 21), the false positive rate is $\approx$ 10% and the false negative is $\approx$ 5%. These rates can not be generalized to the instrument capability since we have a limited number of data points, and the false positive/negative rates depend on the scene of interest.

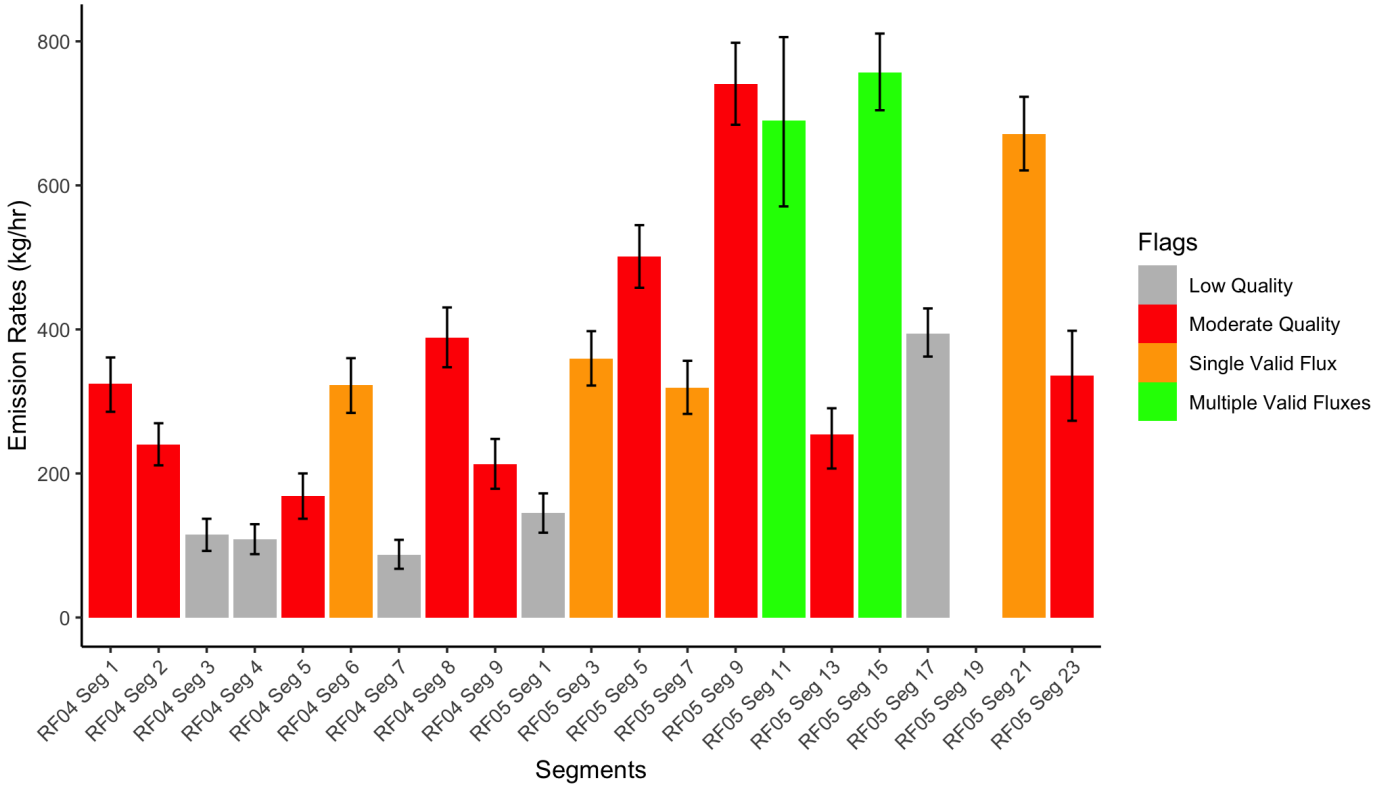


Figure S1: Post-unblinding best estimates of methane emissions from the controlled release experiments during RF04 and RF05 using the mIME method. The endpoints are tagged and color-coded according to the decision tree in Figure S11.

Table S2: RF04 and RF05 combined best estimate results from the single-blind volume-controlled release experiments (32.053 °N, 102.301 °W) and MethaneAIR Level 3 Gaussian filtered data on 30 July 2021 (RF04) and 3 August 2021 (RF05). The times are in UTC. The emission rates, lower bounds, and upper bounds are in kg/hr. The blinded, blinded lower, and blinded upper columns were reported to the Stanford team as MethaneAIR initial emission estimates on 1 February 2022. The unblinded, unblinded lower, and unblinded upper columns were added after the unblinding process. The upper and lower bound estimates were calculated using the method described in Section 2.5.

Seg	Timestamp (UTC)	Blinded	Blinded Lower	Blinded Upper	Unblinded	Unblinded Lower	Unblinded Upper	Flagged
1	30/7/21 15:41	321.74	179.30	471.71	324.55	285.64	361.07	No
2	30/7/21 15:52	232.19	133.63	338.38	240.14	211.28	269.73	No
3	30/7/21 16:07	109.22	62.73	160.03	114.76	92.43	137.06	Yes
4	30/7/21 16:22	102.76	59.80	151.54	108.16	87.96	129.54	Yes
5	30/7/21 16:36	152.31	90.07	222.33	168.53	137.12	200.03	No
6	30/7/21 16:51	225.42	99.69	392.61	322.45	284.06	360.08	No
7	30/7/21 17:06	82.21	33.30	136.95	87.43	67.63	107.83	Yes
8	30/7/21 17:22	362.67	138.79	613.80	388.06	347.49	430.51	No
9	30/7/21 18:08	205.48	89.54	331.16	213.25	178.73	247.93	Yes
1	3/8/21 15:53	140.85	76.46	211.86	140.85	76.46	211.86	Yes
3	3/8/21 16:11	344.58	189.55	512.42	344.58	189.55	512.42	No
5	3/8/21 16:30	472.22	280.28	682.46	472.22	280.28	682.46	No
7	3/8/21 16:49	299.01	197.33	415.01	299.01	197.33	415.01	No
9	3/8/21 17:06	694.32	537.16	859.69	694.32	537.16	859.69	No
11	3/8/21 17:24	648.93	521.87	791.07	648.93	521.87	791.07	No
13	3/8/21 17:41	243.47	166.76	326.58	243.47	166.76	326.58	Yes
15	3/8/21 17:59	730.75	494.07	987.37	730.75	494.07	987.37	No
17	3/8/21 18:15	385.89	288.5	487.62	385.89	288.50	487.62	Yes
19	3/8/21 18:35	NA	NA	NA	NA	NA	NA	No
21	3/8/21 18:53	525.67	342.93	727.14	525.67	342.93	727.14	No
23	3/8/21 19:11	173.21	118.33	231.18	173.21	118.33	231.18	No

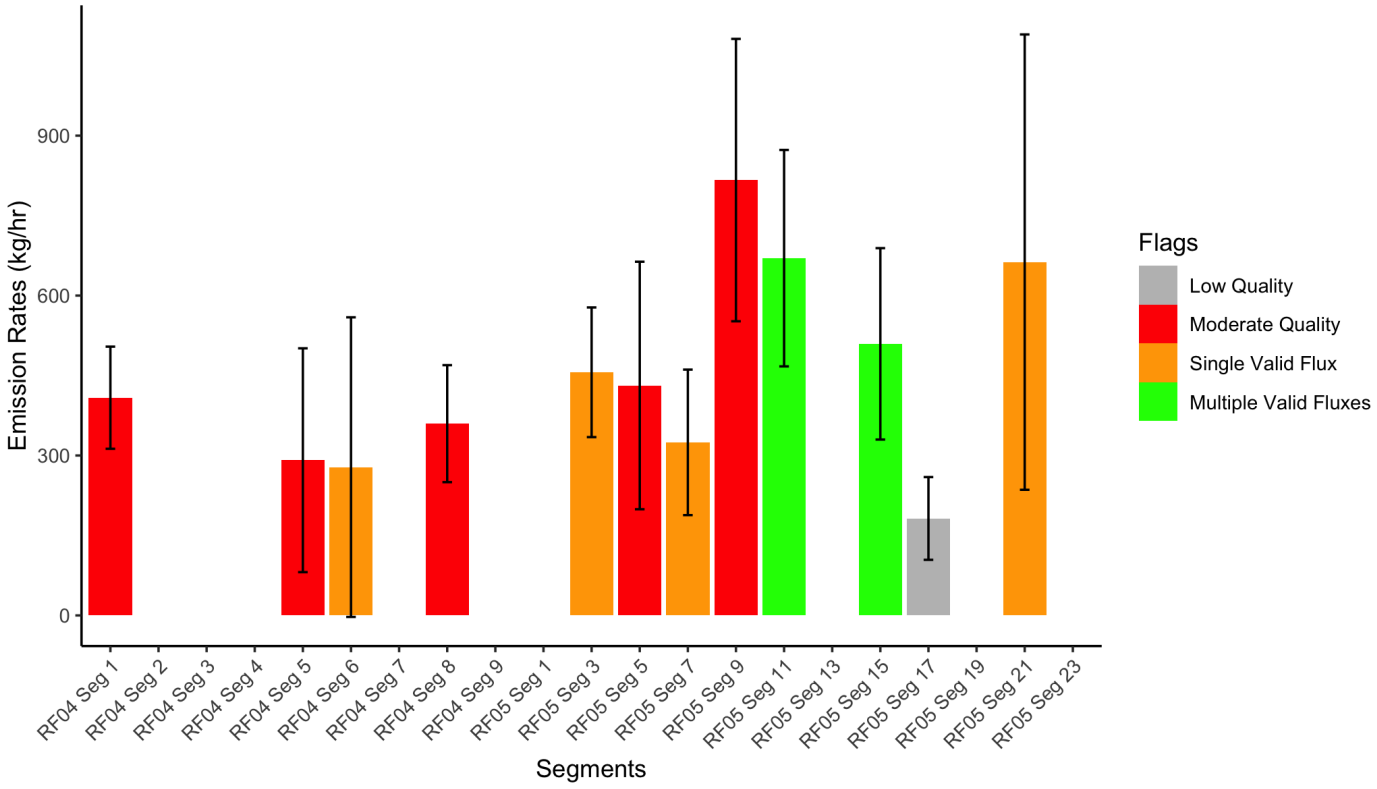


Figure S2: Post-unblinding methane emission estimates from the single-blind volume-controlled release experiments during RF04 and RF05 using the DI method. The endpoints are tagged and color-coded according to the decision tree in Figure S11.

Table S3: DI results from the single-blind volume-controlled release experiments (32.053 °N, 102.301 °W) and MethaneAIR Level 3 Gaussian filtered data on 30 July 2021 and 3 August 2021. The times are in UTC. The emission rates, lower bounds, and upper bounds are in kg/hr.

Seg	Timestamp (UTC)	Unblinded Estimate	Lower Bound	Upper Bound
1	30/7/21 15:41	408.43	312.61	504.25
2	30/7/21 15:52	NA	NA	NA
3	30/7/21 16:07	NA	NA	NA
4	30/7/21 16:22	NA	NA	NA
5	30/7/21 16:36	291.14	81.20	501.09
6	30/7/21 16:51	278.21	-2.83	559.24
7	30/7/21 17:06	NA	NA	NA
8	30/7/21 17:22	359.74	250.00	469.49
9	30/7/21 18:08	NA	NA	NA
1	3/8/21 15:53	NA	NA	NA
3	3/8/21 16:11	456.12	334.50	577.73
5	3/8/21 16:30	431.36	199.19	663.53
7	3/8/21 16:49	324.64	188.14	461.15
9	3/8/21 17:06	816.66	551.81	1081.50
11	3/8/21 17:24	670.18	467.17	873.18
13	3/8/21 17:41	NA	NA	NA
15	3/8/21 17:59	509.45	329.93	688.98
17	3/8/21 18:15	181.93	104.31	259.54
19	3/8/21 18:35	NA	NA	NA
21	3/8/21 18:53	662.77	235.70	1089.84
23	3/8/21 19:11	NA	NA	NA

Table S4: Original results for all methods from the single-blind volume-controlled release experiments (32.053 °N, 102.301 °W) and MethaneAIR Level 3 Gaussian filtered data on 30 July 2021 and 3 August 2021. The emission rates, lower bounds, and upper bounds are in kg/hr. This table was reported to the Stanford team as MethaneAIR initial emission estimates on 1 February 2022. The times are in UTC. The bounds were calculated using the method described in Section 2.5. The SNR is the signal-to-noise ratio. The SL is the scale length of each plume in meters

Seg	Timestamp (UTC)	IME	IME.l	IME.u	Ratio	Ratio.l	Ratio.u	Gaussian	Gauss.l	Gauss.u	SNR	SL
1	30/7/21 15:41	321.74	312.25	330.84	1267.08	1072.46	1419.36	387.04	318.53	455.55	2.33	153.92
2	30/7/21 15:52	232.19	222.98	241.58	765.04	631.16	885.97	NA	NA	4.44	2.45	128.82
3	30/7/21 16:07	109.22	102.16	115.46	206.19	159.37	250.17	NA	NA	NA	2.19	79.98
4	30/7/21 16:22	102.76	94.56	110.81	224.24	189.35	255.61	NA	NA	NA	2.18	87.16
5	30/7/21 16:36	152.31	137.82	165.12	362.74	307	412.7	253.13	158.86	347.39	2.26	106.28
6	30/7/21 16:51	291.55	266.56	313.72	572.28	448.29	685.25	159.29	83.16	235.42	2.5	175.46
7	30/7/21 17:06	82.21	75.89	87.72	131.26	109.96	144.92	NA	NA	NA	2.43	58.3
8	30/7/21 17:22	362.67	339.8	385.66	694.59	589.45	812.25	296.98	260.15	333.82	2.45	196.69
9	30/7/21 18:08	205.48	196.33	214.49	515.09	316.43	651.97	98.32	14.51	182.13	2.13	115.74
1	3/8/21 15:53	140.85	133.45	148.46	215.43	190.36	240.07	NA	NA	NA	2.25	85.46
3	3/8/21 16:11	344.58	322.5	362.84	819.55	664.12	965.76	282.12	223.41	340.84	2.53	185.53
5	3/8/21 16:30	472.22	445.41	497.88	806.07	690.37	972.87	258.84	204.09	313.59	2.49	237.98
7	3/8/21 16:49	299.01	275.76	323.11	707.65	589.56	799.42	128.58	89.51	167.65	2.62	185.79
9	3/8/21 17:06	694.32	644.67	736.81	752.41	616.16	870.63	784.38	710.68	858.07	2.29	363.56
11	3/8/21 17:24	648.93	587.62	711.46	1130.2	784.58	1471.68	1056.21	981.26	1131.17	2.75	203.28
13	3/8/21 17:41	243.47	229.22	256.64	932.64	759.34	1083.86	313.8	252.26	375.35	2.02	128.49
15	3/8/21 17:59	730.75	701.77	761.88	1189.06	1078.02	1297.29	192.44	90.55	294.34	2.52	280.44
17	3/8/21 18:15	385.89	375.8	395.69	818.72	687.49	918.82	NA	NA	3.18	2.48	199.81
19	3/8/21 18:35	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
21	3/8/21 18:53	525.67	490.81	558.78	1242.4	1020.6	1441.07	328.44	207.52	449.37	2.83	171.8
23	3/8/21 19:11	173.21	167.15	178.92	210	178.99	234.64	505.73	447.7	563.76	2.39	53.87

Table S5: Single-blind volume-controlled release data reported by the Stanford team on 1 February 2022. The cr_kgh_CH4_mean90, cr_kgh_CH4_lower90, and cr_kgh_CH4_upper90 represent the 90-second window rolling average, lower, and upper flow rates in kg/hr. The wind speeds are in m/s. The timestamps are in UTC.

Seg	Timestamp (UTC)	cr_kgh_CH4_mean90	cr_kgh_CH4_lower90	cr_kgh_CH4_upper90	Winds
1	30/7/21 15:41	358.48	334.29	384.75	2.8
2	30/7/21 15:52	171.08	154.32	188.05	2.71
3	30/7/21 16:07	49.98	36.06	63.71	2.44
4	30/7/21 16:22	0.00	0.00	0.00	2.29
5	30/7/21 16:36	236.74	217.42	256.81	2.36
6	30/7/21 16:51	295.08	273.58	318.53	2.59
7	30/7/21 17:06	95.05	80.53	109.95	2.68
8	30/7/21 17:22	382.25	356.59	409.77	2.54
9	30/7/21 18:08	63.65	53.99	73.67	3.29
1	3/8/21 15:52	6.72	-2.07	15.42	2.32
3	3/8/21 16:11	495.05	464.61	528.52	2.28
5	3/8/21 16:30	560.57	527.17	596.61	2.27
7	3/8/21 16:48	297.28	277.02	318.78	2.03
9	3/8/21 17:06	654.38	615.38	695.86	1.92
11	3/8/21 17:24	690.80	648.48	734.36	2.07
13	3/8/21 17:41	0.00	0.00	0.00	2
15	3/8/21 17:59	382.28	356.57	409.38	1.64
17	3/8/21 18:15	129.63	113.98	145.39	1.87
19	3/8/21 18:35	56.05	41.93	70.22	2.48
21	3/8/21 18:53	446.05	417.23	476.86	2.42
23	3/8/21 19:11	236.63	217.37	256.66	2.21

Table S6: Single-blind volume-controlled release data reported by the Stanford team on 22 March 2023. The release_rate_kgh is the metered rate reported by the Stanford team. The lower_95CI and upper_95CI are the lower and upper bound of the 95th percentile confidence interval of the metered rates. The MAIR is our blinded best estimate of emissions reported to the Stanford team. The MAIR_lower and MAIR_upper are the lower and upper bound of the 95th percentile confidence interval of our estimates. The timestamps are in UTC.

Seg	Timestamp (UTC)	release_rate_kgh	lower _95CI	upper_95CI	MAIR	MAIR_lower	MAIR_upper
1	25/10/22 17:15	205.23	203.92	206.55	177	121	232
2	25/10/22 17:32	97.04	96.42	97.66	61	18	105
3	25/10/22 17:50	987.13	936.13	1038.13	855	578	1133
4	25/10/22 18:09	982.50	928.47	1036.52	908	662	1155
5	25/10/22 18:27	0.00	0.00	0.00	0	-100	100
6	25/10/22 18:47	468.33	463.82	472.85	669	409	928
7	25/10/22 19:07	24.42	24.31	24.53	0	-100	100
8	25/10/22 19:28	634.57	625.46	643.68	544	374	714
9	25/10/22 20:10	142.14	141.52	142.75	303	183	423
10	25/10/22 20:29	74.65	74.32	74.99	89	-11	189
11	25/10/22 20:46	63.03	62.78	63.29	140	89	191
1	29/10/22 16:25	0.00	0.00	0.00	0	-100	100
2	29/10/22 16:50	340.16	336.53	343.79	196	97	294
3	29/10/22 17:17	671.74	660.33	683.16	787	290	1283
4	29/10/22 17:42	405.59	400.31	410.87	644	271	1017
5	29/10/22 18:07	33.61	33.27	33.94	94	-6	194
6	29/10/22 18:31	337.56	333.89	341.23	388	296	480
7	29/10/22 18:54	818.30	793.86	842.75	958	549	1368
8	29/10/22 19:16	1060.77	1002.64	1118.91	1199	886	1512
9	29/10/22 19:38	1289.35	1221.55	1357.15	1489	1004	1974
10	29/10/22 19:59	585.32	575.53	595.11	579	397	761
11	29/10/22 20:21	199.26	165.06	233.47	252	147	357
12	29/10/22 20:42	493.52	487.02	500.02	584	328	839
13	29/10/22 21:02	168.86	167.19	170.53	112	-21	244

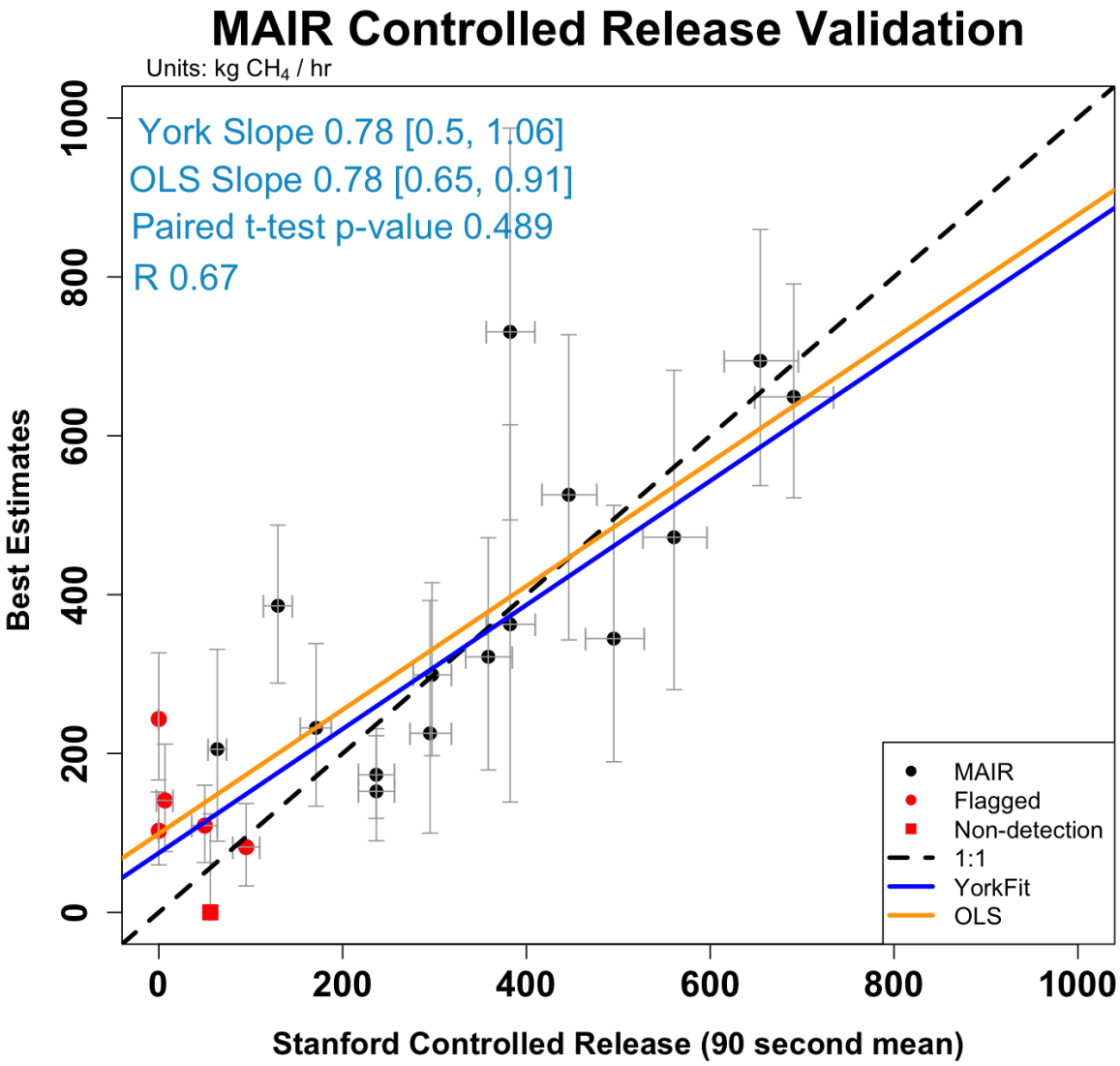


Figure S3: An analogous plot of Figure 3. The results presented in this figure are based on the "best estimates" reported on 1 February 2022. The black circles represent the post-unblinding comparison between the reported and estimated emissions. The red square represents the original comparison, which was later identified as no detection and removed. The red circles represent the flagged data points determined by the decision tree (Section S4.4). The blue solid line is the post-unblinding York fit. The orange is the Ordinary Least Squares (OLS) fit. The comparisons shown by the York and OLS fits are notional and do not actually pertain to the information exchanged in the unblinding. The errors in the MethaneAIR data had not yet been assessed in detail and, in fact, were overestimated. The MethaneAIR team did not know the metered rates and the associated uncertainties were not documented.

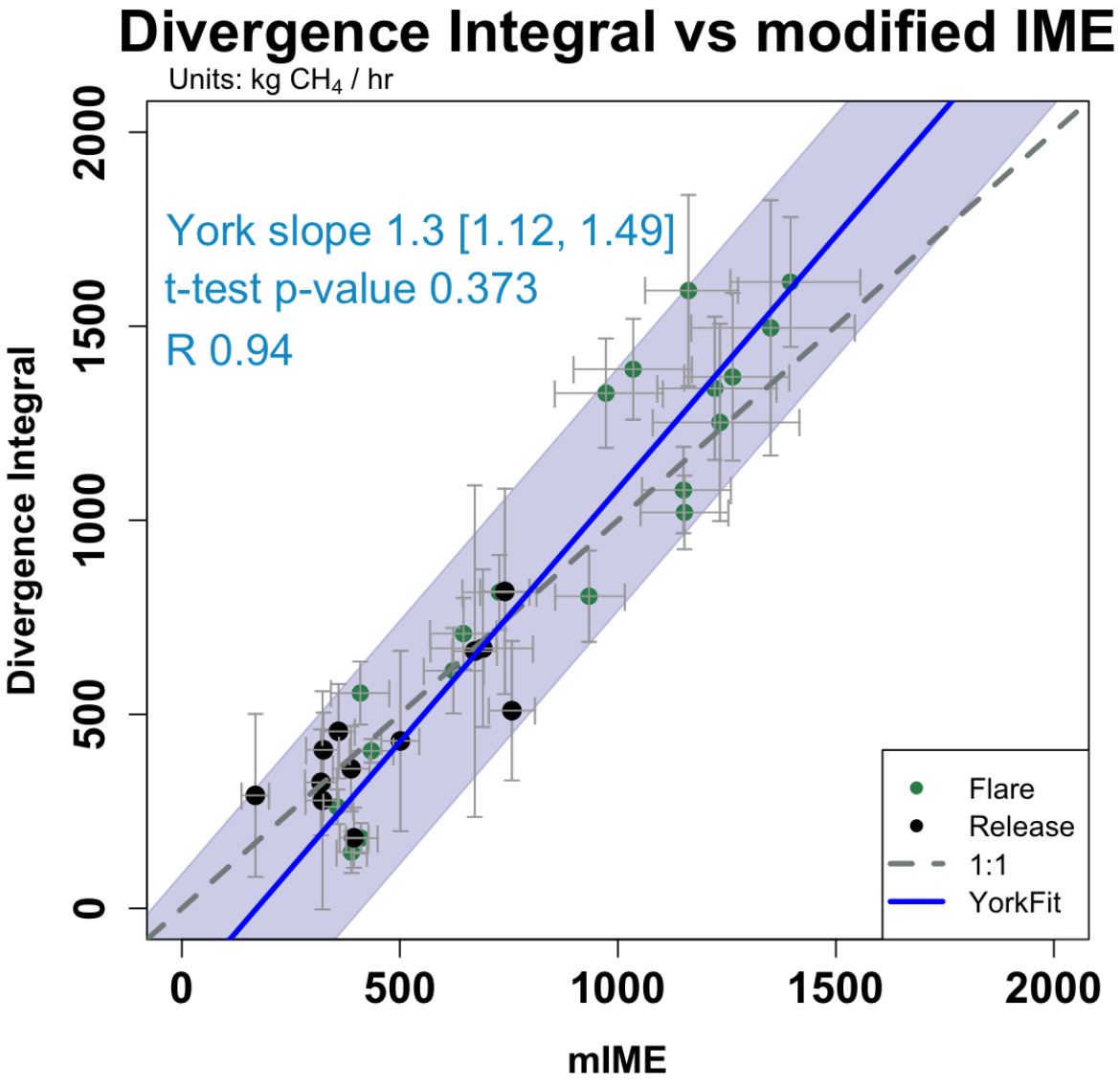


Figure S4: An analogous plot of Figure 3. The results presented in this figure are based on the post-unblinding estimates using the DI results. The black circles represent the post-unblinding comparison between the reported and estimated emissions. The blue solid line is the post-unblinding York fit.

S3.2 Unlit Flare

As shown in Figure S6, the flare emission rates range from 500 kg/hr to 2000 kg/hr.

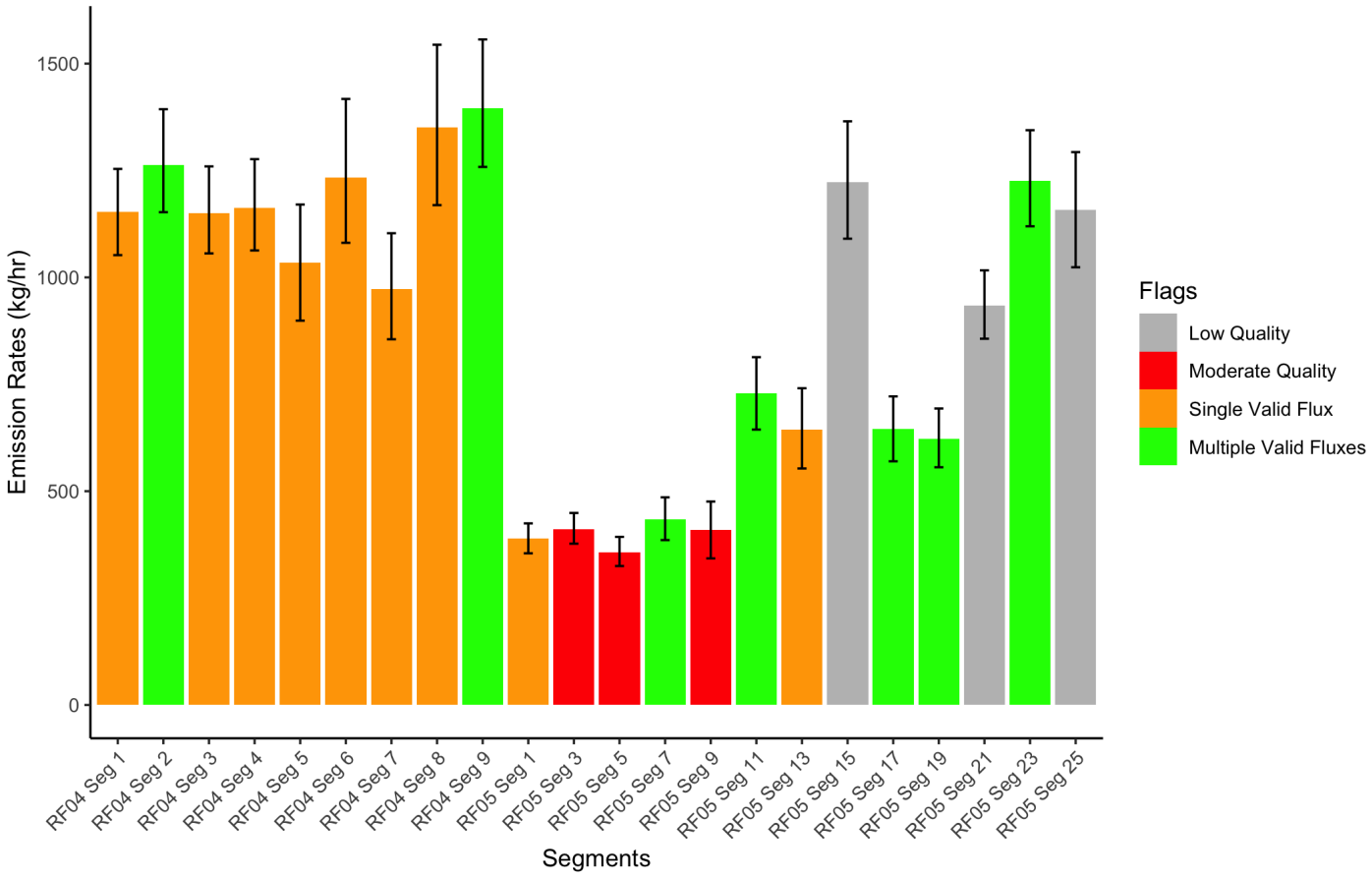


Figure S5: Methane emission estimates from an unlit flare during RF04 and RF05 using the mIME method. The endpoints are tagged and color-coded according to the decision tree in Figure S11.

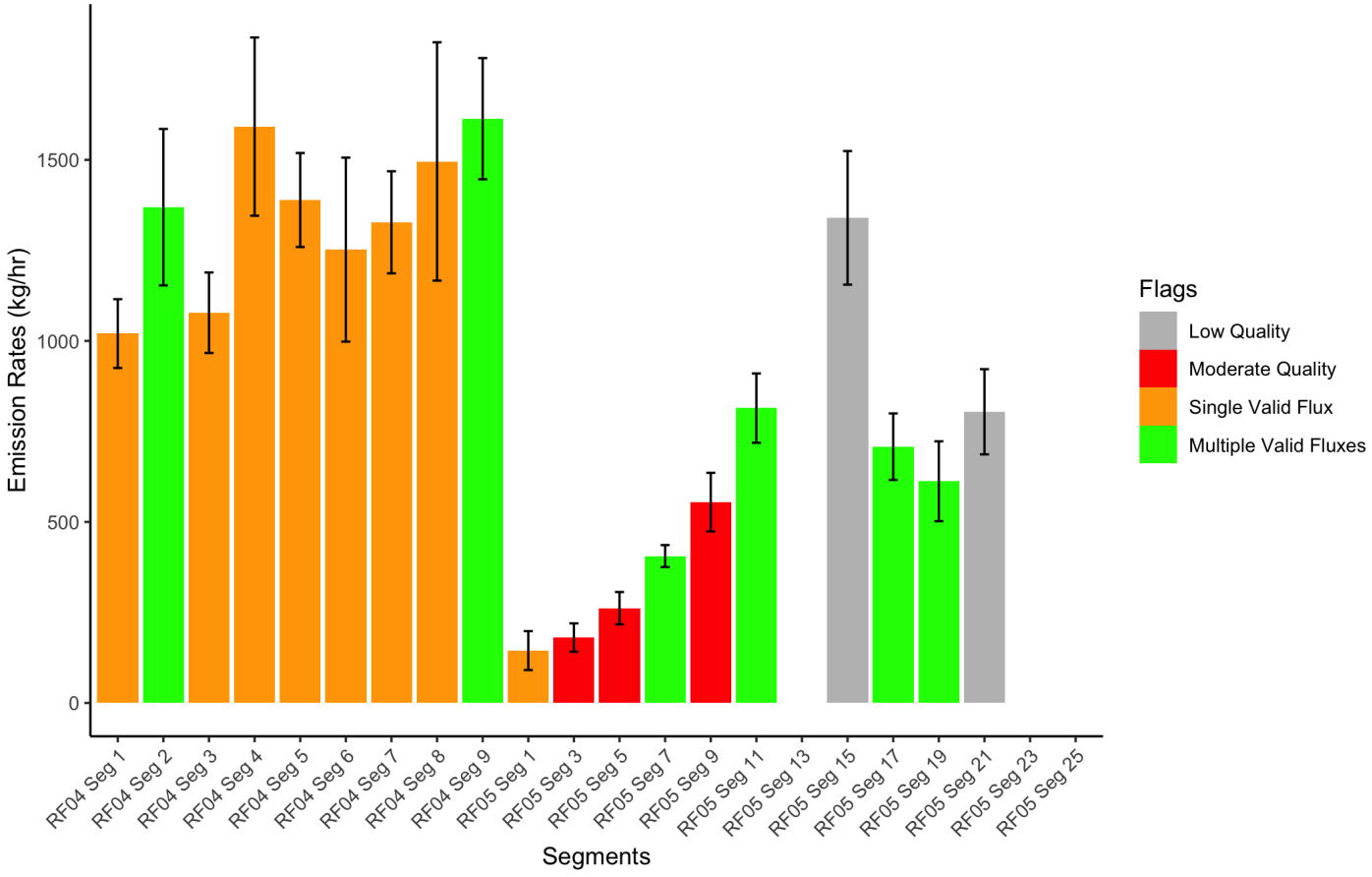


Figure S6: Methane emission estimates from an unlit flare during RF04 and RF05 using the DI method. The endpoints are tagged and color-coded according to the decision tree in Figure S11.

S3.3 MVD & Pipeline

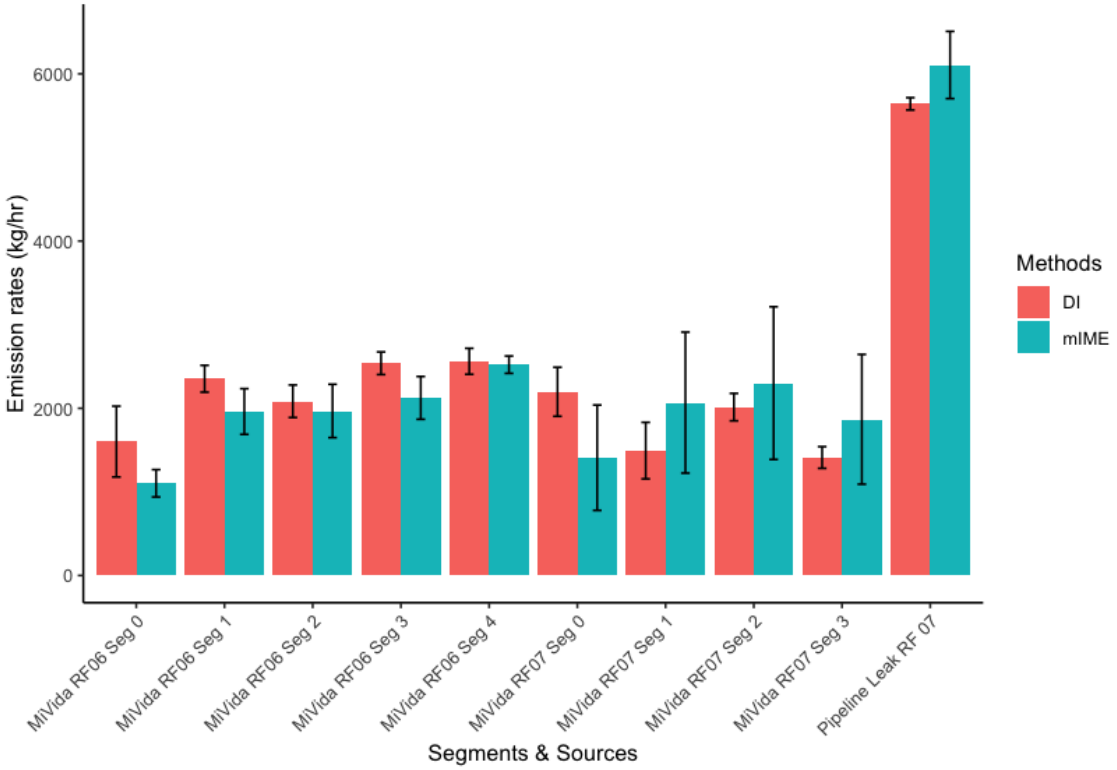
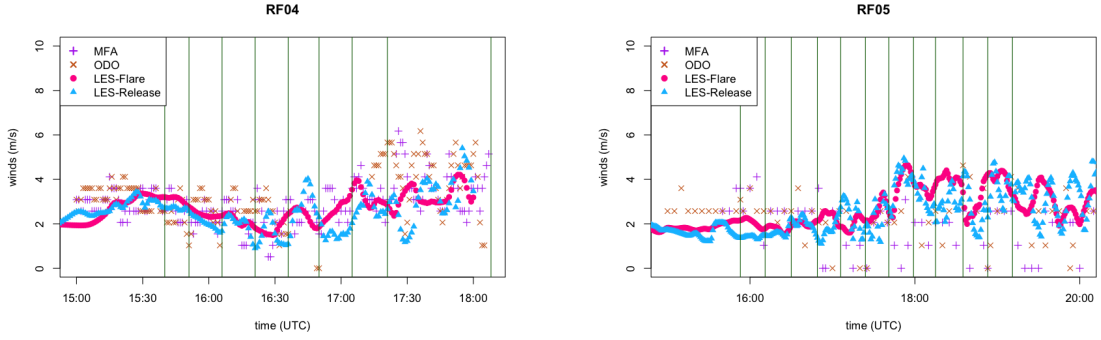


Figure S7: Barplot of the miscellaneous sources quantified in the Delaware Basin.

S4 Analysis Methods

S4.1 Wind Analysis



(a) RF04 30 July 2021

(b) RF05 3 August 2021

Figure S8: ASOS surface wind comparison from the single-blind volume-controlled release experiments during (a) RF04 on 30 July 2021 and (b) RF05 on 3 August 2021. The purple crosses represent ASOS winds from the Midland International Air & Space Port (MAF), which is 15 km from the single-blind volume-controlled release site. The orange x's represent ASOS winds from the Odessa-schlemeyer Field Airport (ODO), which is 17 km from the single-blind volume-controlled release site. The magenta circles represent surface winds from the LES at the flare. The blue triangles represent surface winds from the LES at the single-blind volume-controlled release site. The dark green vertical lines represent times when MAIR flew over the targets. All LES winds during the controlled release experiments and unlit flare observation agree with the ASOS winds (two-sided t-test p-values are greater than 0.05).

S4.2 Wind Product Analysis

To assess our LES winds compared to mesoscale HRRR winds and winds observed at the release site, we plot the estimated emission rates against the known emission rates from the Stanford team using different wind products. As shown in Figure S9, effective wind from the LES provides the highest R^2 value and slope closest to 1. Note that the choice of driven wind product is not limited to HRRR data; any wind products that can drive WRF can run WRF-LES. Although HRRR is the highest temporal and spatial resolution wind product for the Contiguous United States (CONUS), further study of other wind products outside CONUS is needed to determine the most appropriate wind products.

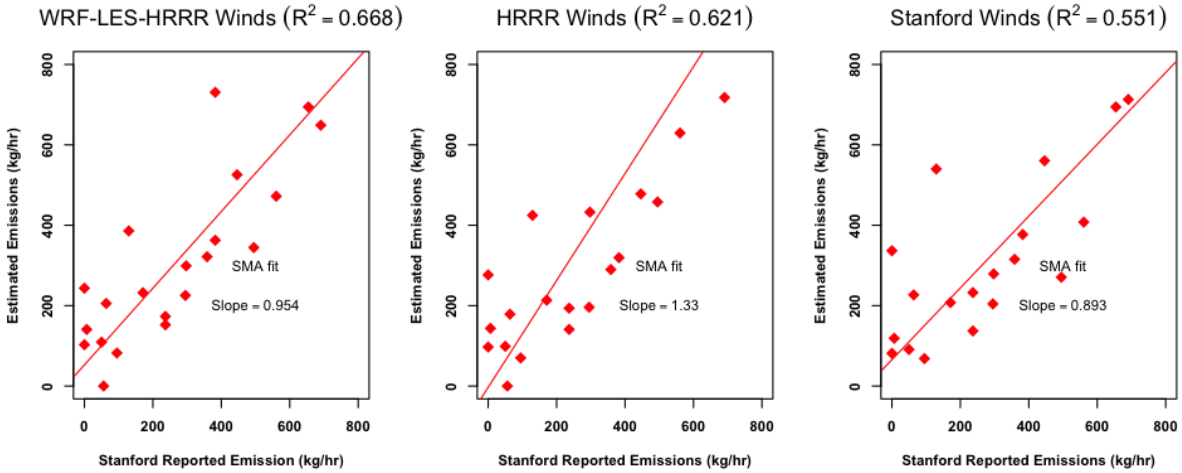


Figure S9: Estimated emission rates with the mIME using different wind products vs. the known emission rates from the Stanford team. From left to right: estimated emission rates based on effective winds from WRF-LES-HRRR, HRRR, and measured winds. The paired t-test between the estimated emissions and Stanford reported emissions suggests that we cannot reject the null hypothesis that a true difference in means is equal to 0 within 95% confidence interval for the estimated emissions based on WRF-LES-HRRR and Stanford observed winds with p-values of 0.166 and 0.309 respectively. However, the test suggests the alternative hypothesis of a true difference in means between the emissions derived from HRRR winds and the reported emissions with a p-value of 0.0408.

S4.3 Threshold Analysis

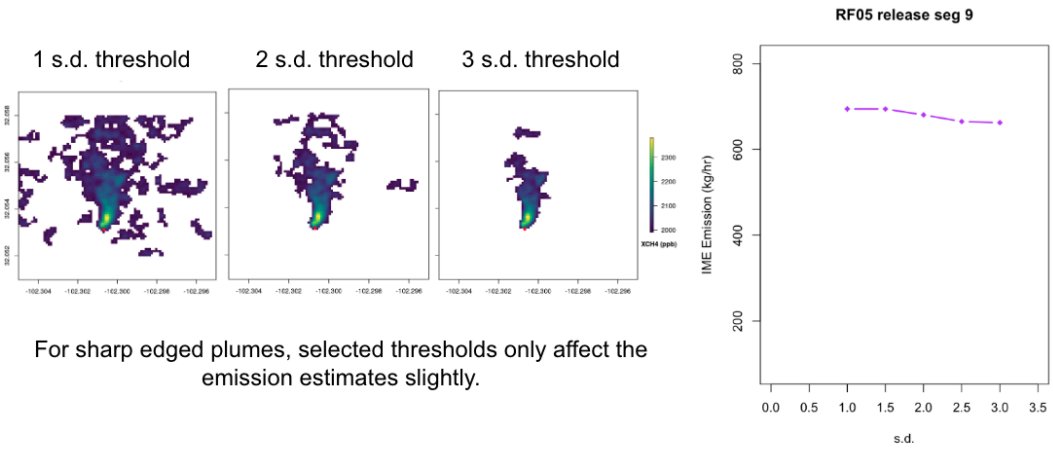


Figure S10: The thresholds do not affect the IME values of the plume of interest.

S4.4 Decision Tree

To understand the relationship between the actual emission rates and the emission estimates, we attempted to use several regression models for our data. York regression was the most appropriate one because York includes correlated errors from both axes. OLS was included in this study since OLS is often used in this type

of experiment. See Sherwin et al. (2022), SI Section S4, for further discussion of tradeoffs associated with York and OLS fits in single-blind tests of methane remote sensing.

Before the actual single-blind volume-controlled release values were revealed, we developed a decision tree workflow that assigns a confidence flag to each case (Section S4.4). The original version of our decision tree was developed purely based on our analysis of our measurement system and algorithms. It was then adjusted slightly after unblinding the controlled release metered emissions. When the signal-to-noise ratio (SNR) for the plume, relative to the background, is below 1 or significant interference is observed, we classify the case as either no detection (Endpoint 0) or poor quality (Endpoint 1). When the SNR is between 1 and 2.5, the estimates are near the detection limit, providing a moderate-quality estimate (Endpoint 2). Only the mIME method can be used, and the confidence will likely be low. When the SNR is above 2.5, we checked if the plume size is greater than one eddy. If the plumes are smaller than one eddy, only the mIME method can be used (Endpoint 3). If the plumes are larger than one eddy, all three methods developed were applied (Endpoint 4). For the last step, the WRF-HRRR-LES were compared to the nearby ASOS winds and later assigned quality flags based on agreement to the observed winds. If the winds disagree (two-sided t-test p-value is less than 0.05), the wind quality is low. When winds agree (two-sided t-test p-value is greater than 0.05), the wind quality is high. All LES winds during the controlled release experiments and unlit flare observation agree within tolerance with the ASOS winds.

As expected, the DI and the Ratio approaches work well with larger sources such as the unlit flare, MiVida Gas Processing Plant, or the pipeline leak plume. However, the ratio approach is limited by the resolution mismatch with the sensor and the LES domain size, making this approach less favorable. So, we decided to proceed with the mIME and the DI methods.

After the metered single-blind volume-controlled release values were revealed, the decision tree was revised to include additional steps that help remove false positive cases and help us determine the detection limits of MethaneAIR.

By filtering out cases that (i) have the absolute differences between mIME emission estimates using a threshold of 1 and 2 standard deviations divided by emission estimate using a threshold of 1 standard deviation greater than 0.5 (i.e., 2 s.d. case has a much different plume shape) or (ii) no emission estimates at the threshold of 2 standard deviations, we designate as "below detection limit" 6 cases from the single-blind volume-controlled release experiment. One was a false negative case, where methane enhancements were not detected. The other four were emission estimates around the detection limit of 200 kg/hr. The part of the decision tree added after unblinding, intended to reduce false positives in the contaminated environment of this test, is boxed in red in Figure S11.

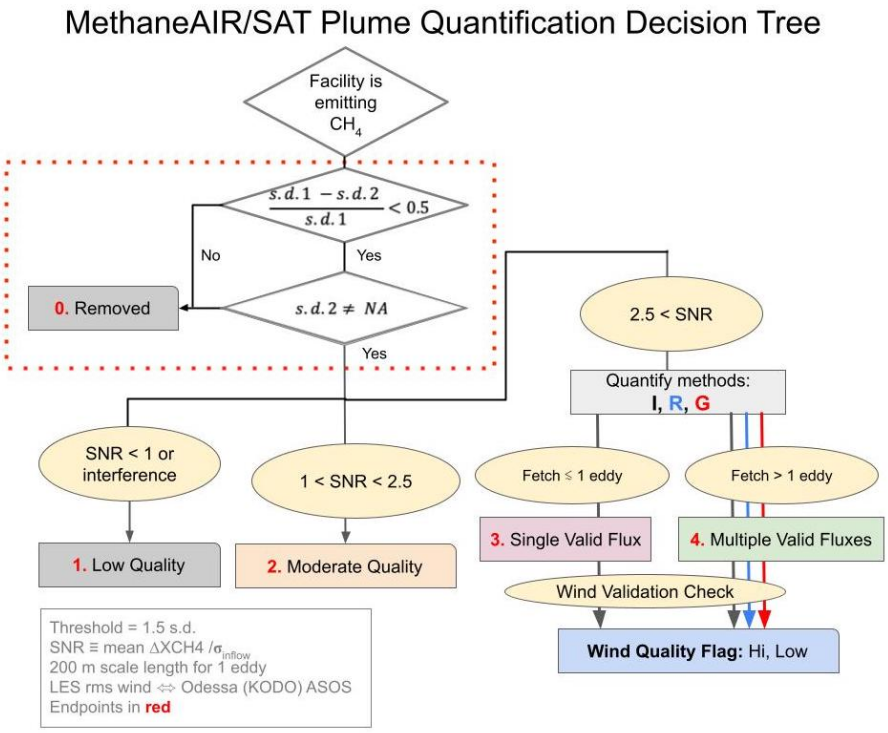


Figure S11: Updated decision tree classifying quality levels of emission estimates. Most of the steps were developed purely based on our analysis of the data and signal processing. The steps in the red dashed box were added after the unblinding process. First, we filter out all the estimates that (i) have the absolute differences between mIME emission estimates using thresholds of 1 and 2 standard deviations divided by emission estimate using a threshold of 1 standard deviation to be greater than 0.5 or (ii) no emission estimates at the threshold of 2 standard deviations. After that, for each estimate with SNR below 1 or significant interference was observed, we classify the case as either no detection (Endpoint 0) or poor quality (Endpoint 1). The estimates are moderate quality when the SNR is between 1 and 2.5 (Endpoint 2). Only the mIME method can be used. When the SNR is above 2.5, we checked if the plume size is greater than one eddy. If the plumes are smaller than one eddy, only the mIME method can be used (Endpoint 3). All three different methods developed were applied if the plumes are larger than one eddy (Endpoint 4). For the last step, the WRF-HRRR-LES were compared to the nearby ASOS winds and later assigned quality flags based on agreement to the observed winds. If the winds disagree, the wind quality is low. When winds agree, the wind quality is high.

S4.5 Additional Attempts to Determine Contamination Near the Methane Plumes

Although our decision tree (Section S4.4) removed six plumes (below detection) from the total 21 plumes based on the signal-to-noise ratio and the size of the plumes, we wanted to investigate the limitations of our sensors and methods further. We decided to perform two more analyses with the hypothesis that these two will help us better understand the limitations.

The first analysis was to perform a k-s test between XCH_4 in the inflow region of each scene and a clear scene defined as latitude 32.10 - 32.11 and longitude -102.32 to -102.31. When the p-values of the k-s test are less than 0.05, we can reject the null hypothesis that the XCH_4 observed in the inflow and clear scene came from the same distribution. A positive outcome from the K-S test suggests that there may be extraneous methane contamination in the inflow.

The second analysis was to filter the clumps with 60 pixels and 100 pixels. The new choices of thresholds are arbitrarily rounded to the multiples of the original 20-pixel threshold and fractions of a large eddy (approximately 100 m). The original 20-pixel threshold or 10 m $\times$ 10 m $\times$ 20 pixels is equivalent to a 45 m $\times$ 45 m plume, which is approximately half of the size of an eddy. The 60-pixel threshold is equivalent to a 77 m $\times$ 77 m plume,

which is approximately two-thirds of the size of an eddy. The 100-pixel threshold is equivalent to a 100 m × 100 m plume, which is approximately the size of an eddy.

Table S7: Summary table of removed (not detected) RF04 plumes based on three methods: decision tree (Section S4.4), 60 pixel threshold, and K-S test.

Segment	Decision Tree	60 pixels	K-S test
1			x
2			x
3	x	x	
4	x	x	
5		x	x
6			
7	x	x	
8			
9	x	x	

Table S8: Summary table of removed RF05 plumes based on three methods: decision tree (Section S4.4), 60 pixel threshold, and K-S test.

Segment	Decision Tree	60 pixels	K-S test
1	x	x	x
3			x
5	x	x	x
7	x	x	
9		x	x
11			
13	x	x	x
15			
17	x	x	
21			
23			

S5 Cropped Blinded-Volume controlled release Plumes

The following figures are cropped single-blind volume-controlled release plumes from RF04 and RF05. The plumes were identified based on the known location of the single-blind volume-controlled release site (grey triangles). To reduce the noise of the raw 1×1 Level 3 data, the data were later filtered by a Gaussian filter (7-pixel x 7-pixel kernel with the full width at a maximum of 2 pixels). We then identified the inflow region and removed the pixels below the 1.5 standard deviation of the inflow threshold. After that, we removed the small clumps with fewer than 20 pixels from the scene of interest.

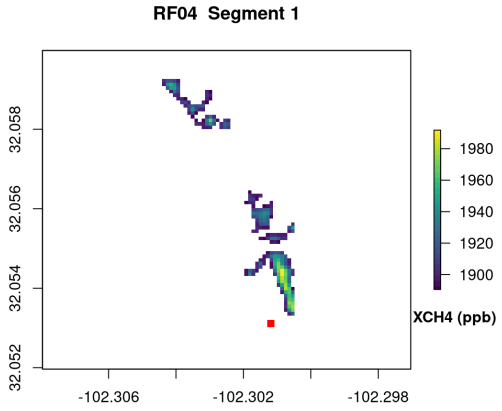


Figure S12: Filtered single-blind volume-controlled release
RF04 Segment 1.

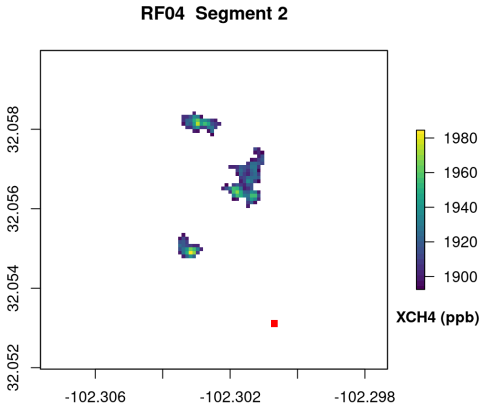


Figure S13: Filtered single-blind volume-controlled release
RF04 Segment 2.

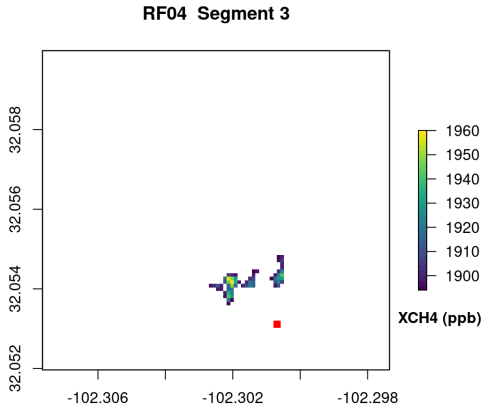


Figure S14: Filtered single-blind volume-controlled release
RF04 Segment 3.

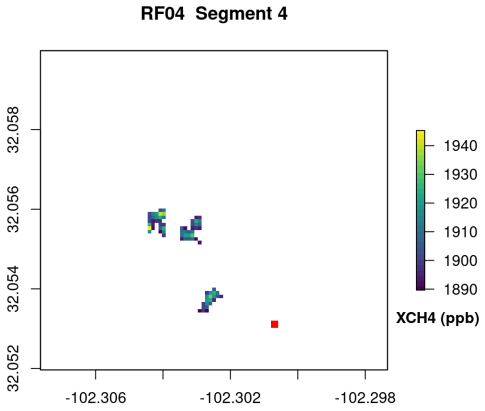


Figure S15: Filtered single-blind volume-controlled release
RF04 Segment 4.

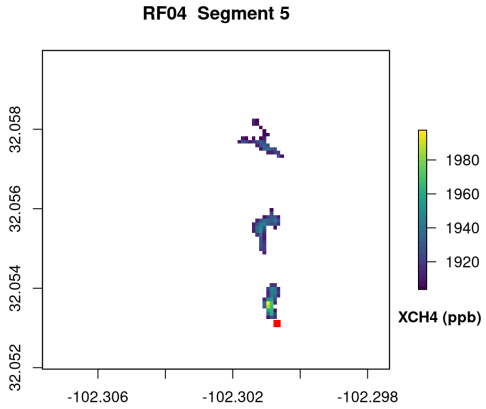


Figure S16: Filtered single-blind volume-controlled release
RF04 Segment 5.

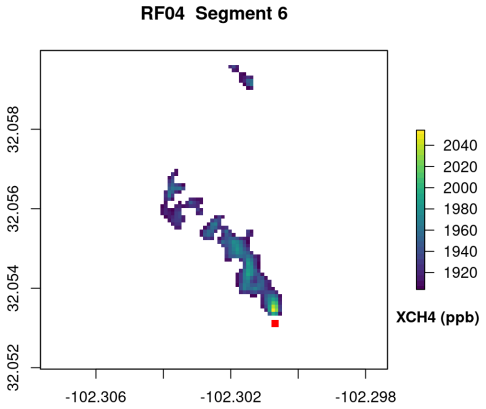


Figure S17: Filtered single-blind volume-controlled release
RF04 Segment 6.

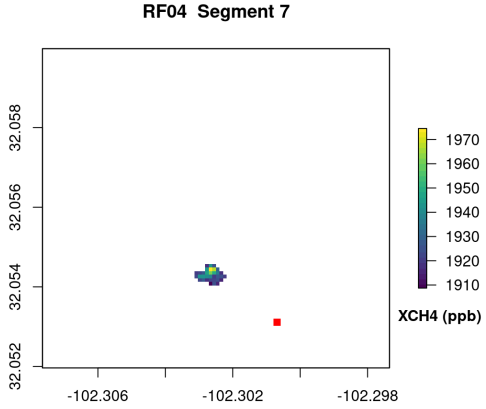


Figure S18: Filtered single-blind volume-controlled release
RF04 Segment 7.

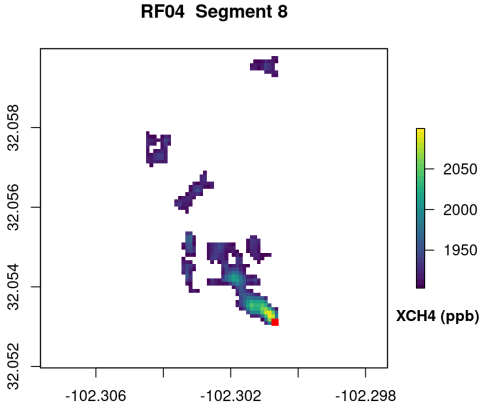


Figure S19: Filtered single-blind volume-controlled release
RF04 Segment 8.

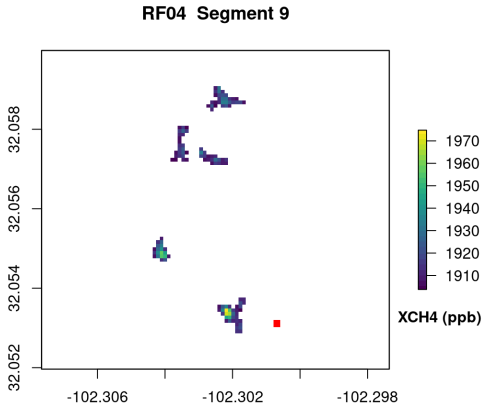


Figure S20: Filtered single-blind volume-controlled release
RF04 Segment 9.

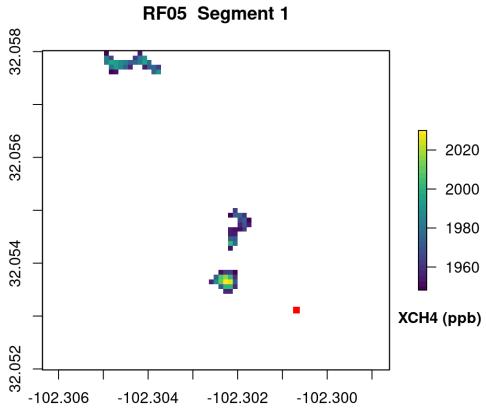


Figure S21: Filtered single-blind volume-controlled release
RF05 Segment 1.

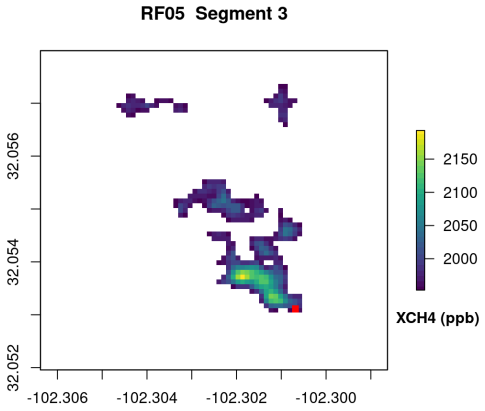


Figure S22: Filtered single-blind volume-controlled release
RF05 Segment 3.

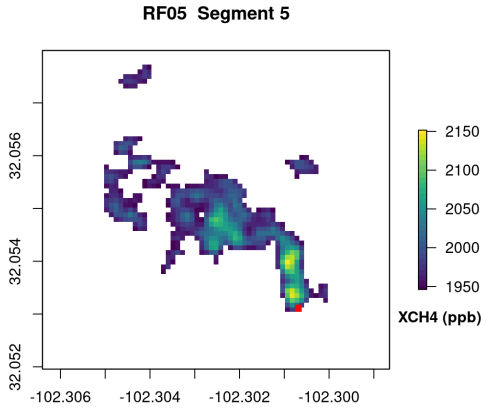


Figure S23: Filtered single-blind volume-controlled release
RF05 Segment 5.

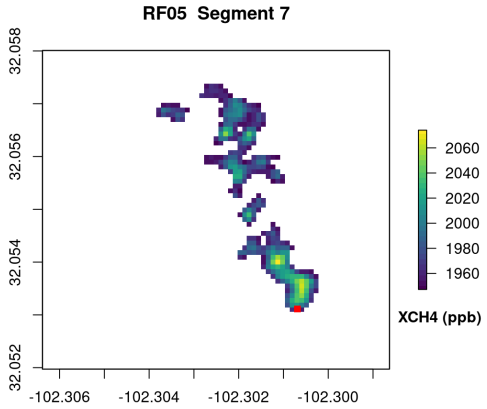


Figure S24: Filtered single-blind volume-controlled release
RF05 Segment 7.

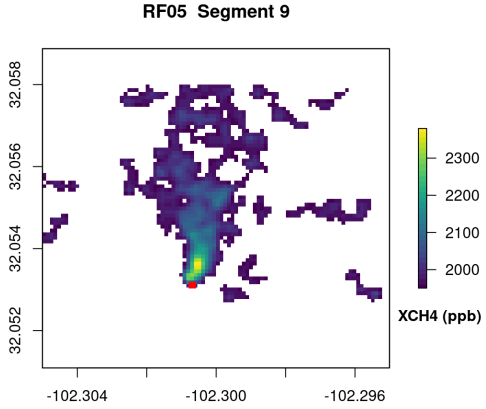


Figure S25: Filtered single-blind volume-controlled release
RF05 Segment 9.

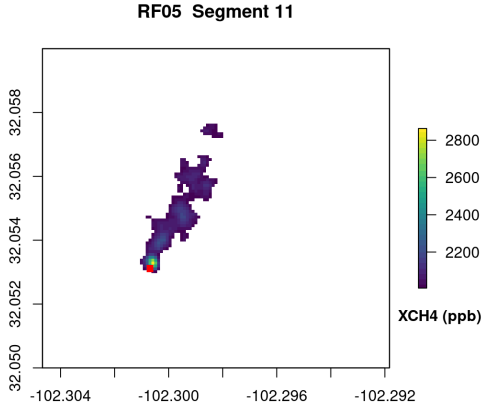


Figure S26: Filtered single-blind volume-controlled release
RF05 Segment 11.

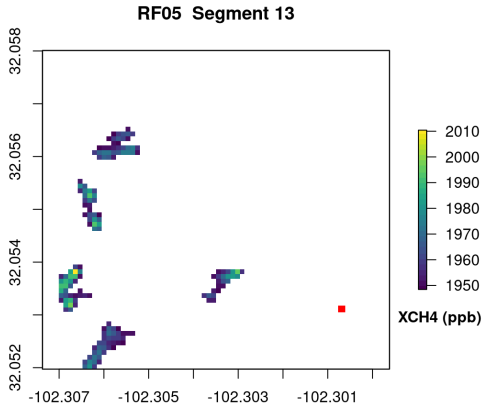


Figure S27: Filtered single-blind volume-controlled release
RF05 Segment 13.

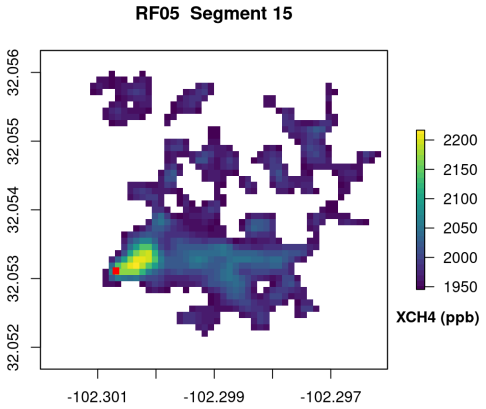


Figure S28: Filtered single-blind volume-controlled release
RF05 Segment 15.

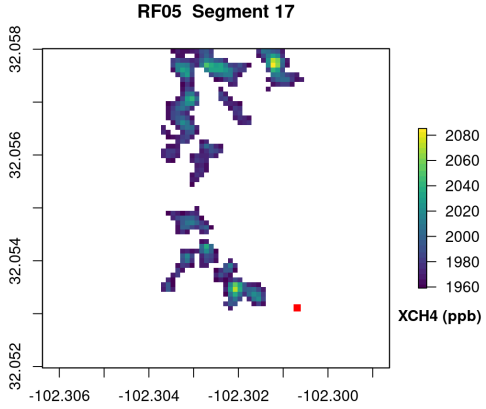


Figure S29: Filtered single-blind volume-controlled release
RF05 Segment 17.

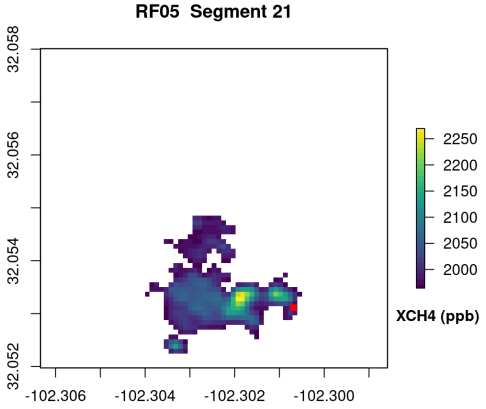


Figure S30: Filtered single-blind volume-controlled release
RF05 Segment 21.

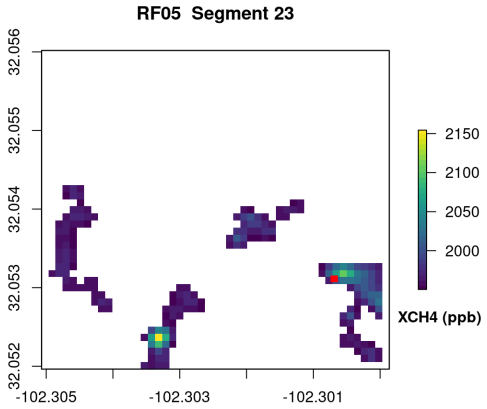


Figure S31: Filtered single-blind volume-controlled release
RF05 Segment 23.

S6 MethaneAIR flight tracks

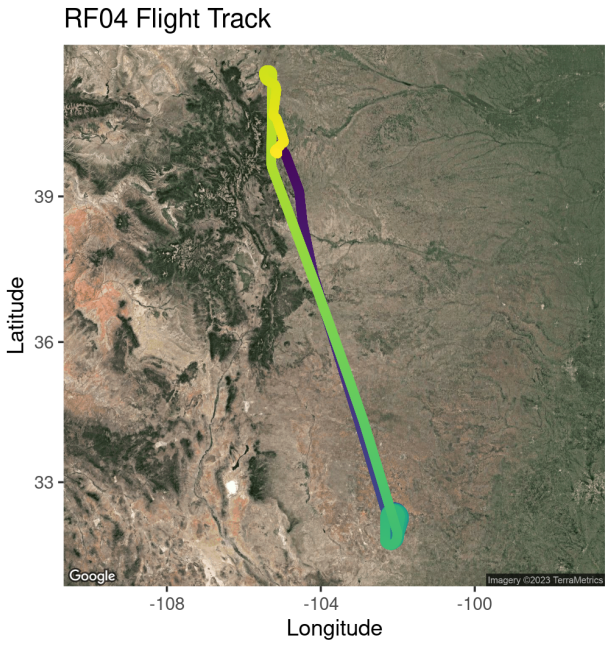


Figure S32: RF04

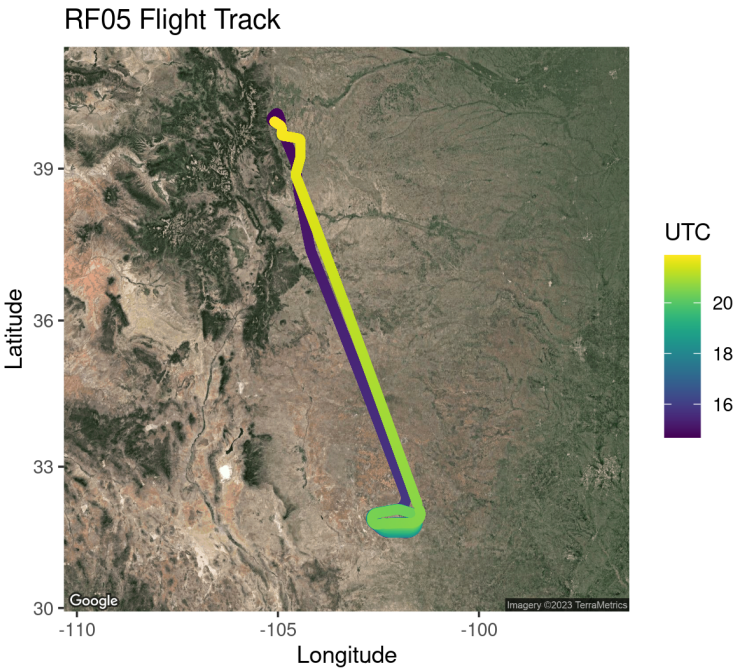


Figure S33: RF05

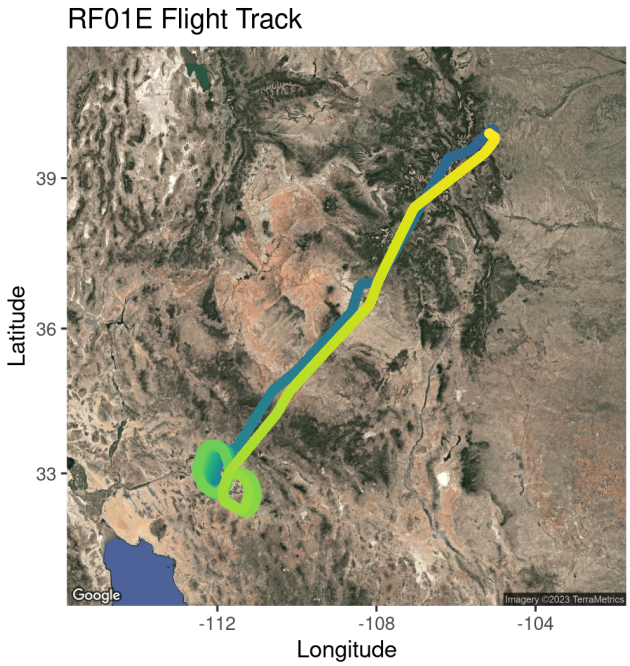


Figure S34: RF01E

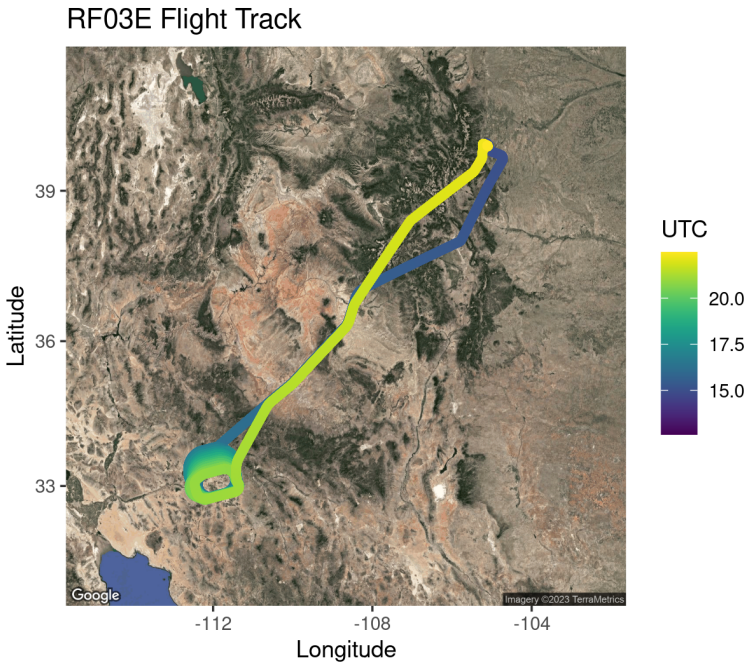


Figure S35: RF03E

References

Conley, S., Faloona, I., Mehrotra, S., Suard, M., Lenschow, D. H., Sweeney, C., Herndon, S., Schwietzke, S., Pétron, G., Pifer, J., Kort, E. A., and Schnell, R.: Application of Gauss's theorem to quantify localized surface emissions from airborne measurements of wind and trace gases, Atmospheric Measurement Techniques, 10, 3345–3358, https://doi.org/10.5194/amt-10-3345-2017, 2017.

Irakulis-Loitxate, I., Guanter, L., Liu, Y.-N., Varon, D. J., Maasakkers, J. D., Zhang, Y., Chulakadabba, A., Wofsy, S. C., Thorpe, A. K., Duren, R. M., Frankenberg, C., Lyon, D. R., Hmiel, B., Cusworth, D. H., Zhang, Y., Segl, K., Gorroño, J., Sánchez-García, E., Sulprizio, M. P., Cao, K., Zhu, H., Liang, J., Li, X., Aben, I., and Jacob, D. J.: Satellite-based survey of extreme methane emissions in the Permian basin, Science Advances, 7, eabf4507, https://doi.org/10.1126/sciadv.abf4507, 2021.

Rutherford, J. S., Sherwin, E. D., Chen, Y., Aminfard, S., and Brandt, A. R.: Evaluating methane emission quantification performance and uncertainty of aerial technologies via high-volume single-blind controlled releases, URL `https://eartharxiv.org/repository/view/5113/`, publisher: EarthArXiv, 2023.

Sherwin, E. D., Rutherford, J. S., Chen, Y., Aminfard, S., Kort, E. A., Jackson, R. B., and Brandt, A. R.: Single-blind validation of space-based point-source methane emissions detection and quantification, URL `https://eartharxiv.org/repository/view/3465/`, publisher: EarthArXiv, 2022.

Sun, K., Zhu, L., Cady-Pereira, K., Chan Miller, C., Chance, K., Clarisse, L., Coheur, P.-F., González Abad, G., Huang, G., Liu, X., Van Damme, M., Yang, K., and Zondlo, M.: A physics-based approach to oversample multi-satellite, multispecies observations to a common grid, Atmospheric Measurement Techniques, 11, 6679–6701, https://doi.org/10.5194/amt-11-6679-2018, publisher: Copernicus GmbH, 2018.

Varon, D. J., Jacob, D. J., McKeever, J., Jervis, D., Durak, B. O. A., Xia, Y., and Huang, Y.: Quantifying methane point sources from fine-scale satellite observations of atmospheric methane plumes, Atmospheric Measurement Techniques, 11, 5673–5686, https://doi.org/10.5194/amt-11-5673-2018, 2018.

York, D.: Least squares fitting of a straight line with correlated errors, Earth and Planetary Science Letters, 5, 320–324, https://doi.org/10.1016/S0012-821X(68)80059-7, 1968.

York, D., Evensen, N. M., Martinez, M. L., and De Basabe Delgado, J.: Unified equations for the slope, intercept, and standard errors of the best straight line, American Journal of Physics, 72, 367–375, https://doi.org/10.1119/1.1632486, publisher: American Association of Physics Teachers, 2004.