Supplementary Material for

Comparison of methane emission estimates from multiple measurement techniques at natural gas production pads.

Clay S. Bell [a], Timothy L. Vaughn [a], Daniel Zimmerle [a*], Scott, C. Herndon [b], Tara I. Yacovitch [b], Garvin A. Heath [c], Gabrielle Pétron [d,e], Shane M. Murphy [f], Anna M. Robertson [f]

[a] Colorado State University, Energy Institute and Mechanical Engineering, 430 N College, Fort Collins, CO 80524, USA

[b]Aerodyne Research, Inc. 45 Manning Road, Billerica, MA, 01821, USA

[c] National Renewable Energy Laboratory, 15013 Denver West Parkway, Golden, CO 80401, USA

[d] NOAA Earth System Research Laboratory, 325 Broadway, Boulder, CO 80305, USA

[e] University of Colorado Boulder, Cooperative Institute for Research in Environmental Sciences, 216 UCB, Boulder, CO 80309, USA

[f] University of Wyoming, Department of Atmospheric Science, 1000 E. University Ave. Dept. 3038, Laramie, WY 82071

[*]email: dan.zimmerle@colostate.edu

47 Pages

15 Figures

9 Tables

Table of Contents

S1. Study Design

Role of Participants

The project was organized as a cooperation between industry partners and the study team, with financial support from study partners, Department of Energy, State of Colorado, and other sources, as indicated in the acknowledgement. The study team and study partners signed a governance agreement that specified operating principals, processes and confidentiality.

The field campaign was designed by the study team, who selected the basin and designed the experimental methods used in the study. Study partners provided information on facilities and operations to support the design weeks prior to the field campaign. While all participants in the study reviewed the design prior to the campaign, the study team retained full control of methods and execution of the campaign.

To the extent possible (see below), partners provided access to facilities during the field campaign with limited prior knowledge of the selected facilities prior to the day of (or night before) measurement. Careful observation by the study team detected no change in facility operations as a result of the study, as observed by either onsite observers or measurement teams operating downwind or from aircraft. During the field campaign, study partners provided access to facilities selected by the study team and data about ongoing operations at those facilities.

Following the campaign, the study team requested additional information about the facility operations, which study partners provided in either face-to-face meetings or via other, typically confidential, correspondence. Study partners provided fundamental data, such as operating modes, flow rates, counts, and similar information. All analysis was performed by, and was under full control of, the study team throughout the project.

Clustered Random Sampling

Clustered random sampling was used in this study. The purpose of the clustered sampling strategy was to (1) increase the overall number of samples by decreasing the travel time between facilities, and (2) increase the number of paired measurements by keeping multiple measurement teams near each other. The clustered sampling strategy served to increase the sample size and improve the logistics during the campaign considering time, resource and cost constraints.

Prior to the field campaign, production sites were clustered using the following methodology:

1. A well list was collected from the state database.
2. Wells within 100 m were grouped to pads.
3. Pads were then grouped with the nearest known gathering compression facility.
4. Starting with a random gathering facility, neighboring gathering facilities were grouped until at least 100 partner production facilities were included in a cluster.
5. The next nearest gathering facility was then used to initiate the next cluster to which neighboring gathering facilities were grouped to create a cluster of at least 100 partner production sites. This was repeated until all facilities were assigned to a cluster.

This process resulted in 17 clusters, 6 of which were randomly selected to measure as shown in Figure S1.

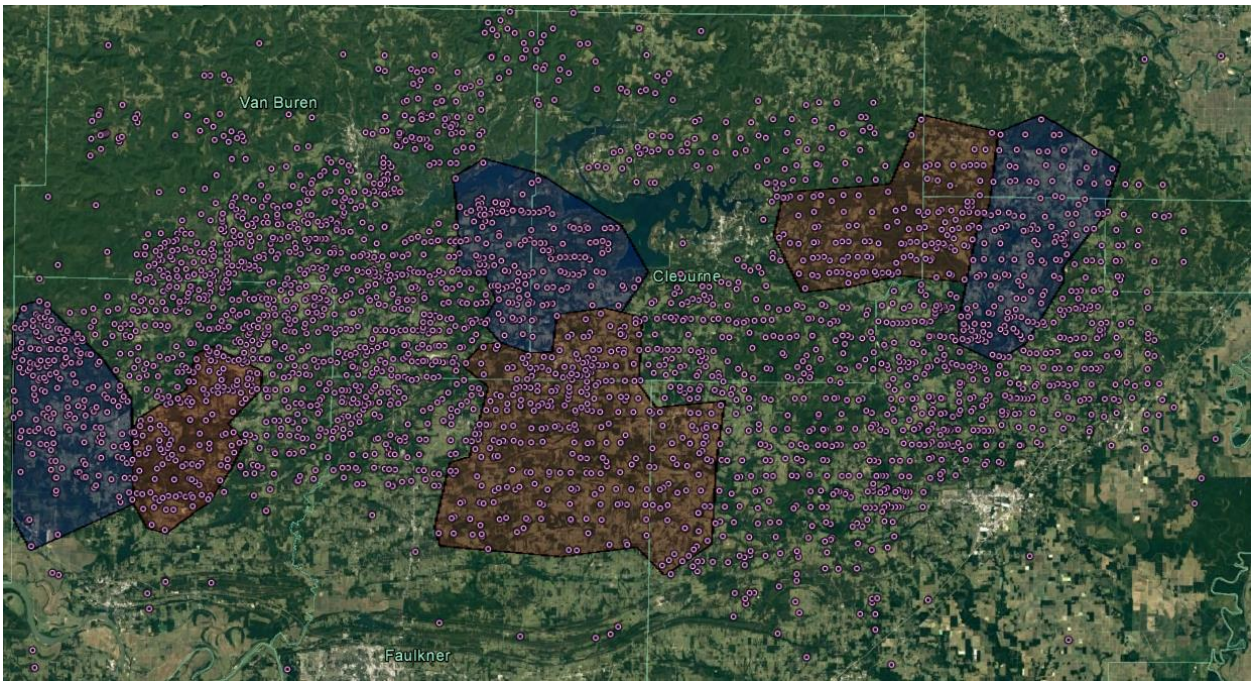

Figure S1: Production facilities in study area. Red and blue shaded areas represent clusters selected for sampling. Dots represent well pad ("production facility") locations.

Site Selection

Measurement teams (Onsite, OTM33A, and Tracer Release) were dispatched to a selected cluster each day. A study coordinator from Colorado State University directed teams to facilities in order to pair as many measurements as possible. The study coordinator kept in regular contact

with all dispatched teams and would select facilities for paired measurements. The study coordinator selected primary measurement targets each morning by reviewing all partner facilities within the cluster for public road access downwind to enable measurement by the dual tracer technique and the OTM33a technique. The study coordinator did not select facilities based on the size or extent of the pad, the number of wells on the facility, the age of the facility, or knowledge of the operations or emissions at a facility from a prior measurement result.

Each measurement team was accompanied by a study partner escort. The role of the escorts was limited to providing access to facilities for measurement, providing knowledge of the operational state at a given facility, and ensuring that safety protocols were followed. Study partners did not participate in the facility selection and were not informed which cluster would be visited in advance of the measurement day to eliminate any intentional changes in facility operation that could affect the measurements.

Representativeness of Sample

The number of active wells per pad was utilized to assess how representative the measurement sample is of the study area population. Well count is indicative of the size (physical extent) of the facility and the amount of equipment on the facility. Due to lease agreements, facilities in the study area generally have 1 wellhead, 1 separator and 1 meter run per well, but typically multiple wells share a common produced water tank. Table S1 and Figure S2 show the fraction of facilities in the study area population and in the sampled population by active well count. The sample includes a smaller fraction of facilities with a single active well and a slightly larger fraction of facilities with greater than 6 active wells relative to the study area population. This is a result of partner facilities having more active wells per pad than non-partner facilities. Facilities with zero active wells may include wells classified by AOGC as abandoned orphaned, dry and abandoned, plugged and abandoned, temporarily abandoned, permitted, completed, expired permit, released water well, or spudded. Five facilities with zero active wells were measured during this study. These 5 facilities included temporarily abandoned wells.

Table S1: Count of facilities by number of active wells per pad in study area, partner facilities, non-partner facilities, and sampled during this study.

Active well count	Count of facilities in study area population (fraction of facilities with active wells)	Count of partner facilities in study area (fraction of partner facilities with active wells)	Count of non-partner facilities in study area (fraction of non-partner facilities with active wells)	Count of facilities sampled (fraction of sampled facilities with active wells)
1	1096 (45%)	739 (40%)	357 (59%)	91 (34%)
2	472 (19%)	327 (18%)	145 (24%)	51 (19%)
3	353 (15%)	283 (15%)	70 (12%)	45 (17%)
4	246 (10%)	233 (13%)	13 (2%)	31 (12%)
5	124 (5%)	119 (7%)	5 (1%)	18 (7%)
6	75 (3%)	68 (4%)	7 (1%)	12 (4%)
>6	64 (3%)	60 (3%)	4 (1%)	21 (8%)

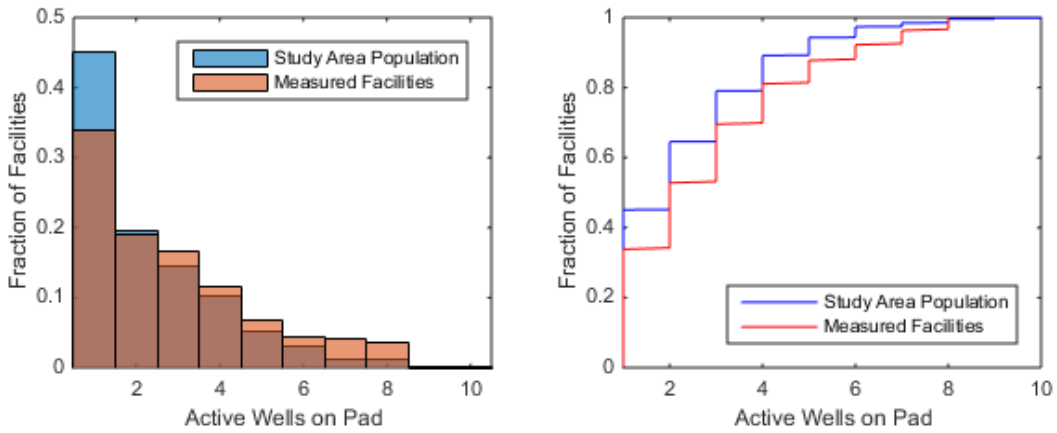

Figure S2: Histogram (left) and cumulative distribution function (right) of number of active wells per pad in study area population and sampled population

Paired Measurements

Table S2 shows the number of facility-level estimates produced by each measurement method, the number of facilities at which a measurement was attempted by each method, and the number of facilities at which the method was paired with another measurement method.

Table S2: Count of facility-level emission estimates by method and by methods paired

			Successful Facility-Level Estimates			
	Successful Facility-Level Estimates	Total Facilities Attempted	Paired with Onsite	Paired with OTM33A	Paired with Tracer	Unpaired
Onsite	261	261	-	43	16	210
OTM33A	50 (38 OTM33A + 12 0BOT)	76 (64 OTM33A + 12 0BOT)	43	-	9	7
Tracer	17 (14 Tracer + 3 LOD)	18	16	9	-	0

Cumulative Distributions of Emission Estimates

A cumulative distribution of each facility level estimate (SOE, OFE, TFE, GFE) is shown in Figure S3. The distribution of measurements is assumed representative of the study area population since facilities were selected using a random sampling methodology. The onsite and OTM33A teams measured a similar distribution of sites during this study. The distribution of tracer measurements suggests facility level emissions are larger than SOE or OFE would suggest however this may be a result of a lower TFE sample size. The distribution of GFE overestimates the emission rate of the majority of facilities by using emission factors and component counts for pneumatics and component leaks; however, the emissions during episodic liquid unloading are underestimated by the GFE relative to the measurement methods used in this study.

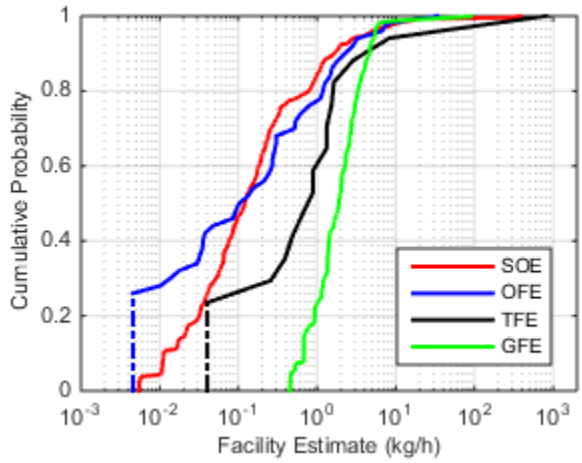

Figure S3: Cumulative distributions of facility-level methane emission estimates including 261 Study Onsite Estimates (SOEs), 51 OTM33a Facility Estimates (OFEs), 17 Tracer Facility

Estimates (TFEs) and 261 GHGRP Facility Estimates (GFEs). Dashed vertical line segments represent the minimum, non-zero estimate for each method.

S2. Onsite Measurement

Onsite Measurement Protocol

Onsite measurement protocol used in this study is documented in Annex 4 Onsite Detection and Measurement Protocol of the final report for RPSEA/NETL contract no 12122-95/DE-AC26-07NT42677 (Zimmerle et al., 2016).

Onsite Direct Measurement Data

Onsite direct measurements (ODMs) are reported in the excel workbook accompanying this supporting information including the facility ID, measurement team, measurement date, equipment category, component category, emission rate as measured by Hi Flow®, and emission rate exception notes (if any). Emission rate exceptions indicate notes about the measurement and include:

(1) Incomplete Capture – Emissions were not completely captured by the device enclosure during sampling, as observed by an onsite observer using optical gas imaging (OGI).
(2) Observed Not Measured – Emissions observed with OGI or laser detector by onsite measurement team, but not measured with high-flow instrument due to safety or accessibility
(3) Above Range – Measured value is above the range of the high-flow instrument (8.00 SCFM or 480 SCFH, equivalent to 9.2 kg/hr).
(4) Below Range – Measured value is below the range of the instrument (0.05 SCFM or 3 SCFH, equivalent to 0.058 kg/hr).(Bacharach, 2015)
(5) Non-Detect – No emissions observed with OGI or laser detector by the onsite measurement team.

Instrumentation

Emissions were detected utilizing a Heath Consultants RMLD™ handheld laser detector, or using optical gas imaging (OGI) cameras such as FLIR™ or EyeCGas cameras. Emissions were measured utilizing a Bacharach HiFlow™ sampler. The Bacharach instrument was calibrated each day using calibration gasses and following the methodology specified by the manufacturer. The reported sensor transition failure (Howard, 2015) is exacerbated by higher VOC to methane ratios and was not expected to impact the study results due to the dry gas production in the study

area with low VOC to methane ratio. A continuous QA check where the methane concentration in the HiFlow exhaust is sampled was not performed in this study. The methane concentration from the measurements can be calculated as a QA check assuming 8scfm motivated flow. This shows out of 299 measurements, only three measurements at 5% transition point and twenty measurements above the 5% transition point, which provides some evidence the transition failure (which would appear as saturation at the 5% mixing ratio) did not occur.

Measurement Teams

Two onsite measurement teams were deployed in this study. One team was a contractor, AECOM, equipped with an OGI camera for leak detection. One team was a measurement team provided by a study partner, equipped with both an RMLD and OGI camera for leak detection. Both teams were equipped with a Bacharach HiFlow sampler for leak quantification. Both measurement teams were provided the same protocol, namely using the OGI cameras and/or RMLD to locate all emission sources on the pad, then using the HiFlow to quantify all emission sources which were safe and accessible for measurement. In accessible or unsafe emission sources were to be noted as observed not measured. Data from the two teams is considered equal in the analysis presented in the main body of this paper, however a comparison of the two onsite measurement teams is provided here.

One team made 220 emission observations at 112 facilities visited in 15 measurement days. The other team made 102 emission observations at 190 facilities visited in 14 measurement days. On average, Team A made approximately 4 times the number of emission observations on a per facility basis (Team A 1.96 observations per facility, Team B 0.54 observations per facility). Team B made zero emission observations at a higher fraction of sites than Team A (66% vs 24% of facilities visited), and one or more emission observations at a lower fraction of sites than Team A (Figure S4).

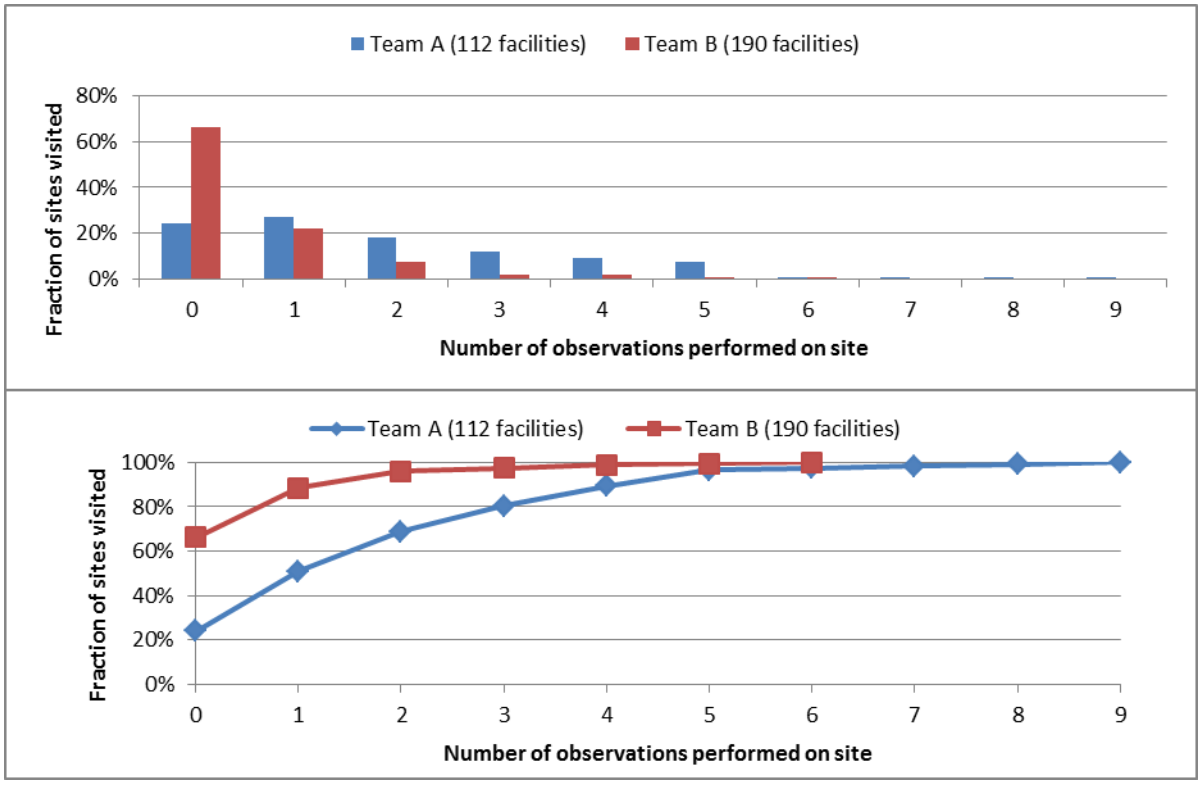

Figure S4: Comparison (top – histogram, bottom – empirical CDF) of number of observations per site made by two different onsite teams.

The highest emission rate measured by Team B was 0.51 kg/h from a PRV on a compressor. The top 10% of measurements made by Team A were above 0.51 kg/h (Figure S5), all of which were on tanks (9 measurements) or compressors (11 measurements). Collectively these 20 measurements accounted for 79% of cumulative emissions measured by Team A (Figure S6). The top 4 measurements (all on tanks) account for 60% of cumulative emissions measured by Team A, and include 2 measurements marked as incomplete capture indicating the actual emission rates were greater than the measured value. The fraction of observations in each emission rate exception category (Below HiFlow Range, Incomplete Capture, Observed Unmeasured, None) were nearly equally distributed between teams (Figure S7), however Team B made a slightly higher fraction of below range observations than Team A (60% vs 54% of observations), as well as a higher fraction of "observed not measured" sources (11% vs 5% of observations). Team B made no direct measurements of emissions from tanks; however 10 emission sources from tanks were reported as "observed not measured" sources by Team B. The differences between measurement teams in both the empirical CDF of emission rates (Figure S5)

and the cumulative emissions estimates (Figure S6) may be attributed to the lack of emission rate estimates from these 10 unmeasured sources (10% of emission observations made by Team B). Although Team B measurements lack these important high emission rate sources, a two sample k-s test performed at the 0.1% significance level ($\alpha = 0.001$) does not reject the null hypothesis that the two measurement distributions are from the same underlying population.

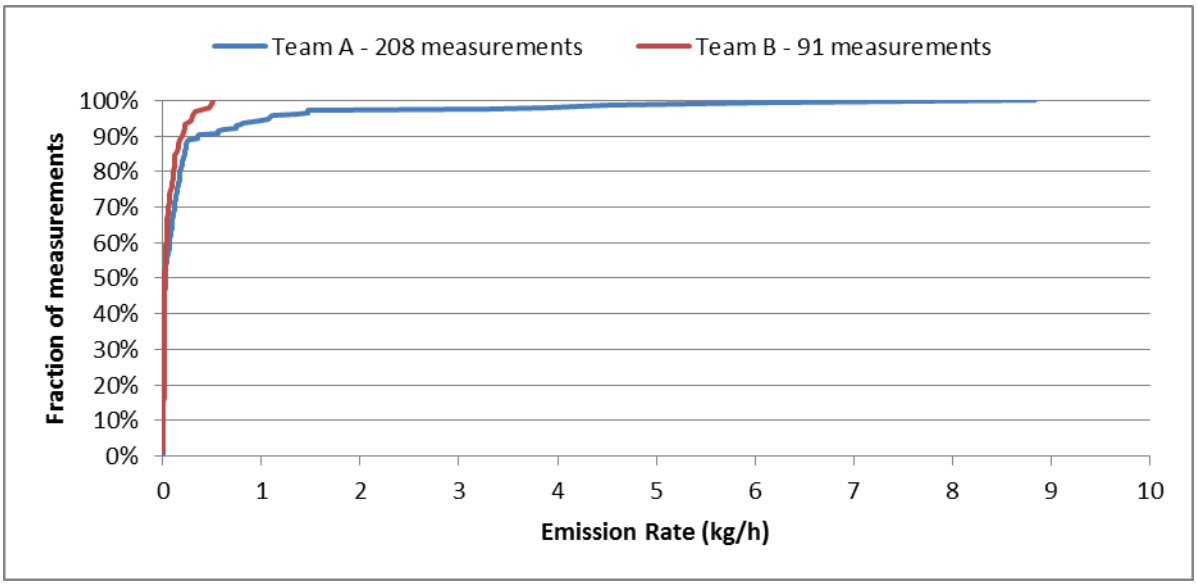

Figure S5: Comparison (Empirical CDF) of emission rate from measurements performed by two onsite measurement teams. Top 20 measurements (10%) performed by Team A were higher rates than largest measurement performed by Team B, 0.51kg/h.

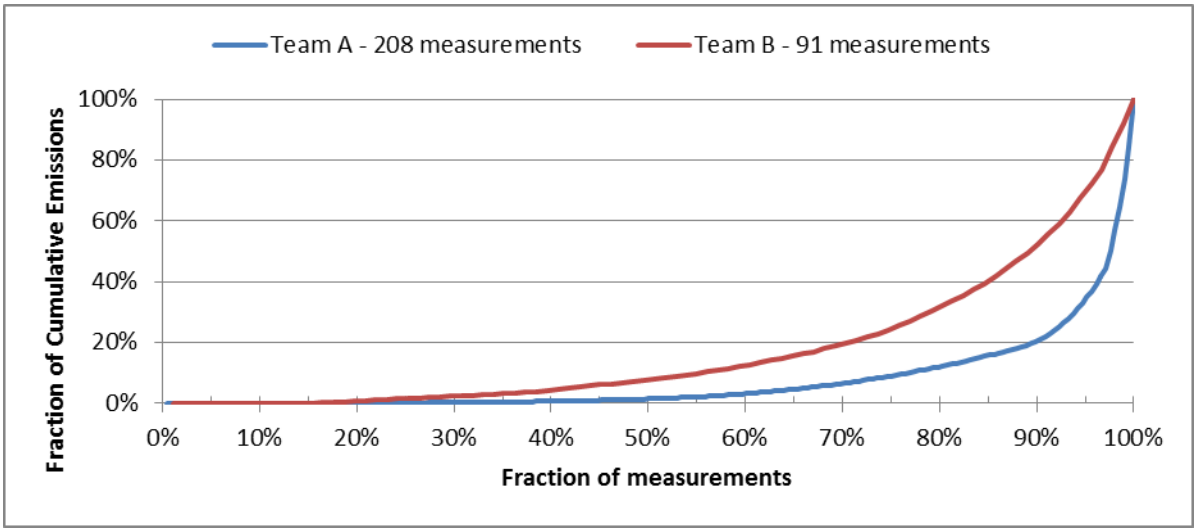

Figure S6: Comparison (Empirical CDF) of cumulative emission rate from measurements performed by two measurement teams. Top 20 measurements (10%) performed by Team A account for 79% of cumulative emissions measured by the team.

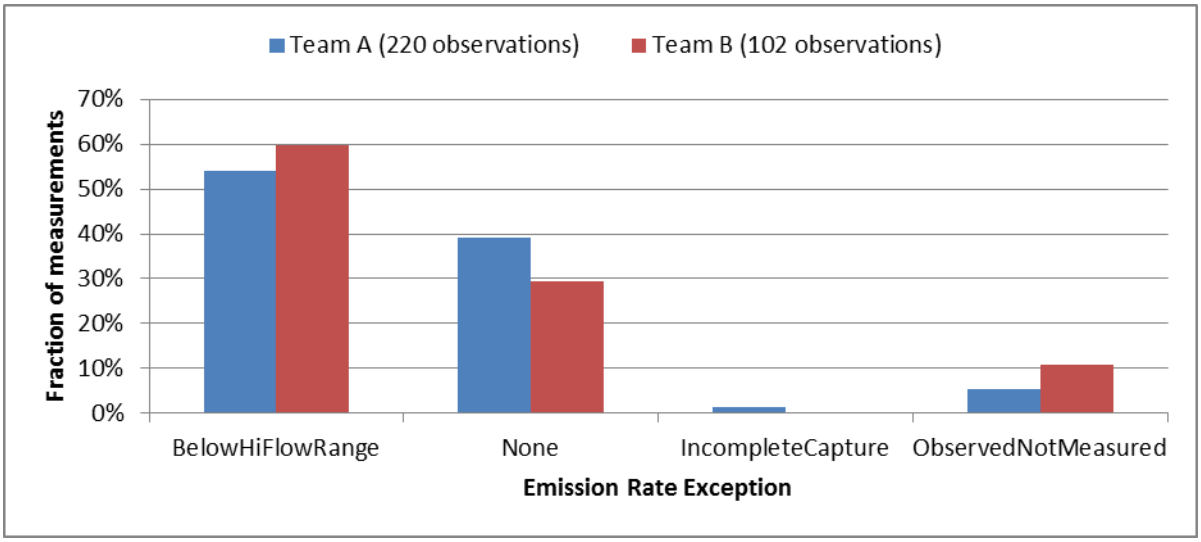

Figure S7: Comparison of emission rate exceptions from two onsite teams. All tank emission sources observed by Team B are included in "Observed Not Measured" category (i.e. no tank measurements were performed by Team B team), which may account for difference in cumulative emissions estimated by the two teams.

S3. Study Onsite Estimates

Calculation Methodology

A study onsite estimate (SOE) and uncertainty estimate was developed for each facility visited by an onsite team. The SOE provides an estimate of facility-level emissions, as well as the proportion of emissions contributed by each SOE category. This allows an analysis to be performed in which sites are grouped by which category is the largest contributor to the SOE. SOEs are composed of ODMs and engineering estimates of four unmeasured source categories developed using a Monte-Carlo model.

- Observed/Unmeasured Sources – Sources which were identified by the onsite team but not measured due to limited accessibility or safety concerns. The most common observed not measured source was inaccessible tank vents.
- Pneumatic Venting – Pneumatic devices at production well pads in the study area were observed to be primarily intermittent venting devices installed in separator level control applications. The actuation of the devices was observed to be infrequent, often with no actuations occurring while measurement teams were onsite. Due to resource constraints, onsite teams did not attempt to measure pneumatic devices which vent under routine operation.
- Compressor Exhaust – Exhaust of natural gas powered compressors includes methane due to incomplete combustion, often referred to as methane "slip". No direct measurements of combustion gasses were made during the study, but study partners provided previously completed stack tests for Caterpillar 3300 series engines.
- Vented Liquid Unloadings – Liquid unloadings are a routine operation in which natural gas is vented to the atmosphere for a controlled period of time, usually achieved by routing the gas flow to the produced water tank instead of to the sales line, in order to reduce back pressure on the well and enhance the removal of liquids accumulated in the well bore.

Each engineering estimate is made iteratively using a Monte Carlo model outlined in Figure S8. Activity data and emission rate data sources for each engineering estimate are listed in Table S3. The study onsite estimate $\dot{m}_{j,i}$ for facility j on Monte Carlo iteration i is calculated as

$$\dot{m}_{j,i} = \dot{m}_{ODM} + \dot{m}_{Unmeasured,i} + \dot{m}_{Pneumatics,i} + \dot{m}_{Compressor,i} + \dot{m}_{Unloadings,i} \qquad \text{Eq. 1}$$

For each facility SOE the mean, median, and two-sided 95% confidence bounds are calculated empirically from Monte-Carlo simulation results.

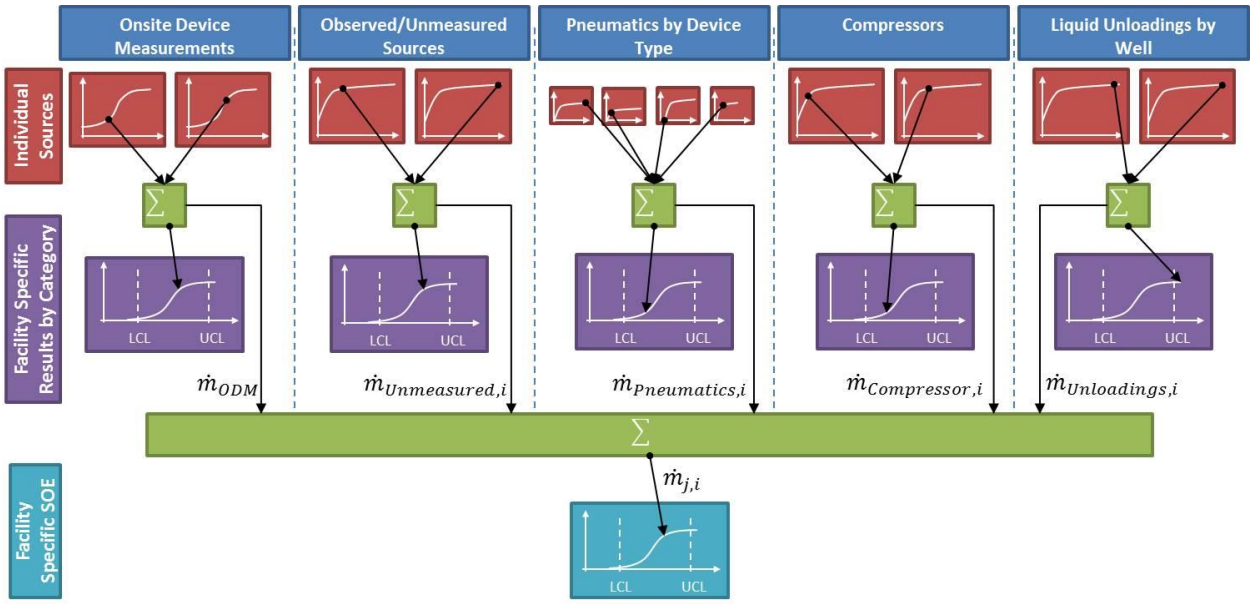

Figure S8: Calculation of study onsite estimate and facility specific results by source category.

Table S3: Data sources for engineering estimates used to develop comprehensive SOE.

Engineering Estimate	Activity Data Source	Emission Rate Data Source
(1) Observed unmeasured sources	Notes from onsite measurement team in this study	ODMs from this study on the same equipment category
(2) Gas powered pneumatic valves and pumps under routine operation	Partner provided device counts	Mid-Con Data from (Allen, Pacsi, et al., 2015)
(3) Engine exhaust	Partner provided compressor inventory	CAT 3300 series engines - Partner provided stack test data Non-CAT 3300 series engines - Following 40 CFR 98.233(z)
(4) Liquid unloadings	Manual – Partner provided unloading inventory during study period including date and duration Plunger – Partner provided annualized count and average duration	Mid-Con Data from (Allen, Sullivan, et al., 2015)

Onsite Direct Measurements – The total methane emissions associated with onsite direct measurements (ODMs) at a given facility is calculated for iteration i as

$$\dot{m}_{ODM,i} = \sum_{k=1}^{N} f_i * ODM_k \qquad \text{Eq. 2}$$

where N is the number of valid measurements at the particular facility and ODM_k is the emission rate of an individual ODM. f_i is a factor drawn from a normal distribution (mean = 1, σ = 0.05) each iteration to account for instrument uncertainty specified in the Bacharach HiFlow ® manual (+/- 10%).

ODMs marked with a below range emission rate exception were included in the ODM sum. Instrument uncertainty of below range measurements was increased linearly from +/- 10% at the high-flow lower detection limit to +/-100% at a reported emission rate of zero.

Invalid measurements indicated by incomplete capture or above range emission rate exceptions are not included in this sum and are instead included as a modeled emission source in the summation of the observed/unmeasured source category.

Observed/Unmeasured – The total methane emissions associated with sources observed by the study team but not measured due to limitations such as poor accessibility or unsafe conditions at a given facility is calculated for iteration i as

$$\dot{m}_{Unmeasured,i} = \sum_{k=1}^{N} \dot{m}_k \qquad \text{Eq. 3}$$

where N is the number of observed/unmeasured sources at the particular facility and $\dot{m}_k$ is the emission rate of an individual source k sampled from the rate distribution of ODMs of the same equipment category as the unmeasured source. When resampling from measurements for "observed/unmeasured" sources, all measurements on the same equipment type (including incomplete capture and over-range) are included in the distribution from which resampling occurs. If, at random, an incomplete capture or over-range measurement is drawn from the distribution, it is then modeled on that iteration as such.

Invalid measurements indicated by incomplete capture or above range emission rate exceptions are also included in this category in addition to observed/unmeasured sources. These sources with emission rate exceptions are modeled using a right triangular probability distribution extending from maximum probability at the measured value to a probability of zero at 2x the measured value. The factor of 2x the measured value was selected during the analysis assuming that the measurement team would not have attempted to measure a device emitting greater than 480 SCFH (2 x the upper range of the Hi-Flow®) or would have captured at least half the emissions during an incomplete capture.

Pneumatic Venting Devices and Chemical Injection Pumps – Counts of pneumatic powered chemical injection pumps and pneumatic devices classified as continuous high-bleed, continuous low-bleed, or intermittent-bleed devices were provided by study partners on a per well basis. The total methane emissions associated with routine operation of pneumatic devices at a given facility is estimated for iteration i as

$$\dot{m}_{Pneumatics,i} = \sum_{T} \sum_{k=1}^{N} (\dot{Q}_{T,k} * X_{CH_4} * 19.2) \qquad \text{Eq. 4}$$

where N is the number of devices of type T at the particular facility and $\dot{Q}_{T,k}$ is the whole gas emission rate of an individual device k sampled from the distribution of type T. Whole gas emission rates in SCFH were sampled for each intermittent bleed, low bleed and high bleed pneumatic device using Mid-Continental data from (Allen, Pacsi, et al., 2015). Whole gas emission rates from chemical injection pumps were sampled from (Allen et al., 2013). Conversion from whole gas volumetric flow rates to methane flow rates was performed using the methane mol fraction X_{CH_4} in the gas composition of the well. In cases where gas composition was not available for a specific well, the methane mol fraction was sampled from a distribution of gas composition from other wells in the study area. A conversion factor of 19.2 g methane/scf was applied to convert the volumetric methane flowrate to a mass flowrate.

Compressor Exhaust – An estimate of methane in the exhaust of operating compressors as a result of incomplete combustion is included in the SOE. The total methane emissions from compressor exhaust at a given facility is estimated for iteration i as

$$\dot{m}_{Compressor,i} = \sum_{k=1}^{N} \dot{m}_k \qquad \text{Eq. 5}$$

where N is the number of operating compressors at the facility and $\dot{m}_k$ is the emission rate of an individual compressor k.

If the compressor was driven by a Caterpillar 3300 series the emission rate for the engine was estimated as

$$\dot{m}_{k,CAT3300} = E_{CH_4} * Load * HP \qquad \text{Eq. 6}$$

where the emission rate E_{CH_4}in g/hp-hr was sampled from a distribution of stack test data taken on Caterpillar 3300 series engines provided by study partners, $Load$ is the percent load the engine was operating at, and HP was the rated power of the engine.

If the compressor was driven by another engine make or model (which stack test data was unavailable for) the methane emission rate was estimated using Greenhouse Gas Reporting Program Subpart W methodology assuming a combustion efficiency of 99.5% (0.5% of fuel remains unburnt in exhaust).

$$\dot{m}_{k,SubW} = \dot{m}_{fuel} * Load * HP * 0.005 \qquad \text{Eq. 7}$$

Liquid Unloading – Vented liquid unloadings are routine operations where back pressure on the well resulting from liquids that have accumulated in the well bore is reduced by venting the well directly to atmosphere. This operation may be assisted by a plunger, when present, and may be triggered automatically by a control system or manually by an operator. The frequency of liquid unloading at a given well is not consistent and may vary over days and longer. Because of associated measurement challenges, it was decided not to directly measure vented liquid unloadings. The episodic emissions during vented liquid unloading increase facility emissions for a period of time (duration 1 minute to 24 hours), but do not represent an anomalous condition. Such increased emissions are not indicative of annual or daily average emission rates from the same facilities.

At one facility (Facility ID = 371) onsite measurement was paired with OTM33A and Dual Tracer methods during a manual liquid unloading. The emission rate during the manual liquid unloading was modeled in the SOE using the whole gas rates reported for manual unloadings

marked as mid-continental by (Allen, Sullivan, et al., 2015). The whole gas emission rate, $\dot{Q}$, was converted to a methane mass flow rate using the gas composition of the well, X_{CH_4}, and the conversion factor of 19.2 g methane/scf.

$$\dot{m}_{Manual\ Unloading} = \dot{Q} * X_{CH_4} * 19.2 \qquad \text{Eq. 8}$$

Plunger liquid unloadings were modeled as a time averaged emission rate over a one hour sampling period in order to make the SOE as comparable as possible to the downwind measurements. The model accounts for uncertainty in the number of plunger unloadings that occurred during the period, the duration of each plunger unloading, and the emission rate during plunger unloading as illustrated in Figure S9.

The timing of plunger unloadings which occurred during a one hour timespan was modeled using the annual number of plunger unloadings vented to atmosphere and the average duration provided by study partners. At facilities where onsite measurements were paired with tracer release or OTM33A methods, study partners provided an event count specific to the day of measurement in order to improve the accuracy of the SOE in comparisons to contemporaneous estimates. Starting at the beginning of the one-hour time period, a time t_1 was selected from a uniform distribution between 0 and the maximum time between unloadings. This time represents the start of the first unloading during the sampling period. The duration of the first unloading was then selected from a triangular distribution of +/-50% about the average duration. Starting at the end of the first unloading, a time t_2 was selected from a triangular distribution between the minimum and the maximum time between unloadings. The duration of the second unloading was then modeled as done for the first unloading. This method continued until the end of the one hour time span is reached.

The emission rate during plunger liquid unloading was then modeled using the whole gas rates reported for automated plunger unloadings marked as mid-continental by (Allen, Sullivan, et al., 2015). The whole gas emission rate, $\dot{Q}$, was converted to a methane mass flow rate using the gas composition of the well, X_{CH_4}, and the conversion factor of 19.2 g methane/scf.

$$\dot{m}_{Plunger\ Unloading} = \dot{Q} * X_{CH_4} * 19.2 \qquad \text{Eq. 9}$$

The time averaged emission rate was then calculated as

$$\dot{m}_{Plunger\ Unloading,Time\ Ave.} = \frac{\sum_{i=1}^{N} \dot{m}_{Plunger\ Unloading} * d_i}{1hr} \qquad \text{Eq. 10}$$

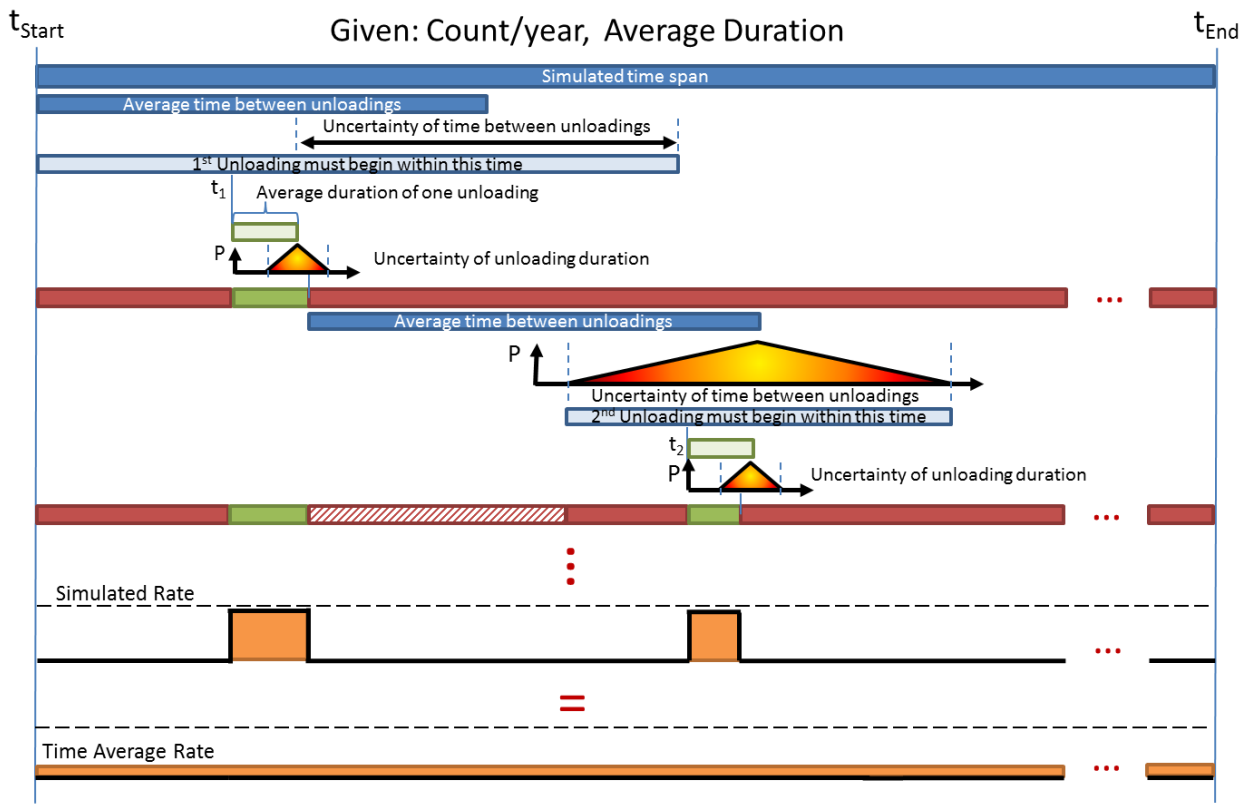

Figure S9: Modeling time averaged emissions from plunger liquid unloading during 1 hour measurement period using well specific annual unloading count and average duration provided by study partners and considering uncertainty using triangular probability distribution about average. Emission rate simulated using mid-continental data from (Allen, Sullivan, et al., 2015). Time averaged emission rate during the measurement period calculated by integrating total emissions over simulated time span and dividing by length of sampling period, t_{End}-t_{Start}.

Activity Data

Activity data required to calculate the SOE and/or the GFE are included in the excel workbook accompanying this supporting information. The tables include:

1. Pneumatic device counts
2. Compressor characteristics for combustion emissions calculation
3. Manual liquid unloading data

4. Plunger liquid unloading data

Calculated SOEs

The calculated SOE is reported in the excel workbook accompanying this supporting information. The table includes the mean and empirical two-sided 95% confidence bounds (reported as the 2.5 and 97.5 percentiles) of the calculated SOE for each facility. The mean and confidence bounds of each model category are also reported.

Figure S10 separates 41 facilities with vented liquid unloading events occurring during the time of onsite measurement from 220 facilities without unloading events occurring during the time of onsite measurement. In both groups, emission distributions are skewed, meaning a small number of measurements are responsible for a significant portion of the total emissions. Simulated emissions from liquid unloadings account for 88% of the cumulative emission estimate from all 261 facilities. Onsite direct measurements and simulated emissions from observed, unmeasured sources (sources which were not accessible for direct measurements but observed to be emitting) account for 78% of the cumulative emissions estimate at facilities without simulated liquid unloadings in the SOE.

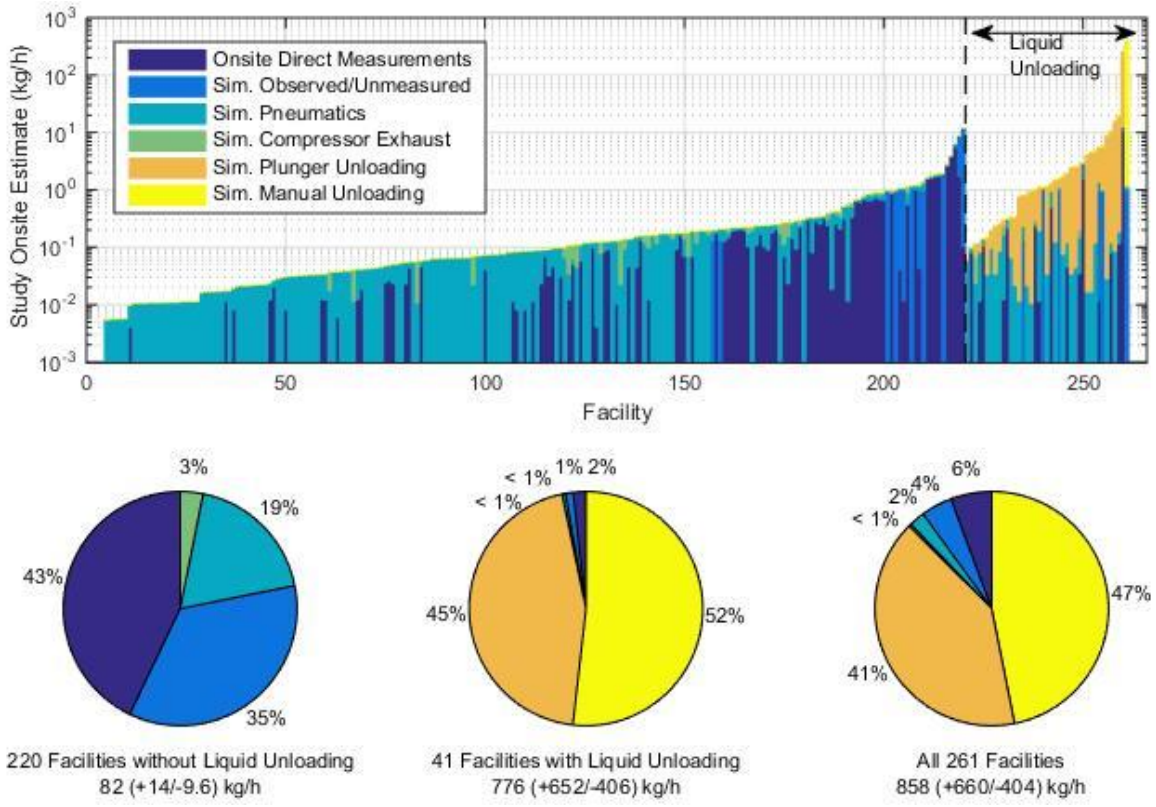

Figure S10: Study Onsite Estimate (SOE) for the 261 production facilities measured in the study area of the Fayetteville shale play during September-October 2015. Bars: Facilities are grouped by whether liquid unloadings are modeled in the SOE, and both groups illustrate skewed emission distributions where a few sites account for a large portion of emissions. Pie charts summarize the cumulative emissions from facilities without liquid unloadings, facilities with liquid unloadings, and all 261 facilities. Unloadings dominate cumulative emissions (88% from all 261 facilities), while only 6% of the cumulative estimated emissions were directly measured during the field campaign, however at 220 facilities without unloadings emissions from onsite direct measurements and simulated inaccessible (observed/unmeasured) sources account for 78% of cumulative emissions.

S4. Tracer Estimates

Tracer estimates are reported in (Yacovitch et al., n.d.) and included in the excel workbook accompanying this supporting information for 17 production facilities, including 3 facilities using the limit of detection method (LOD). 11 of these estimates were paired with OFE's and 16 were paired with SOEs.

S5. OTM33A Estimates

Facility-level methane emission estimates at production pads made during this campaign using OTM33A are reported for 51 facilities, including 13 zero based on transect (0BOT) facilities in the MC study area in (Robertson et al., 2017) and included in the excel workbook accompanying this supporting information. Nine of these estimates were paired with tracer release and 43 were paired with onsite measurement.

OTM33A measurements of partial facilities were not included in this analysis. In some cases more than one OTM33A facility-level measurement is reported at a facility. This work considers the average of these estimates and uncertainties calculated using quadrature as:

$$\overline{OFE} = \frac{\sum_{i=1}^{N} OFE_i}{N} \qquad \text{Eq. 11}$$

$$\overline{OFE_+} = \sqrt{\sum_{i=1}^{N} \left(\frac{OFE_{+,i}}{N}\right)^2} \qquad \text{Eq. 12}$$

$$\overline{OFE_-} = \sqrt{\sum_{i=1}^{N} \left(\frac{OFE_{-,i}}{N}\right)^2} \qquad \text{Eq. 13}$$

S6. GHGRP Facility-Level Estimates (GFE)

Owners or operators of facilities that emit 25,000 metric tons (expressed as carbon dioxide equivalent) or more of greenhouse gases (GHG) per year are required to report GHG data to EPA under the Greenhouse Gas Reporting Program (GHGRP). (Note that a GHGRP facility represents all production operations for one operator in a basin, and is not comparable to "production facility" terminology utilized in this study.) Annualized emissions from natural gas production facilities are reported by each operator aggregated to the basin level.

To make comparisons to the facility-level methane emissions estimates in this study, a facility emission estimate was also constructed following GHGRP (e-CFR: TITLE 40—Protection of Environment, n.d.) methods. To make facility estimates comparable to the other methods used in this study, the GHGRP calculations are performed for each facility individually, and are not aggregated across facilities to produce a study area estimate as required by GHGRP reporting. Since the GHGRP emission factors are not provided with CIs, the CI for GHGRP estimates includes only the variability due to uncertainty in well properties, equipment numbers, and activity counts where these data were not available from study partners and does not include variability due to uncertainty in the emission factors. Only the methane components of the GHGRP methodology were calculated and methane emissions were not converted to CO_2 equivalent.

Emission estimates from the following sections of EPA Title 40 Code of Federal Regulations Part 98 Subpart W were included in the estimate:

- 98.233(a) Natural gas pneumatic device venting.
- 98.233(c) Natural gas driven pneumatic pump venting.
- 98.233(f) Well venting for liquids unloadings.
- 98.233(r) Population count and emission factors.
- 98.233(z) Onshore petroleum and natural gas production and natural gas distribution combustion emissions.

The emissions estimates do not include other paragraphs which were not applicable to the equipment observed onsite, or the operational state of the facilities. These include:

- 98.233(d) Acid gas removal (AGR) vents.
- 98.233(e) Dehydrator vents.
- 98.233(g) Gas well venting during completions and workovers from hydraulic fracturing.
- 98.233(h) Gas well venting during completions and workovers without hydraulic fracturing.
- 98.233(i) Blowdown vent stacks.
- 98.233(j) Onshore production storage tanks.
- 98.233(k) Transmission storage tanks.
- 98.233(l) Well testing venting and flaring.
- 98.233(m) Associated gas venting and flaring.
- 98.233(n) Flare stack emissions.
- 98.233(o) Centirfugal compressor venting.
- 98.233(p) Reciprocating compressor venting.
- 98.233(q) Leak detection and leaker emission factors.
- 98.233(s) Offshore petroleum and natural gas production facilities.
- 98.233(w) EOR injection pump blowdown.

98.233(a) Natural gas pneumatic device venting & (c) Natural gas driven pneumatic pump venting.

Natural gas pneumatic device venting and pneumatic pump venting are calculated under paragraphs (a) and (c) as:

$$Mass_{t,i} = Count_t * EF_t * GHG_i * Conv_i * T_t \qquad \text{Eq. 14}$$

where:

$Mass_{t,i}$ = Annual total mass GHG emissions in metric tons CO_2e per year from a natural gas pneumatic device vent of type "t", for GHG_i;

$Count_t$ = Total number of natural gas pneumatic devices of type "t" (continuous high bleed, continuous low bleed, intermittent bleed, pneumatic pump);

EF_t = Population emission factors for natural gas pneumatic device venting listed in Tables W-1A for onshore petroleum and natural gas production facilities;

GHG_i = Concentration of GHG_i, CH4, or CO2 in natural gas;

$Conv_i$ = Conversion from standard cubic feet to metric tons CO2e equal to 0.000403 for CH4, and 0.00005262 for CO2;

T_t = Average estimated number of hours in the operating year the devices, of each type t, were operational. Default is 8760 hours.

This equation was modified in this study to determine a methane emission rate in g/hr as:

$$\dot{m}_t = Count_t * EF_t * X_{CH_4} * 19.2 \qquad \text{Eq. 15}$$

Where:

$\dot{m}_t$ = Methane emission rate in g/hr;

$Count_t$ and EF_t are the count and emission factor for each device type as defined above;

X_{CH_4} = Average concentration of methane in the gas at the facility;

19.2 = Conversion factor from scf methane to g methane.

This study used device counts provided by study partners for each well, and the following emission factors (SCFH per component) from Table W-1A for the Western U.S. (Arkansas is Western US per Table W-1D).

Table S4: Pneumatic device emission factors from Table W-1A of 40 CFR Part 98 Subpart W for the Western U.S.

Device Type	EF (SCFH)
Low Continuous Bleed Pneumatic Device Vents	1.39
High Continuous Bleed Pneumatic Device Vents	37.3
Intermittent Bleed Pneumatic Device Vents	13.5
Pneumatic Pumps	13.3

98.233(f) Well venting for liquids unloadings.

Manual unloading

Total emissions for well venting for manual liquids unloadings are calculated using Calculation Methodology 2 under paragraph (f) as:

$$E_{s,n} = \sum_{p=1}^{W} \left[V_p \times \left((0.37 \times 10^{-3}) \times CD_p^2 \times WD_p \times SP_p \right) + \sum_{q=1}^{V_p} \left(SFR_p \times \left(HR_{p,q} - 1.0 \right) \times Z_{p,q} \right) \right] \quad \text{Eq. 16}$$

Where:

$E_{s,n}$ = Annual natural gas emissions at standard conditions, in cubic feet/year;

W = Total number of wells with well venting for liquids unloading for each sub-basin;

0.37×10^{-3} = {3.14 (pi)/4}/{14.7*144} (psia converted to pounds per square feet);

CD_p =Casing internal diameter for each well, p, in inches;

WD_p = Well depth from either the top of the well or the lowest packer to the bottom of the well, for each well, p, in feet;

SP_p = Shut-in pressure or surface pressure for wells with tubing production and no packers or casing pressure for each well, p, in pounds per square inch absolute (psia) or casing-to-tubing pressure of one well from the same sub-basin multiplied by the tubing pressure of each well, p, in the sub-basin, in pounds per square inch absolute (psia);

V_p = Number of vents per year per well, p;

SFR_p= Average flow-line rate of gas for well, p, at standard conditions in cubic feet per hour;

$HR_{p,q}$ = Hours that each well, p, was left open to the atmosphere during unloading, q;

1.0 = Hours for average well to blowdown casing volume at shut-in pressure;

$Z_{p,q}$ = If $HR_{p,q}$ is less than 1.0 then $Z_{p,q}$ is equal to 0. If $HR_{p,q}$ is greater than or equal to 1.0 then $Z_{p,q}$ is equal to 1.

This equation was modified to calculate an average rate during the one manual unloading observed by the onsite team during contemporaneous measurement with the OTM33A and tracer methods as:

$$\dot{m} = \frac{\left[\left((0.37 \times 10^{-3}) \times CD^2 \times WD \times SP\right) + (SFR \times (HR - 1.0) \times Z)\right] * X_{CH_4} * 19.2}{HR} \quad \text{Eq. 17}$$

Where:

X_{CH_4} = Average concentration of methane in the gas at the facility;

19.2 = Conversion factor from scf methane to g methane.

Plunger Unloading

Total emissions for well venting from plunger liquids unloadings are calculated using Calculation Methodology 3 under paragraph (f) as:

$$E_{s,n} = \sum_{p=1}^{W} \left[V_p \times \left((0.37 \times 10^{-3}) \times TD_p^2 \times WD_p \times SP_p\right) + \sum_{q=1}^{V_p} \left(SFR_p \times \left(HR_{p,q} - 0.5\right) \times Z_{p,q}\right) \right] \quad \text{Eq. 18}$$

Where:

$E_{s,n}$ = Annual natural gas emissions at standard conditions, in cubic feet/year;

W = Total number of wells with well venting for liquids unloading for each sub-basin;

0.37×10^{-3} = {3.14 (pi)/4}/{14.7*144} (psia converted to pounds per square feet);

TD_p =Tubing internal diameter for each well, p, in inches;

WD_p = Tubing depth to plunger bumper for each well, p, in feet;

SP_p = Flow-line pressure for each well, p, in pounds per square inch absolute (psia), using engineering estimate based on best available data;

V_p = Number of vents per year per well, p;

SFR_p= Average flow-line rate of gas for well, p, at standard conditions in cubic feet per hour;

$HR_{p,q}$ = Hours that each well, p, was left open to the atmosphere during unloading, q;

0.5 = Hours for average well to blowdown tubing volume at flow-line pressure;

$Z_{p,q}$ = If $HR_{p,q}$ is less than 0.5 then $Z_{p,q}$ is equal to 0. If $HR_{p,q}$ is greater than or equal to 0.5 then $Z_{p,q}$ is equal to 1.

This equation was modified to calculate the average methane emission rate due to plunger unloadings from each well using the average duration, HR_p, and annual count, V_p, provided by study partners as:

$$\dot{m} = \frac{\left[V_p \times \left((0.37 \times 10^{-3}) \times TD_p^2 \times WD_p \times SP_p\right) + V_P \times \left(SFR_p \times \left(HR_p - 0.5\right) \times Z_p\right)\right] * X_{CH_4}}{8760} \cdot \quad \text{Eq. 19}$$

98.233(r) Population count and emission factors.

Total emissions from valves, connectors, open-ended lines, and pressure relief valves at natural gas production facilities are calculated using emission factors and average component counts by major equipment piece under paragraph (r) as:

$$E_{s,i} = Count_s * EF_s * GHG_i * T_s \quad \text{Eq. 20}$$

Where:

$E_{s,i}$ = Annual volumetric GHG emissions at standard conditions from each component type in cubic feet.

$Count_s$ = Total number of this type of emission source at the facility. For onshore petroleum and natural gas production, average component counts are provided by major equipment piece in Tables W-1B.

EF_s = Population emission factor for the specific component type, as listed in Table W-1A

GHG_i = For onshore petroleum and natural gas production facilities, concentration of GHG_i, CH_4 or CO_2, in produced natural gas.

T_s = Average estimated time that each component type associated with the equipment leak emission was operational in the calendar year, in hours, using engineering estimate based on best available data.

This equation was modified to calculate the average methane emission rate from valves, connectors, open-ended lines, and pressure relief valves for each facility as:

$$\dot{m} = \sum_s (Count_s * EF_s) * X_{CH_4} * 19.2 \qquad \text{Eq. 21}$$

Where:

$\dot{m}$ = Mass flow rate of methane emissions from valves, connectors, open-ended lines, and pressure relief valves at the facility.

X_{CH_4}= Average methane concentration in gas composition at the facility.

19.2 = Conversion factor from scf methane to g methane.

This study used emission factors (SCFH per component) from Table W-1A and average component counts by major equipment piece from Table W-1B for the Western U.S. (Arkansas is Western US per Table W-1D) shown here in Table S5 and Table S6 respectively. Major equipment counts were assumed as 1 well head, 1 separator, 1 meter/piping, 0 in-line heaters, and 0 dehydrators per well. Compressor counts were used from data provided by study partners.

Table S5: Population emission factors from Table W-1A of 40 CFR Part 98 Subpart W for the Western U.S.

Device Type	EF (SCFH)
Valve	0.121
Connector	0.017
Open-ended Line	0.031

Pressure Relief Valve	0.193

Table S6: Average component counts by major equipment piece from Table W-1B of 40 CFR Part 98 Subpart W for the Western U.S.

Major Equipment	Valves	Connectors	Open-ended Lines	Pressure Relief Valves
Wellheads	11	36	1	0
Separators	34	106	6	2
Meters/Piping	14	51	1	1
Compressors	73	179	3	4
In-line heaters	14	65	2	1
Dehydrators	24	90	2	2

98.233(z) Onshore petroleum and natural gas production and natural gas distribution combustion emissions.

Total methane emissions from combustion at natural gas production facilities are calculated under paragraph (z) for combustion units that combust field gas as:

$$E_{a,CH_4} = V_a * (1 - \eta) * Y_{CH_4} \qquad \text{Eq. 22}$$

Where:

E_{a,CH_4} = Contribution of annual CH_4 emissions from portable or stationary fuel combustion sources in cubic feet, under actual conditions.

V_a= Volume of gas sent to combustion unit in cubic feet, during the year.

η = Fraction of gas combusted for portable and stationary equipment determined using engineering estimation. For internal combustion devices, a default of 0.995 can be used.

Y_{CH_4} = Concentration of methane constituent in gas sent to combustion unit.

This equation was modified to calculate an average methane emission rate as

$$\dot{m} = \frac{V_a * (1-\eta) * Y_{CH_4} * 19.2}{8760} \quad \text{Eq. 23}$$

The volume of gas sent to the combustor was calculated as:

$$V_a = \frac{HR * HP * Load * OP}{HV} \quad \text{Eq. 24}$$

Where:

HR = Engine heat rate in BTU/hp-hr.

HP = Engine horsepower.

$Load$ = Average engine percent load.

OP = Annual operating hours of the compressor as provided by study partners.

HV = Fuel heating value of 992.25 BTU/scf (the average of gas composition data provided by study partners)

Calculated GHGRP Methane Emissions Estimates.

Methane emissions estimates were calculated using the modified GHGRP methods described in this section for each facility an SOE was calculated for. If data was unavailable from study partners for a particular parameter, a value was sampled from available partner data using a Monte-Carlo model. The mean results calculated for each site are included in the excel workbook accompanying this supporting information.

Comparison of SOE to GFE

The SOE and GFE results are summarized in **Error! Reference source not found.** by omparable categories. The cumulative SOE (858 kg/h) is 16% larger than the cumulative GFE (742 kg/h) for these 261 facilities. Even though the cumulative GFE is within the confidence bounds of the SOE, the emissions by category are quite different (Table 4). For instance, cumulative ODMs and simulated observed unmeasured sources in the SOE (86 kg/h) are half of the GFE for component leaks (172 kg/h) using GHGRP emission factors and average component counts by major equipment piece. Cumulative emissions from pneumatics in the SOE (19 kg/h) are 5.2% of the cumulative GFE estimate (356 kg/h) using device counts and GHGRP emission factors. Emissions from incomplete combustion in compressors are calculated using very similar

assumptions for both SOE and GFE and thus account for 2.9 kg/h in the SOE and 3.5 kg/h calculated following 98.233(z). Cumulative emissions from liquid unloadings in the SOE (751 kg/h), which utilizes recent time-series measurements (Allen, Sullivan, et al., 2015) and was reasonably confirmed by contemporaneous tracer measurements (statistically similar estimates at all four sites where these methods were paired), are roughly 3.5 times cumulative emissions from liquid unloadings in the GFE (211 kg/h). These differences in cumulative device-level emission estimates likely reflect differences between emission factors specific to the study area and nationally-averaged ones from GHGRP, and differences between operational practices.

Table S7: Cumulative study onsite estimate (SOE) and GHGRP facility estimate (GFE) by source category for 261 facilities with an SOE estimate.

SOE Category	SOE (kg/h)	GHGRP Category	GFE (kg/h)	SOE/ GFE
ODMs & Observed Unmeasured Sources	85.8	98.233(r) Component Leaks	172	50%
Pneumatics	18.7	98.233(a) Pneumatic devices & (c) Pneumatic pumps	356	5.2%
Combustion	2.9	98.233(z) Combustion	3.5	82%
Liquid Unloading	751	98.233(f) Well venting	211	356%
Total	858	Total	742	116%

S7. Cumulative Emission Estimates and Variance Weighted Least Squares Regression

Cumulative emissions estimates are reported with error bounds corresponding to 95% confidence intervals. Where cumulative emission estimates are reported for SOE, the uncertainty in the cumulative estimate was calculated by bootstrapping facility specific results. This accounts for skewed emissions estimates at the facility level (distributions produced by the Monte-Carlo SOE model for each facility are not normally distributed). Uncertainty estimates in cumulative TFE and cumulative OFE were calculated by adding the upper and lower uncertainties in quadrature.

Since the facility-level emissions data spans several orders of magnitude, minimizing the residual between the largest data point and a linear fit tends to minimize the total residuals in a normal (and orthogonal) least squares regression without much attention paid to the lower-magnitude data. An alternative regression method, variance weighted least squares (VWLS) as described by (Neri et al., 1989) is used here. Weights were applied using the variance in both the x and y directions, where the weights are the reciprocal of the square of the standard deviation of each data point. The y intercept was constrained to zero in the regression since no bias between methods is expected at sites with zero emissions. Sites with no error bounds reported (facilities reported as zero based on transect by the OTM33a measurement team) were not included in the regression.

Confidence intervals for the variance weighted least squares regressions were calculated using a bootstrap method where each data point is resampled and a regression is performed on the resampled data. When resampling the SOE a new value was drawn from a site specific distribution generated by the SOE Monte-Carlo model. This allows the non-Gaussian distribution of emission estimates from the SOE at each facility to be accounted for in the VWLS regression. When resampling the TFE or OFE data were sampled from a normal distribution with the 95% confidence bounds reported by the method at each site.

The goodness of fit between data and the VWLS regression is using the coefficient of determination, R^2. It is important to note R^2 only accounts for variance in the y-axis, and does not reflect the variance in the data on the x-axis. It is also important to note that R^2 may be negative due to performing the regression without an intercept (i.e. fixing intercept at $x = 0$, $y = 0$ and only varying slope during linear regression).

All Paired Measurements

All facilities where SOE was paired with a downwind method are shown in Figure S12. VWLS regression in panel (a) indicates TFE is statistically higher than SOE with TFE to SOE ratio = 1.7 $\left(\frac{+0.54}{-0.4}\right)$ even though confidence intervals overlap at 15 of 16 individual facilities. VWLS regression in panel (b) indicates OFE is lower although not statistically different than SOE with OFE to SOE ratio = 0.49 $(\frac{+0.58}{-0.20})$ when the regression is performed including all paired measurements even though confidence intervals overlap at only 28 of 43 individual facilities.

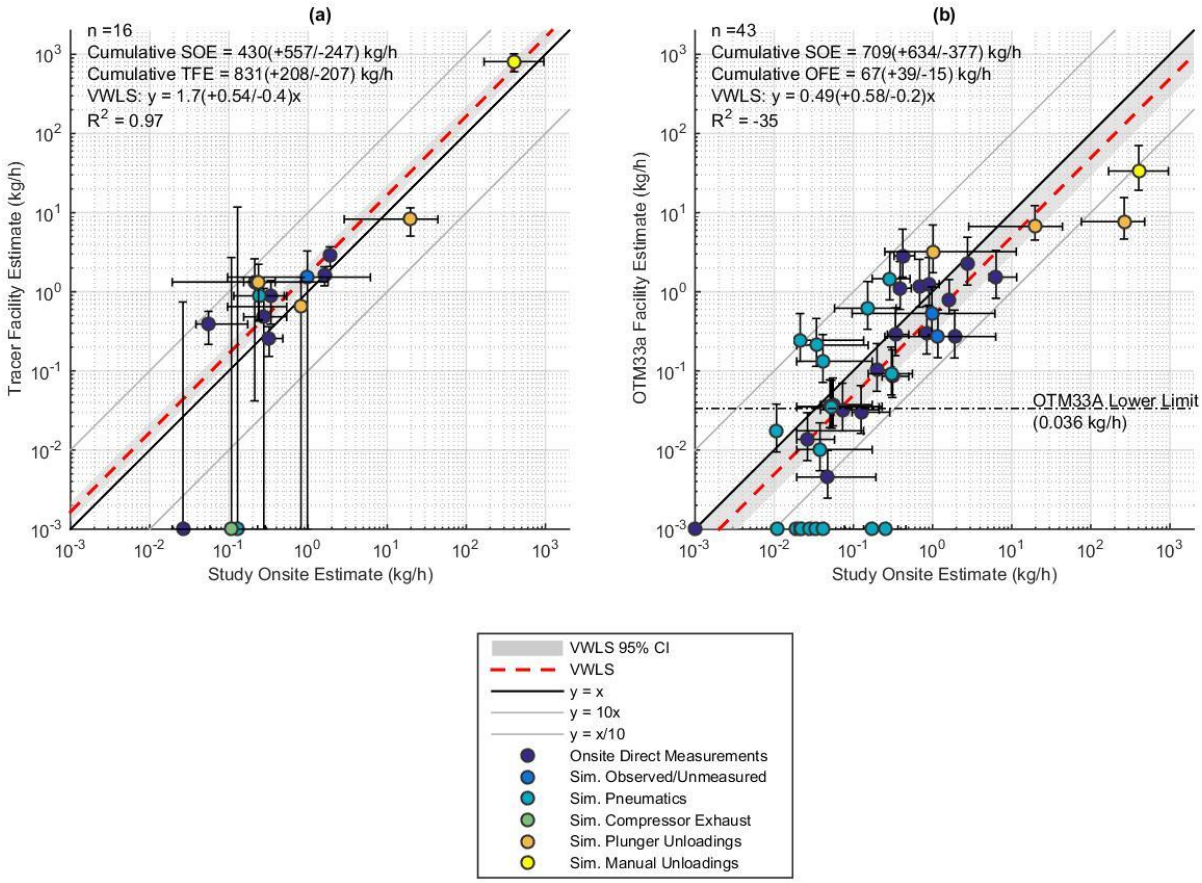

Figure 11: (a) 9 Tracer facility estimates (TFE) and (b) 20 OTM33A facility estimates (OFE) each compared to paired study onsite estimates (SOE) at subset of facilities where continuous emissions measured by onsite direct measurements or documented as an observed unmeasured source contribute the most emissions to the SOE. Comparisons are made using orthogonal least

squares regression (OLS). Markers are colored by the largest category contributing to SOE and shown with 95% confidence intervals. Measurements indicating zero emissions are plotted on log-log axes at 10^{-3} kg/h for display purposes only. The lower detection limit of TFE is facility-specific and shown as the upper confidence bar of measurements indicating zero emissions. The lower detection limit of OTM33A is not facility-specific. Correlation coefficient (Pearson R) indicates a stronger correlation in TFE vs SOE comparison than in OFE vs SOE comparison. OLS indicates TFE is higher than SOE ($TFE = 0.19(\frac{+0.66}{-0.37}) + 1.3(\frac{+1.2}{-1.1}) \times SOE$) in (a), while (b) shows OFE is lower than SOE ($OFE = 0.34\left(\frac{+0.37}{-0.58}\right) + 0.32(\frac{+1.1}{-0.2}) \times SOE$). OLS 95% confidence intervals calculated by bootstrapping the data and regression indicate methods are not statistically different. Higher coefficients of determination (R^2) in tracer vs SOE comparison than in OTM vs SOE comparison indicates the fit does not capture the wide variation seen in the data (most paired estimates are within factor of 10 illustrated by y = 10x, y = x/10 bounds).

Liquid Unloadings

Paired measurements at facilities where simulated emissions from liquid unloading is the largest category in the SOE are shown in Figure S12. Simulated emissions from liquid unloading is the largest SOE category at 4 facilities where SOE was paired with TFE, including the facility where a manual unloading was observed. At all four facilities TFE and SOE have overlapping confidence bounds indicating the estimates from the two methods are not statistically different. The cumulative estimates produced by the two methods at these four facilities are also not statistically different ($cumulative\, SOE = 424\,(+558/-247)$ kg/h, $cumulative\, TFE = 821\,(+/-207)$ kg/h). While the sample size is limited ($n = 4$), agreement between individual paired estimates as well as the cumulative TFE and SOE estimates at these facilities provides some validation of the liquid unloading model used in the Monte-Carlo to estimate these emission sources in the SOE.

Simulated emissions from liquid unloading is also the largest SOE category at 4 facilities where SOE was paired with OFE. The SOE and OFE are statistically different (confidence bounds do not overlap) at two of these four facilities, including the facility where a manual unloading was observed during the measurements. Although the confidence bounds of the VWLS regression include a slope of 1, the cumulative estimates produced by the two methods are statistically

different ($cumulative\ SOE\ =\ 688\ (+622/-376)$ kg/h, $cumulative\ OFE\ =\ 51\ (+38/-15)$ kg/h).

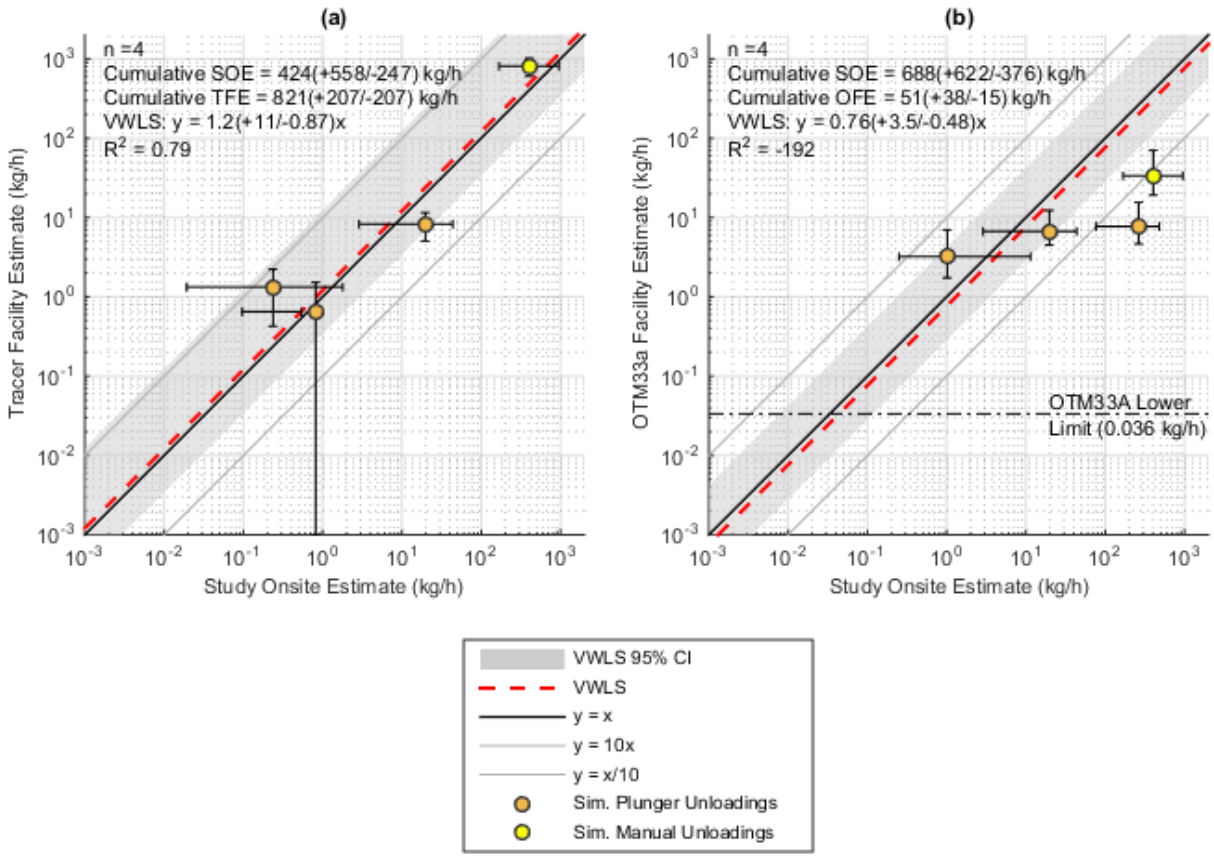

Figure S12: (a) 4 Tracer facility estimates (TFE) and (b) 4 OTM33A facility estimates (OFE) compared to paired study onsite estimates (SOE) at facilities where simulated emissions from liquid unloading is the largest source category in the SOE. Each panel includes a fit line from a variance weighted least squares regression (VWLS). The shaded region corresponds to 95% CI calculated by bootstrapping the VWLS regression. Markers are colored by the largest source category in the SOE at the facility and include whiskers corresponding to the 95% CI reported by each method.

Pneumatic Controllers

Paired measurements at facilities where simulated emissions from pneumatic devices is the largest category in the SOE are shown in Figure S13. The figure shows 2 facilities where onsite measurement was paired with TFE and at 19 facilities where onsite measurement was paired

with OFE. SOE and OFE are not statistically different at 15 of 19 facilities where the measurements were paired and simulated emissions from pneumatics are the largest source in the SOE. The VWLS regression in Figure S13(b) shows the OFE is higher than the SOE at these 19 facilities, however the 95% CI of the VWLS regression includes a slope of 1 indicating the estimates are not statistically different. Cumulative estimates from the two methods at these 19 facilities are also not statistically different ($cumulative\ SOE = 1.6\ (+0.64/-0.43)$ kg/h, $cumulative\ OFE = 2.9\ (+1.9/-0.75)$ kg/h).

Exact pneumatic actuation counts during the downwind measurement period are not available. Given the infrequent actuation of pneumatic controllers in the study area, pneumatic devices may or may have not have actuated during the downwind measurements. If OTM33A or tracer measured these instantaneous emissions, the downwind estimate could indicate a higher emission rate than the emission rate simulated in the SOE and conversely if these instantaneous emissions were not measured the downwind estimate could indicate a lower emission rate than simulated in the SOE.

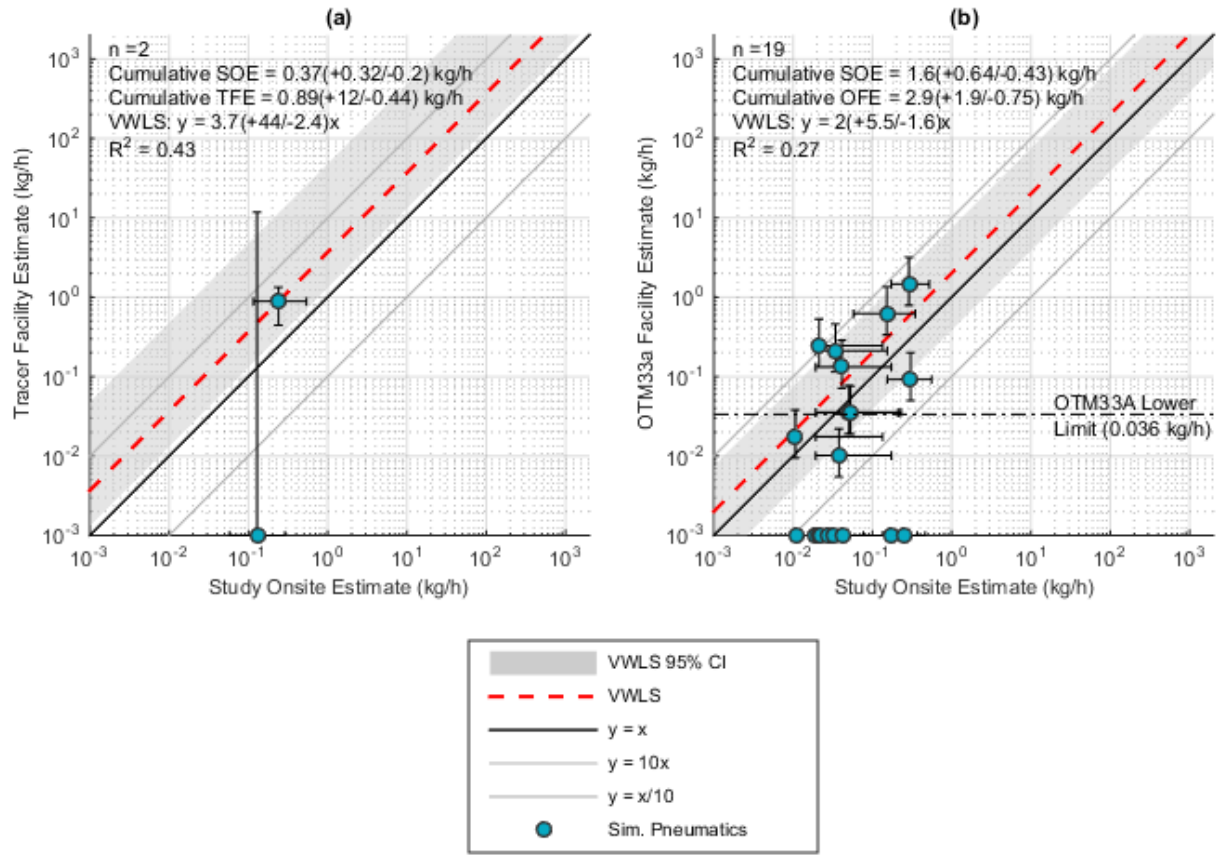

Figure S13: (a) 2 Tracer facility estimates (TFE) and (b) 19 OTM33A facility estimates (OFE) compared to paired study onsite estimates (SOE) at facilities where simulated emissions from normally-functioning, intermittent, pneumatics are the largest source in the SOE. Each panel includes a fit line from a variance weighted least squares regression (VWLS). The shaded region corresponds to 95% CI calculated by bootstrapping the VWLS regression. Markers are colored by the largest source category in the SOE at the facility and include whiskers corresponding to the 95% CI reported by each method. Measurements of zero are shown at 10^{-3} on the log-log axes for display purposes only.

Comparison of TFE to OFE

Contemporaneous measurements using TFE and OFE were made at 9 facilities (Figure S14). The TFE and OFE have overlapping confidence bounds in 6 of these 9 facilities. OFE underestimates relative to TFE as indicated by a variance-weighted least squares linear regression of $TFE = 5.0\ (+3.0/-3.8) * OFE$, and cumulative emissions estimates ($cumulative\ OFE =$

$42\ (+37/-15)$ kg/h, $cumulative\ TFE = 826\ (+207/-207)$ kg/h, Table S8). Although cumulative TFE is 20X cumulative OFE, this difference is driven by low OFE relative to TFE at 2 facilities where simulated liquid unloadings were the largest source in the SOE which for 99% of the cumulative TFE from the 9 paired facilities. At 6 paired facilities where the SOE is dominated by ODMs, cumulative OFE (1.9 kg/h) is 26% of cumulative TFE (7.3 kg/h).

Table S8: Cumulative tracer facility estimate (TFE) relative to cumulative OTM33A facility estimate (OFE) at paired facilities.

			Paired Facilities Subdivided by Largest SOE Category	
		All Facilities with paired TFE & OFE	Simulated unloadings	ODM or simulated observed/unmeasured
Number of facilities		9	2	6
Statistically similar estimates		6	1	4
Cumulative OFE (kg/h)		42 (+37/-15)	40.2	1.9
Cumulative TFE (kg/h)		826 (+207/-207)	819	7.3
Fraction of Cumulative TFE		100%	99%	1%
Cumulative OFE / Cumulative TFE		5%	5%	26%
Correlation coefficient, r		0.98	1.00	0.52
Variance Weighted Least Squares[1,2]	Slope	5.0 (+3.0/-3.8)	1.42 (+1.00/-0.85)	8.13 (+7.34/-4.43)
	R^2	0.29	-0.81	-5.11

1 Linear regression with intercept of 0.
2 Note R^2 is a measure of goodness of fit in y-axis only and does not consider goodness of fit in x-axis. Negative R^2 is a result of regression with no intercept.

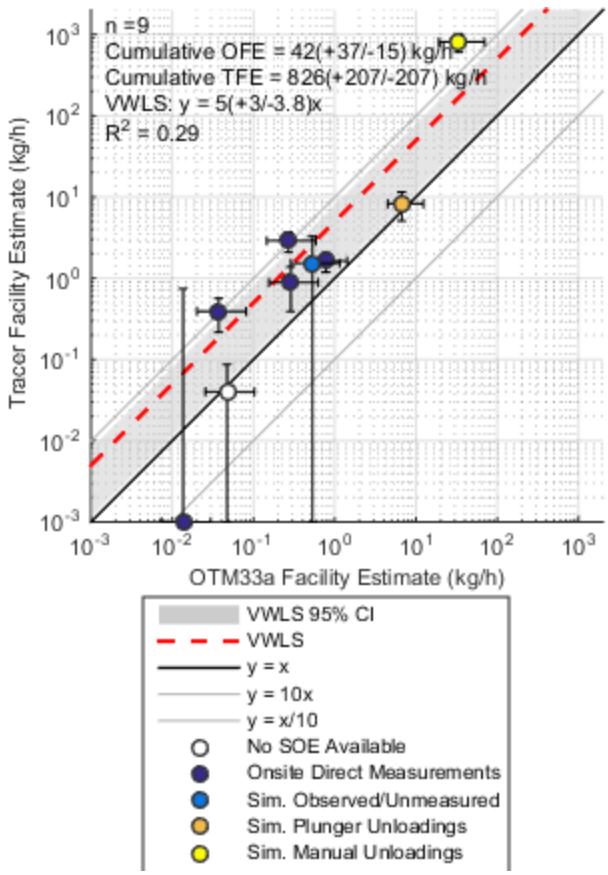

Figure S14: 9 Tracer facility estimates (TFE) compared to paired OTM33A facility estimates (OFE) (n=9). Markers are colored by the largest emission category in the study onsite estimate (SOE) where available. Estimates from the two methods are statistically different at 3 of 9 facilities. Variance-weighted least squares linear regression (VWLS) suggests TFE is statistically different and estimates five times the emissions estimated by OFE ($TFE = 5.0(+3.0/-3.8) \times OFE$). Cumulative estimates produced by the two methods are also statistically different (cumulative OFE = 42 (+37/-15) kg/h, cumulative TFE = 826 (+207/-207) kg/h) however this factor of 20 difference is driven primarily by the facility with a manual unloading.

S8. Orthogonal Least Squares Regression.

Orthogonal least squares (OLS) is an alternate regression method used to compare two estimates of the same parameter made using different methods with associated errors in which the orthogonal residuals between data and a line of best fit are minimized. A coefficient of determination R_y^2 is calculated for each regression to indicate the fraction of variance in y explained by the regression as:

$$R_y^2 = 1 - \frac{\sum(y - \hat{y})^2}{\sum(y - \bar{y})^2}$$

where $\hat{y}$ is the predicted value from the regression $\hat{y} = a + bx$, and $\bar{y}$ is the average of y data.

Similarly, a coefficient of determination R_x^2 is calculated for each regression to indicate the fraction of variance in x explained by the regression as:

$$R_x^2 = 1 - \frac{\sum(x - \hat{x})^2}{\sum(x - \bar{x})^2}$$

where $\hat{x}$ is the predicted value from the regression $\hat{x} = (y - a)/b$, and $\bar{x}$ is the average of x data.

Confidence intervals for the OLS regressions were calculated using a bootstrap method where each data point is resampled and a regression is performed on the resampled data. When resampling the SOE a new value was drawn from a site specific distribution generated by the SOE Monte-Carlo model. This allows the non-Gaussian distribution of emission estimates from the SOE at each facility to be accounted for in the bootstrapped OLS regression. When resampling the TFE or OFE data were sampled from a normal distribution with the 95% confidence bounds reported by the method at each site.

Intermittent Pneumatic Controllers

We use OLS regression to compare facilities where we expect an intercept in the regression, namely facilities where emissions from intermittent pneumatic controllers are the primary source in the SOE. Figure S15 shows paired measurements between (a) TFE and SOE and (b) OFE and SOE at facilities where simulated emissions from intermittent pneumatic controllers is the primary source in the SOE. The OLS regression between OFE and SOE indicates OFE estimates

higher emissions at these sites than SOE with OFE to TFE ratio = 6.4 ($\frac{+22}{-2.9}$) however a negative intercept, $y_{x=0} = -0.39(\frac{+0.24}{-1.9})$, indicates SOE includes some emissions when OFE is zero. At these facilities the SOE model includes emissions from intermittent pneumatic controllers using average emission rates measured by Allen, Pacsi, et al., 2015 and therefore always contains some level of emissions, however the downwind methods may not measure any emissions if no pneumatic actuations occurred while the measurement was being performed. In contrast, if controller actuations did occur during downwind measurements, the downwind methods may estimate higher emission rates than estimated using the time average rates in the SOE.

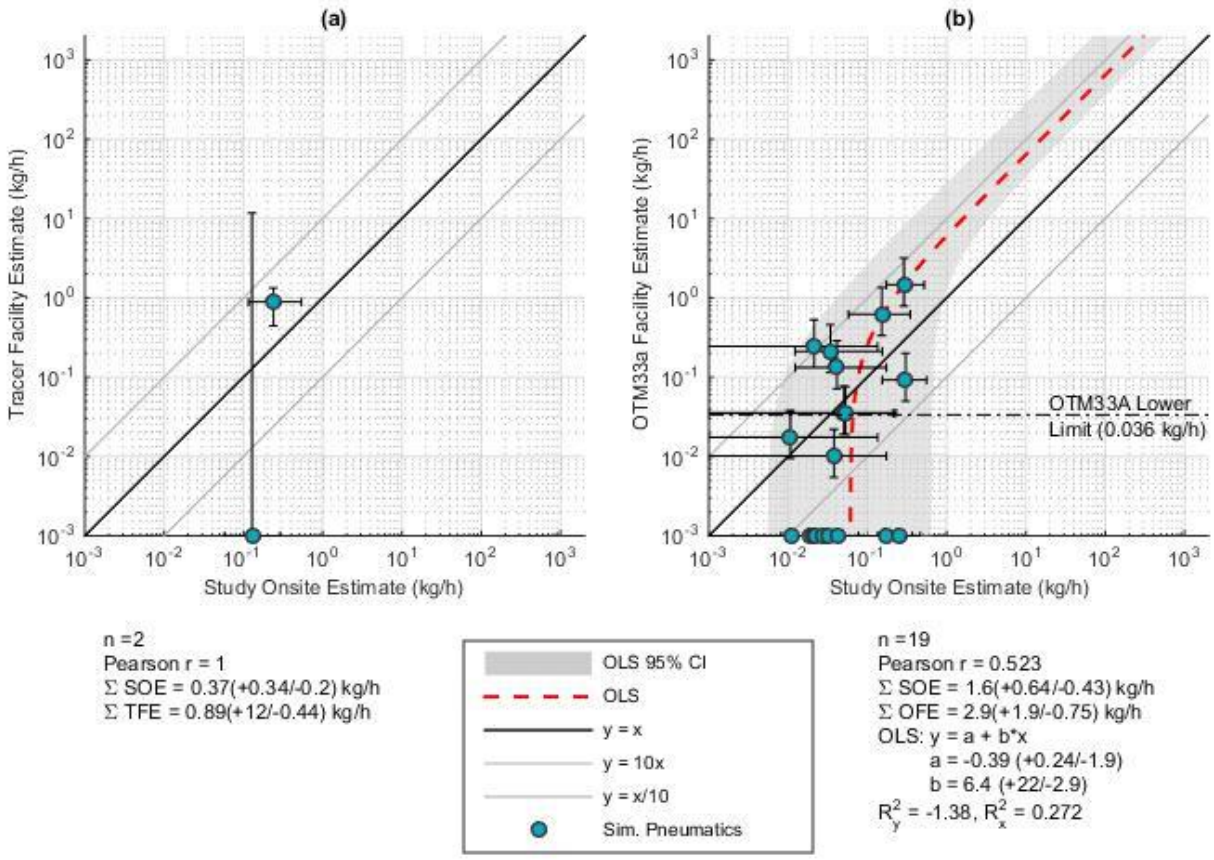

Figure S15: (a) 2 Tracer facility estimates (TFE) and (b) 19 OTM33A facility estimates (OFE) compared to paired study onsite estimates (SOE) at facilities where simulated emissions from normally-functioning, intermittent, pneumatics are the largest source in the SOE. Panel (b) includes a fit line from a orthogonal least squares regression (OLS). The shaded region corresponds to 95% CI calculated by bootstrapping the OLS regression. Panel (a) does not

include OLS regression due to the small sample. Markers are colored by the largest source category in the SOE at the facility and include whiskers corresponding to the 95% CI reported by each method. Measurements of zero are shown at 10^{-3} on the log-log axes for display purposes only.

S9. Sensitivity Analysis

Six facilities at which paired measurements were performed on different days were included in the analysis presented in this paper. These facilities were sites 513, 926, 927, 2072, 2599 and 2779. Table S9 summarizes the paired measurements at these facilities. Repeating the analysis without including these facilities in regressions or cumulative totals does not significantly impact the conclusions of this work.

Table S9: Facilities where paired measurements were performed on different days

Facility ID	Method 1	Method 2	Days apart	Statistically Different Estimates?
513	OTM33A (10/2/2015)	Onsite (10/12/2015)	10	NO
926	Onsite (9/28/2015)	OTM33A(9/29/2015)	1	NO
927	Onsite (9/28/2015)	OTM33A (9/29/2015)	1	NO
2072	OTM33A (10/2/2015)	Onsite (10/12/2015)	10	NO
2599	Onsite (10/5/2015)	OTM33A (10/6/2015)	1	NO
2599	Onsite (10/5/2015)	Tracer (10/6/2015)	1	NO
2779	Onsite (10/5/2015)	OTM33A (10/6/2015)	1	NO

References

Allen DT, Pacsi AP, Sullivan DW, Zavala-Araiza D, Harrison M, Keen K, Fraser MP, Daniel Hill A, Sawyer RF, Seinfeld JH. 2015. Methane Emissions from Process Equipment at Natural Gas Production Sites in the United States: Pneumatic Controllers. *Environ Sci Technol* **49**(1): 633–640. doi: 10.1021/es5040156

Allen DT, Sullivan DW, Zavala-Araiza D, Pacsi AP, Harrison M, Keen K, Fraser MP, Daniel Hill A, Lamb BK, Sawyer RF, et al. 2015. Methane Emissions from Process Equipment at Natural Gas Production Sites in the United States: Liquid Unloadings. *Environ Sci Technol* **49**(1): 641–648. doi: 10.1021/es504016r

Allen DT, Torres VM, Thomas J, Sullivan DW, Harrison M, Hendler A, Herndon SC, Kolb CE, Fraser MP, Hill AD, et al. 2013. Measurements of methane emissions at natural gas production sites in the United States. *Proc Natl Acad Sci* **110**(44): 17768–17773. doi: 10.1073/pnas.1304880110

Bacharach. 2015. HI FLOW® Sampler For Natural Gas Leak Rate Measurement Instruction 0055-9017 Operation and Maintenance Rev. 7 – July 2015. Author. Available at http://heathus.com/products/hi-flow-sampler/. Accessed 2016 Feb 19.

e-CFR: TITLE 40—Protection of Environment. n.d. TITLE 40—Protection of Environment Electronic Code of Federal Regulations § Part 98—Mandatory Greenhouse Gas Reporting Part 98—Mandatory Greenhouse Gas Reporting. Available at http://www.ecfr.gov/cgi-bin/text-idx?tpl=/ecfrbrowse/Title40/40cfr98_main_02.tpl. Accessed 2016 Jul 1.

Howard T. 2015. University of Texas study underestimates national methane emissions at natural gas production sites due to instrument sensor failure. *Energy Sci Eng* **3**(5): 443–455. doi: 10.1002/ese3.81

Neri F, Saitta G, Chiofalo S. 1989. An accurate and straightforward approach to line regression analysis of error-affected experimental data. *J Phys [E]* **22**(4): 215. doi: 10.1088/0022-3735/22/4/002

Robertson AM, Edie R, Snare D, Soltis J, Field RA, Burkhart MD, Bell CS, Zimmerle D, Murphy SM. 2017. Variation in Methane Emission Rates from Well Pads in Four Oil and Gas Basins with Contrasting Production Volumes and Compositions. *Environ Sci Technol* **51**(15): 8832–8840. doi: 10.1021/acs.est.7b00571

Yacovitch TI, Daube C, Vaughn TL, Bell C, Roscioli JR, Knighton B, Nelson DD, Zimmerle D, Pétron G, Herndon SC. n.d. Natural gas facility emission measurements by dual tracer flux in two US natural gas producing basins. *Elem Sci Anth*. In press. http://doi.org/10.1525/elementa.251

Zimmerle D, Pétron G, Pickering C, Vaughn T, Bell C, Schwietzke S, Mielke-Maday I, Heath G, Nummedal D, Howell C. 2016. Reconciling Top-down and Bottom-up Methane Emission Estimates from Onshore Oil and Gas Development in Multiple Basins: Report on Fayetteville Shale Study. Report No.: RPSEA Contract Number: 12122-95.