Oil & Natural Gas Technology

DOE Award No.: DE-FE0014055

FINAL REPORT (10/1/2013 – 3/31/2017)

Measurements and modeling to quantify emissions of methane and VOCs from shale gas operations

Submitted by:
Albert A. Presto

Signature
Carnegie Mellon University
DUNS #: 05-218-4116
5000 Forbes Ave
Pittsburgh, PA 15213
e-mail: apresto@andrew.cmu.edu
Phone number: (412) 721-5203

Prepared for:
United States Department of Energy
National Energy Technology Laboratory

6/30/2017

Office of Fossil Energy

Major goals of the project

The objectives of the project were to determine the leakage rates of methane and ozone-forming Volatile Organic Compounds (VOCs) and the emission rates of air toxics from Marcellus shale gas activities. Methane emissions in the Marcellus Shale region were differentiated between "newer" sources associated with shale gas development and "older" sources associated with coal or conventional natural gas exploration.

Accomplishments of the project

Overview

This project conducted measurements of methane and VOC emissions from both shale and non-shale natural gas resources. The initial scope of the project was the Marcellus Shale basin, and measurements were conducted in both the western wet gas regions (southwest PA and WV) and eastern dry gas region (northeast PA) of the basin. During this project, we obtained additional funding from other agencies to expand the scope of measurements to include additional basins. The data from both the Marcellus and other basins were combined to construct a national analysis of methane emissions from oil & gas production activities.

Results of the project are described in detail below. Briefly, we can summarize the major accomplishments and findings as follows:

1. **Methane emissions measurements and attribution:** Emission rates of methane were measured for both shale and non-shale gas sources in the Marcellus Shale. These measurements contributed substantially to the existing body of literature quantifying methane emissions from shale gas resources. Our measurements of methane emissions from non-shale gas wells in PA and WV seem to be the largest data set contributed in this space since the Gas Technology Institute studies in the 1990s. In PA and WV, total methane production is dominated by newer, high producing shale wells, but emissions are almost equally split between new wells and older, less productive conventional wells.
2. **VOC emissions:** We measured emission rates of non-methane VOCs from shale and non-shale sources in the Marcellus Shale and other basins. Overall, we observed small emission rates that are not well correlated with methane emissions, indicating that the sources of VOC emissions are often distinct from the sources of methane emissions (e.g., tank venting versus process leaks).
3. **Method development:** We contributed to the methods used to measure emissions from oil & gas facilities. Most measurements relied on the previously developed tracer flux method, which is often viewed as the gold standard. We evaluated the method limits of the tracer flux method, with specific focus on errors associated with placement of tracer release points relative to methane emissions points. We also employed two newer methods, EPA OTM 33a and a Gaussian plume based inverse modeling method, and compared emissions measurements from all three methods.
4. **National emissions:** Using a combination of data collected in this and related projects along with previously published data, we have developed a national emission estimate for the U.S. The national picture quantifies total emissions as well as basin-to-basin differences in methane leakage rate (i.e., percent of production lost) due to different mixes in facility age and production volumes. Our analysis also informs the fraction of emissions that can be avoided by targeting specific high-emitting facilities, sometimes called super emitters. An ongoing study, funded by NASA, will enable us to compare our

methane inventory to existing inventories, and to implement the new inventory into global chemical transport models.

The research conducted in this project will contribute to a minimum of five publications. All analyses for the publications listed below are complete, and we expect to submit all the manuscripts by the end of the 2017 calendar year.

- Omara, M.; Sullivan, M.R.; Li, X.; Subramanian, R.; Robinson, A.L.; Presto, A.A. Methane emissions from conventional and unconventional natural gas production sites in the Marcellus Shale basin. *Environ. Sci. Technol.*, **2016**, DOI: 10.1021/acs.est.5b05503
- M. Omara, N. Zimmerman, X. Li, A.E. Ellis, M. Sullivan, R. Subramanian, A.L. Robinson, A.A. Presto Estimates of U.S. onshore methane emissions from active oil and gas production sites based on site-level measurements. *In preparation for Nature Communications*. **Submission expected by August 2017.**
- N. Zimmerman, M. Omara, X. Li, A.E. Ellis, M. Sullivan, R. Subramanian, A.A. Presto, A.L. Robinson. Modified Gaussian plume models as an effective surveying tool for quantifying fugitive methane from natural gas production. *In preparation for Atmos. Meas. Tech.* **Submission expected by November 2017**
- Li, X.; Zimmerman, N.O.; Omara, M.; Sullivan, M.; Subramanian, R.; Robinson, A.L.; Presto, A.A. VOC emissions from natural gas production in multiple basins. *In preparation.* **Submission expected by November 2017.**
- Li, X.; Presto, A.A.; Adams, P.J. Comparison of methane emissions inventories. *In preparation.* **Submission expected by December 2017.**

Significant Results

1.0 Methods

1.1 Tracer Flux Measurements

1.1.1 Overview

The tracer flux method uses a mobile platform to intercept plumes of target emissions (methane in the examples below) and intentionally emitted tracers. Figure 1 shows sample plumes of a tracer (N_2O) and methane detected downwind of a shale gas well. The temporal correlation of the plumes indicates a common source.

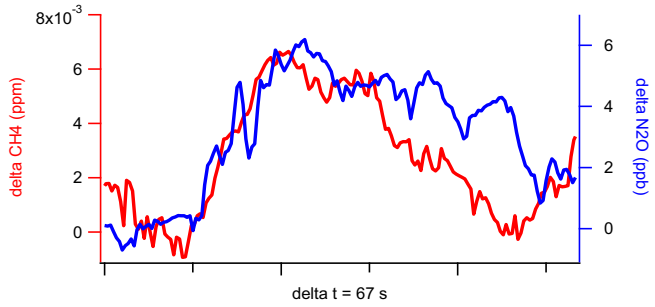

Figure 1: Example of a methane (red) and N_2O (blue) plume detected down wind of a shale gas well. Concentrations are background corrected.

The methane flux is calculated from each plume intercept using the following equation:

$$Q_{CH4} = Q_{tracer}\left[\frac{C_{CH4} - C_{CH4,bkd}}{C_{tracer} - C_{tracer,bkd}}\right]$$

Where Q_{CH4} and Q_{tracer} are the fluxes (flow rates) of methane and the tracer, C_{CH4} and C_{tracer} are the concentrations of methane and the tracer measured in the plume, and $C_{CH4,bkg}$ and $C_{tracer,bkg}$ are the background concentrations.

Concentrations of methane and the tracer are determined by analyzing data for individual plumes following a predetermined decision tree. Optimal plumes have concentrations of CH_4, C_2H_2, and N_2O that are highly correlated, as in Figure 2. Under such conditions, concentrations of the tracers and methane at each time point in the plume can be correlated, as shown in the upper panels of Figure 2. Thus, Q_{CH4} can be calculated using both tracers, and the measured C_2H_2:N_2O ratio can be compared to the known ratio of the flow rates of these species as an internal calibration check. Measured C_2H_2:N_2O ratios that are consistent with the tracer release rates are an indication of equidispersion of methane and the tracers.

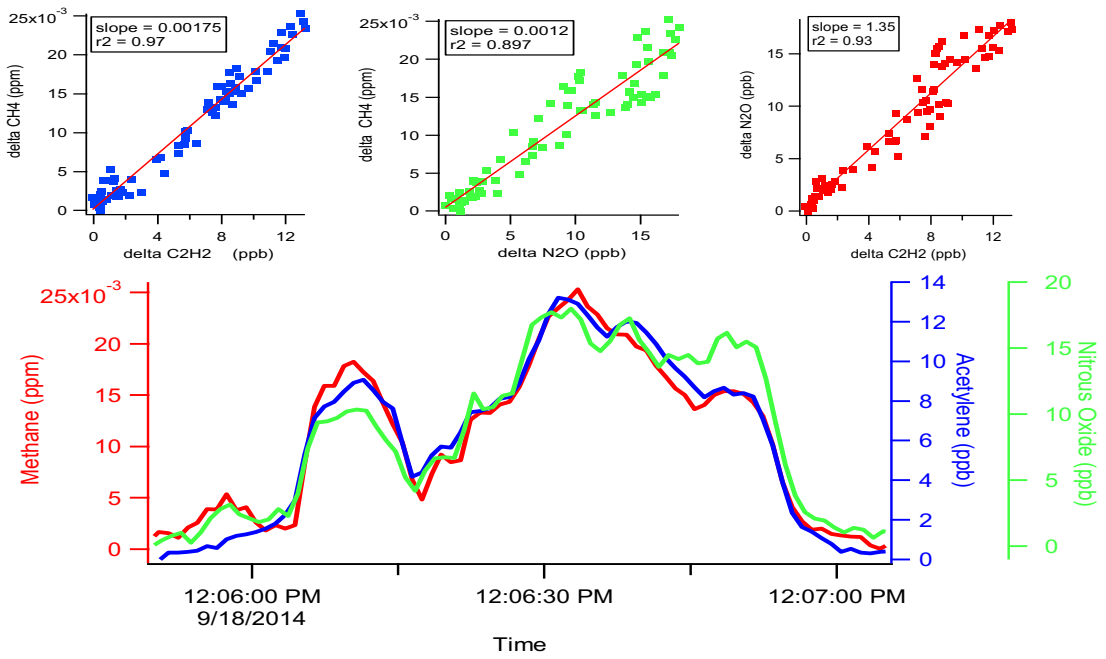

Figure 2: Example of well-correlated plumes intercepted 168 m downwind of the tracer release point. The upper panels show regressions of methane versus acetylene (blue), methane versus N_2O (green), and N_2O versus acetylene (red). The bottom panel shows the background-corrected plumes of each species.

In some cases the point-by-point analysis method shown in Figure 2 cannot be used. This often occurs when there is a slight offset in the methane and tracer plumes, or when the plumes have small peak magnitudes. In these cases, the total area under each plume is calculated, as shown in Figure 3. Methane fluxes determined by this "area" method are less certain than by the method shown in Figure 2. This uncertainty is reflected in the data with larger error bars.

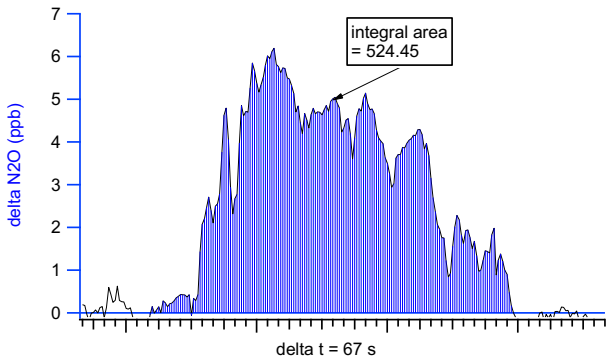

Figure 3: Example of plume integration for an N_2O plume.

Multiple plumes are detected for each location. The average emission for each site is determined by computing the WAFLER (Weighted Average Facility Level Emission Rate). The emission rate calculated for each plume is weighted by its specific uncertainty as determined from the recovered N_2O/acetylene ratio.

One challenge of the tracer flux method is to separate intentional plumes (e.g., the released tracers and methane from the sampling site) from any other potential plumes. This is accomplished in two ways. First, we release two tracers, and identify concurrent plumes of both tracers with the correct mass ratio of the tracers as part of QA. Second, we verify the location of the plumes with measured wind speed and direction data, as shown in Figure 4.

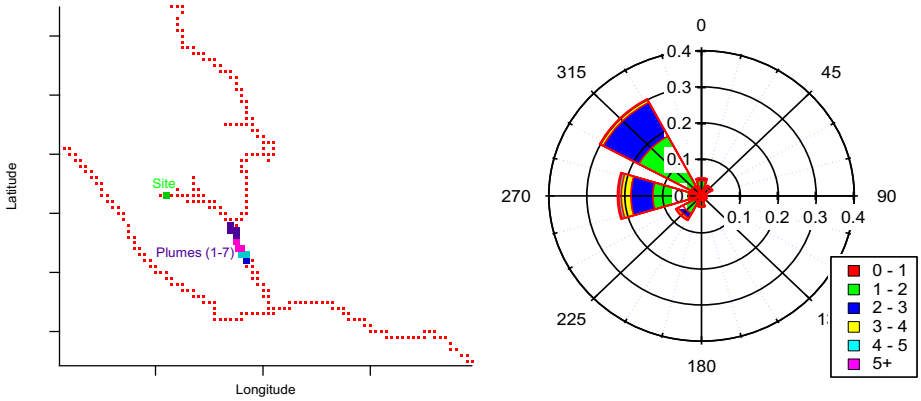

Figure 4: Plumes identified during sampling correspond to locations downwind of the sampling site, as indicated by the wind rose.

1.1.2 Calibration experiments

In addition to making measurements at natural gas sites, we conducted several calibration experiments to verify the accuracy of the tracer release method. In these experiments, methane (99.9% purity) and tracers (Acetylene and Nitrous Oxide) were released at known flow rates in an urban park on three different days. The release rate of methane was varied from 15 SLPM to 45 SLPM

In two of the tests, the two tracers were placed at close proximity to the release point for methane (approximately 10 to 15 m apart). In the third test, only acetylene was placed close to the methane release point, while nitrous oxide was released at different distances, spanning 10 m to 75 m from the methane release point. In all three tests, the downwind distances at which plumes were intercepted varied from approximately 100 m to 460 m.

Overall, calculated methane release rates from the dual tracer technique gave results that closely matched actual release rates (Table 1, Figure 5). For all plumes, the uncertainty associated with the technique was ± 4.4 SLPM. Uncertainty was defined from the relative error in tracer recovery, which was calculated as the difference between the nitrous oxide/acetylene slope of the plumes intercepted downwind and the onsite ratio, normalized by the onsite ratio. Each plume emission rate estimate was used to calculate the weighted average methane emission rate, weighted by its uncertainty ($1/\sigma^2$).

Table 1: Summary of results of calculated methane emission rates from the tracer release calibration tests.

Actual CH4 Emission Rate (SLPM)	Weighted Average CH4 Emission Rate (SLPM)	Uncertainty	Minimum	Maximum
15	13.2	2.1	11.2	22.1
30	33.8	2.2	25.3	37.6
45	45.3	5.4	41.9	51.1

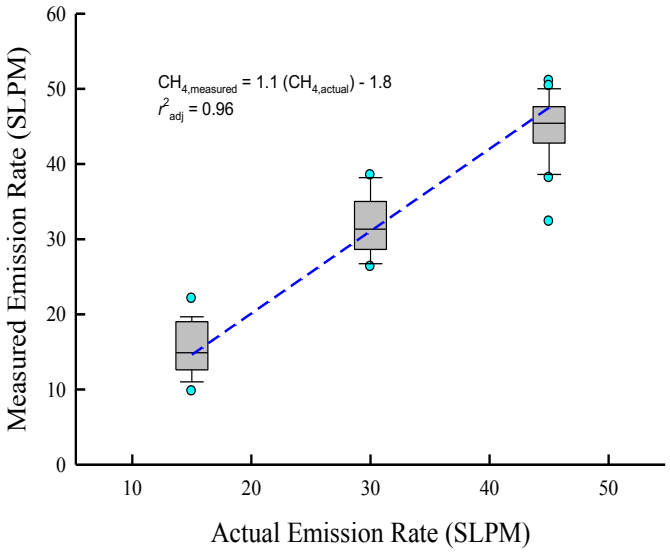

Figure 5: Plot of measured methane emission rate versus actual methane release rate (SLPM). The boxes show the 25th and 75th percentiles, the line inside each box shows the median, whiskers show the 10th and 90th percentiles, and the circles show the extremes.

High correlation (i.e., $r^2 > 0.75$, obtained from regressions of background corrected methane against nitrous oxide) were obtained for all tracer distances ranging from 10 m to approximately 50 m from the methane release point. At 75 m, wind speed of approximately 5 MPH, and downwind distance of approximately 230 m, the nitrous oxide plume was separated in space from the methane and acetylene, and was generally smaller in area and concentration, suggesting

unequal dispersion for the two tracers (Figure 6).

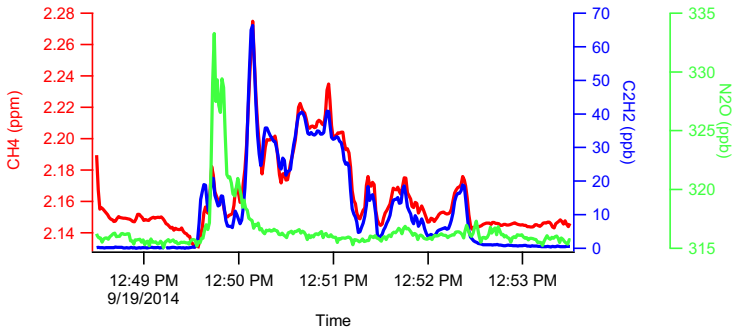

Figure 6: Raw data from a plume intercepted 230 m from methane source. Here, nitrous oxide was released approximately 75 m from the methane source. Note the spatial separation and the difference in lengths of the nitrous oxide plume and that of methane and acetylene. Here, the downwind ratio of the background corrected tracer areas was approximately 9-fold greater than that onsite. Therefore, only the acetylene plume was accepted.

From these data, the placement of the tracers relative to the methane source is critical in assessing the accuracy of the method. Collocated sources are indicated by good correlation between methane and the tracers (here defined as regressions with $r^2 > 0.75$). As a quality criterion, when plumes from one of the tracers and methane are spatially separated (as in Figure 6), the data is still useful if a check of the ratio of the areas under the plumes for the tracers is within a factor of 1.5 of the release ratio onsite.

1.2 Other methods: OTM 33A and Gaussian inverse modeling of drive-by plumes

Tracer flux measurements require access on or near emissions sites to place the tracers, and tracer placement can be time consuming. We investigated two other sampling methods that rely on plume capture but do not require tracer release or direct site access.

Both the OTM33A and drive-by methods were evaluated against the standard dual tracer flux approach. The tracer species, methane, and other analytes were measured approximately 100 m to 2 km downwind of the target facility At least five plumes were collected at each facility, and the facility level emission rate (FLER) for methane can be calculated from the measured tracer release rate and the background-corrected methane-to-tracer ratios derived from each plume transect.

For the drive-by analysis, downwind drive-by measurements were performed using the mobile laboratory equipped with real-time instruments to measure methane/acetylene (Model G2203, Picarro Inc., Santa Clara, CA), and nitrous oxide and ethane (quantum cascade tunable infrared laser differential absorption spectrometer (QC-TILDAS), Aerodyne Research, Inc., Billerica, MA.) For each selected facility, the mobile lab drives past the facility on a downwind road located 50-200 m from the facility. The drive-by approach utilizes a standard Gaussian dispersion model to back-calculate facility-level emissions.

$$C = \frac{Q}{u\sigma_z\sigma_y 2\pi} e^{-y^2/2\sigma_y^2} \left(e^{-(z_r - H_e)^2/2\sigma_z^2} + e^{-(z_r + H_e)^2/2\sigma_z^2} \right)$$

In the Gaussian dispersion equation, y is the distance away from the plume center point, σ_x, σ_y, and σ_z are dispersion coefficients determined from wind speed (u), current meteorology and downwind distance, x, from facility. Following the work of Yacovitch et al. (2015), measurement and release heights (z_r and H_e) are set to zero. The method works by first assuming the facility-level emission rate (Q) is equal to 1 g/s to calculate the downwind methane concentration (C). The slope of the linear regression fit between the modeled and measured downwind methane concentration provides the actual facility level emission rates.

Prior to data analysis, the combined planned and accidental drive-by sites totaled 94 between the Denver-Julesburg and Uintah campaigns (funded by a second project from NOAA, NA14OAR4310135), of which 28 sites also had tracer flux measurements. Sites with both tracer flux and drive-by measurements were used to develop quality assurance and quality control (QA/QC) criteria for sites that were only subjected to the drive-by measurement approach. From this assessment, it was concluded that plumes for which wind speed exceeded 2.5 m/s and during which the standard deviation of the wind during the plume was less than 25% generally agreed with the tracer flux estimate within the margin of error first proposed in Yacovitch et al. (2015) (up to a factor of three discrepancy). Upon applying these QA/QC criteria and comparing the measured and modelled emission rates of methane and the two tracer gases, agreement within the range suggested by Yacovitch et al. (2015) was generally observed (Figure 7).

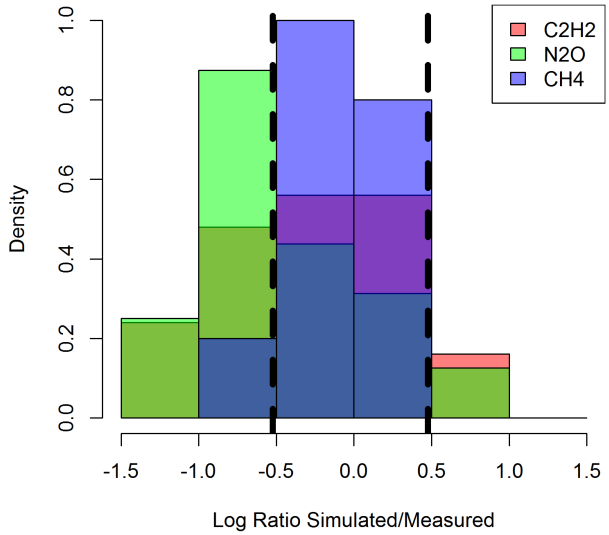

Figure 7: Histogram of observed ratios of simulated/modelled methane emission rates using the drive-by assessment vs. those measured using tracer flux or the known tracer release rates. The black dashed lines are the suggested range of agreement from Yacovitch et al. (2015).

As wind-related parameters had the greatest impact on the modelled emissions estimate, for the drive-by analysis 5 wind speed scenarios were considered. These scenarios included: (1) the average wind speed scenario (Avg. WS), in which the average wind speed during the plume us used; (2 and 3): average wind speed +/- the standard deviation of the wind during the plume (Avg. WS +/- SD); and (4 and 5): maximum and minimum wind speeds observed during the plume and during the estimated transport time (Max./Min. WS).

It was concluded that candidate plumes for drive-by analysis need an average wind speed greater than 2.5 m/s and a standard deviation of the wind during the plume of less than 25%. Upon applying these metrics, agreement between tracer flux and drive-by was marginally improved; however, the drive-by analysis tended to underestimate concentrations (Figure 8).

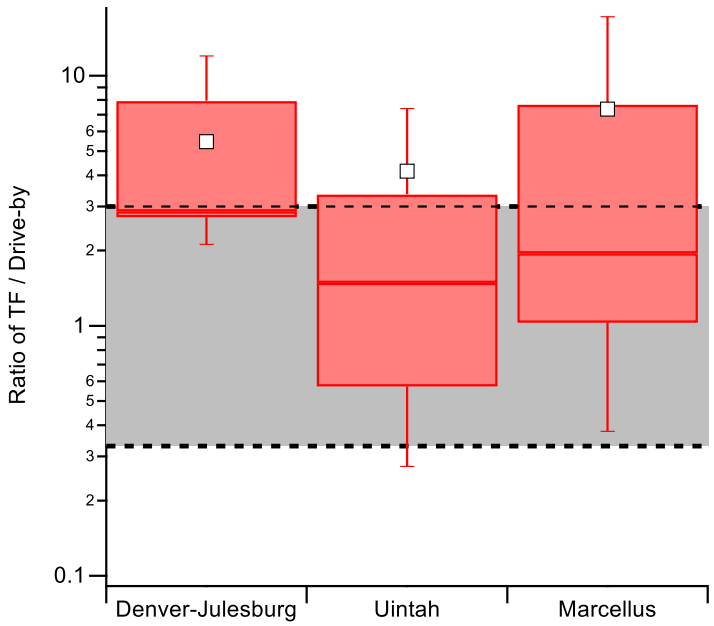

Figure 8. Agreement between tracer flux and drive-by across the campaigns. The grey area is the suggested range of agreement by Yacovitch et al. (2015).

For the OTM33A analysis, the mobile laboratory is positioned at a stationary location 100-200m downwind of the target facility for a 15-30-minute sampling period. During this sampling period, measured methane data were binned in 10-degree increments by wind direction, and fitted to a Gaussian function.

Both compressor stations (n = 7) and production well pads (n = 8) were sampled using OTM 33A. Comparing the OTM33A and tracer flux measurements, OTM 33A measurements were mostly within a factor of 2 of the tracer flux measurements (Figure 9).

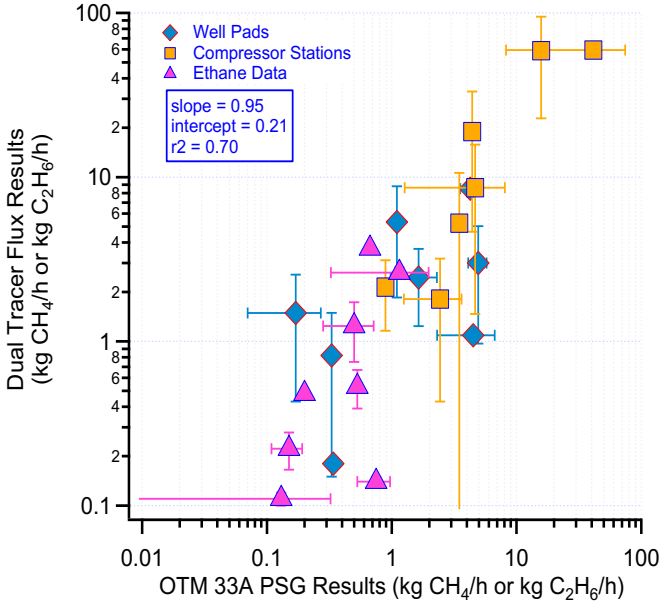

Figure 9: Comparison of methane emissions estimates using dual tracer flux and the proposed EPA OTM33A protocol. Error bars represent one standard deviation.

In general, the OTM 33A approach yielded lower emissions estimates compared to tracer flux, likely because of (i) the assumption of point source in estimates of emissions using the point source Gaussian approach, and (ii) the fact that these two measurements were not performed simultaneously. The data obtained were not comprehensive (n = 2 for most sites). Also, it is important to note that the tracer flux and the OTM 33A measurements were not obtained simultaneously, that is, OTM 33A measurements were performed after a 2 to 3-hour tracer flux measurement.

2.0 Measurement results

2.1 Methane emissions: Southwestern Marcellus (PA and WV)

Methane emissions were measured at 31 sites in the western portion of the Marcellus Shale Basin. The sites are a combination of producing shale gas wells, producing conventional wells, and shale gas wells during flowback, and compressor stations. There were site-to-site differences in the equipment installed at each site. For example, some sites included an active vapor recovery system to capture emissions from condensate tanks, while others did not.

There were also operational differences between sites. For example at one site, we collected data during "normal" operations as well as during a condensate unloading event. During condensate unloading natural gas liquids were emptied from a storage tank into a tanker truck.

Figure 10 shows measured methane emissions at the sampled sites. Emissions are higher at flow back sites than producing wells (either shale or conventional). Emissions are also slightly higher at shale wells than conventional wells, however the raw emission rate does not account for differences in production between the shale and conventional sites. In general, the shale gas wells are larger (multi-well pads) than conventional wells, and also have large gas production. When normalized against production, shale gas wells (both flow back and producing) emit < ~1% of total production.

Emissions from gas wells are expected to be a "fat tail" problem, with a small number of large emitters dominating the total emissions. While we are yet to identify a specific super emitter, we are beginning to see the distribution of emissions across sites, with higher emissions at a few wells and many sites with low emissions. The largest emissions were observed at a compressor station during a blow down event (Site 29). Emissions are discussed in more detail in previous reports and by Omara et al (2016).

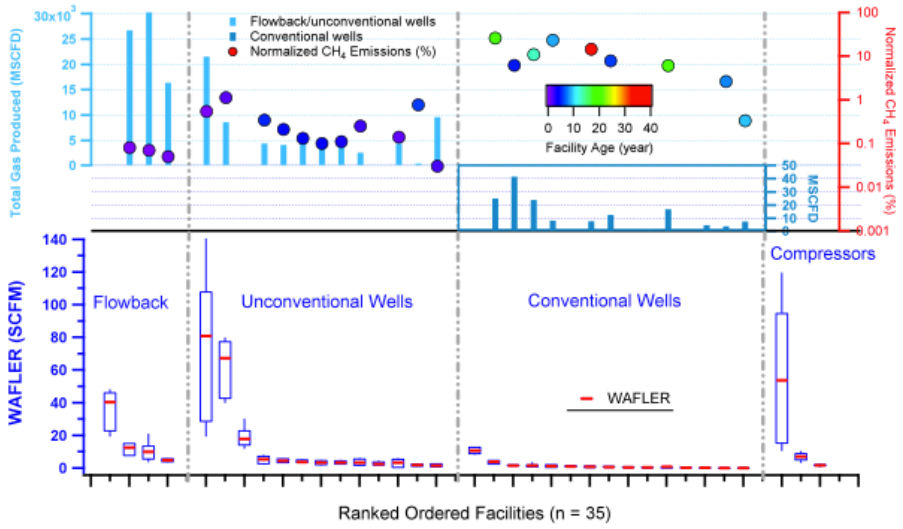

Figure 10. WAFLER for each of the sites sampled in southwest PA and WV (bottom left axis), total gas production at well sites (top left axis), and the normalized emissions as a fraction of gas production at well sites (top right axis). Normalized emission rates are colored by the facility age.

2.2 Methane emissions: Northeastern PA

We conducted tracer flux, drive-by, and OTM 33a measurements at sites in northeastern Pennsylvania (NEPA) in April-May 2016. This campaign represented an opportunity to collect additional Marcellus Shale emissions data and to determine if emissions in the northeastern part of the play, which has drier gas, are different than in the southwestern part of the basin. Additionally, gas wells in NEPA are dominated by newer shale wells. Older conventional wells are much less common in this area. We sampled 47 unique production sites in Susquehanna, Bradford, Wyoming, and Sullivan counties (Figure 11).

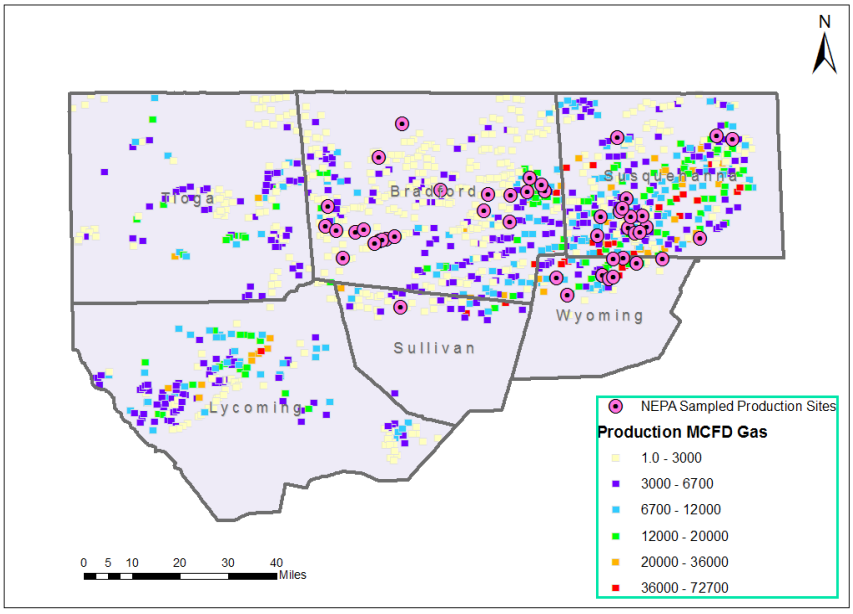

Figure 11. Map of shale gas wells and sampled locations in NEPA.

We took several steps to ensure site representativeness in the NEPA measurements. Specifically, we compared distributions of site age, site-level gas production, and site operator between our sampled sites and the full distribution of production sites in NEPA. An example of the comparison for site age is shown in Figure 12. Overall, the sampled sites matched reasonably well with the full distribution of sites for all three metrics. E.g., when comparing site age distributions, we had a similar mean in sampled sites as for all sites (approximately 5 years), though we were slightly under sampled in sites <2 years old.

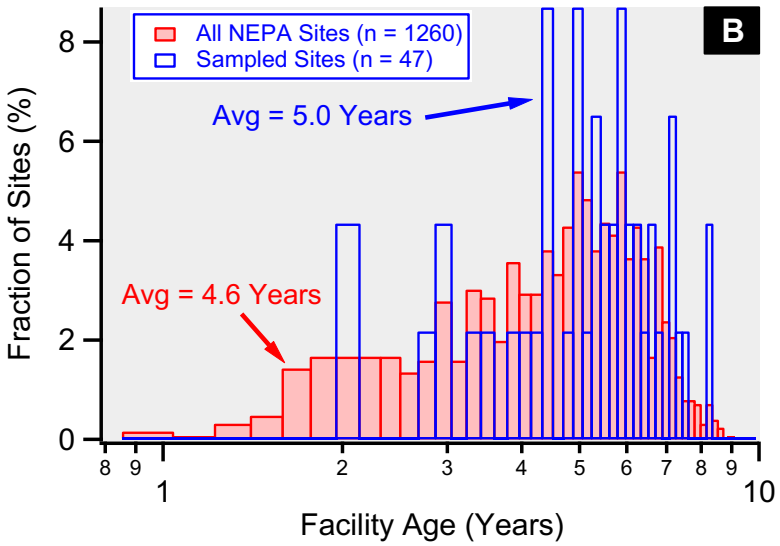

Figure 12. Comparison of the histogram of site ages for NEPA for all sites (red) and sampled sites (blue).

Methane emissions from production sites in NEPA follow what is now known to be a common pattern – absolute emissions span a range of several orders of magnitude, and display a skewed, "heavy tail" distribution. The results are shown in Figure 13. Emissions measurements at individual sites were also skewed. The four largest emission rates were measured as episodic plumes at three sites. These episodic emissions were substantially higher than emissions measured in other plumes measured from the same sites, and strongly skewed calculations of per-site average emissions.

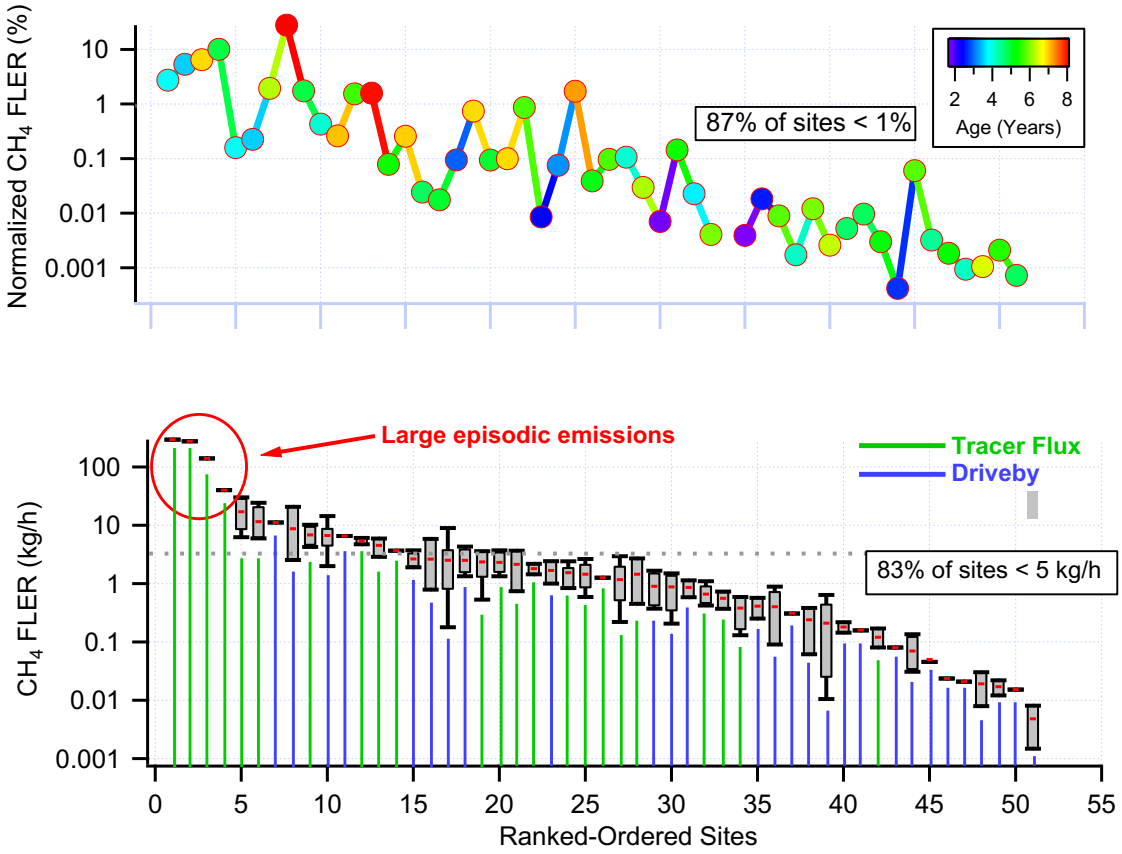

Figure 13. Methane emissions from production facilities in NEPA. Absolute facility-level emissions (bottom panel) span a range of several orders of magnitude. Normalized emissions (emissions divided by production, top panel) are typically small.

The normalized emission rate was typically small. 87% of sites had normalized emissions less than 1% of production, similar to the emission rates observed at shale gas wells in the southeastern portion of the Marcellus Shale in PA and WV. Oil and gas operations in NEPA are dominated by newer shale gas wells. Thus, unlike the southwestern portion of the Marcellus shale, there is not a significant contribution from emissions from older, conventional wells.

2.3 Analysis of VOC data

VOCs were measured inside of identified emission plumes through a combination of canister samples and a real-time GC-FID. A SRI 8610C Gas Chromatograph (GC) was equipped

on the mobile lab to measure in total 14 VOCs species from C3 to C13, including Benzene, Toluene, Ethylbenzene and Xylenes (BTEX), which are all EPA Hazardous air pollutants. We collected in total 47 BTEX samples in the Denver-Julesburg Basin, the Uintah Basin, and the Northeastern PA (NEPA). The site and sample count in each region is summarized in Table 2.

Table 2. Site and sample count of the VOC measurements

Basin	Denver-Julesburg	Uintah	Northeastern PA	Total count
Site count	16	17	9	42
Sample count	20	18	9	47

Our measured VOC concentrations from all three regions are shown in Figure 14. Except ethane, concentrations of other VOC species are generally less than 1 ppb. VOC concentrations of the upwind background samples are also shown in Figure 7. For many VOC species the concentration measured downwind of facilities are not substantially higher than the upwind background concentrations. The VOC concentrations measured in NEPA are lower and less variable than the VOC concentrations measured in the Denver-Julesburg and Uintah Basin. We compared our measurements with the averaged VOC concentrations of 28 cities reported by Baker et al. (2008) (https://doi.org/10.1016/j.atmosenv.2007.09.007) and the ethane, ethylbenzene and xylene concentrations downwind of natural gas facilities are higher than the averaged city concentrations, but other VOC species have either similar or lower concentrations compared with city background. We identified two high VOC emitters in the Denver-Julesburg basin, which emitted about 2-3 orders of magnitude higher than other sites. No high VOC emitters are detected in Uintah Basin and NEPA.

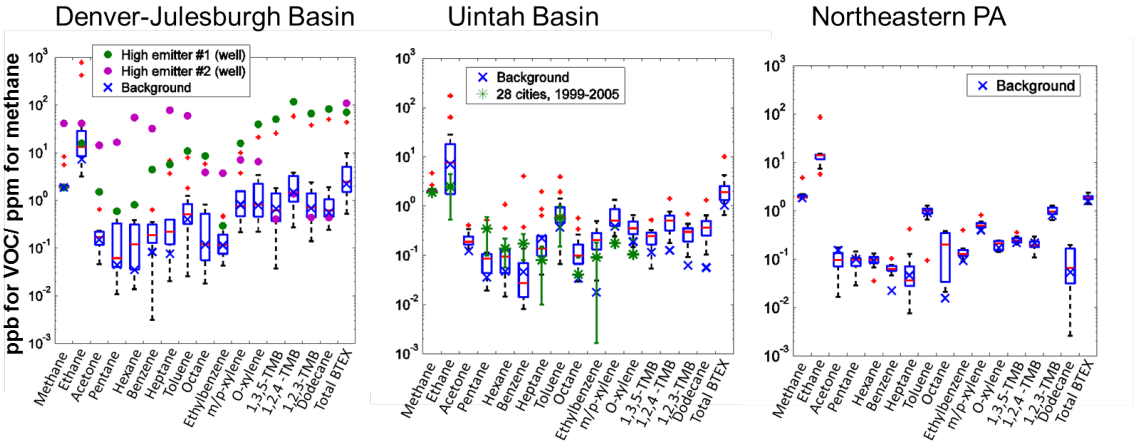

Figure 14. All VOC concentrations (in ppb) and methane concentrations (in ppm) measured in the Denver-Julesburg Basin, the Uintah Basin and the Northeastern PA. We use box-and-whisker plots to represent all measurements for each species. The red line in the middle of the box represents the median concentration; while the top and the bottom of the box represent 75th and 25th percentile, respectively. The background measurements of each basin are shown as blue crosses.

In order to compare the VOC emissions in different regions we calculated averaged background-corrected VOC-to-methane ratios of natural gas production facilities in each region. The result is shown in Figure 15. It indicates that well pads in the Denver-Julesburg Basin tend

to have higher and much variable VOC-to-methane ratio than sites in Uintah basin and NEPA. But since the amount of sites we studied is small, the result may not be representative for all sites in both basins.

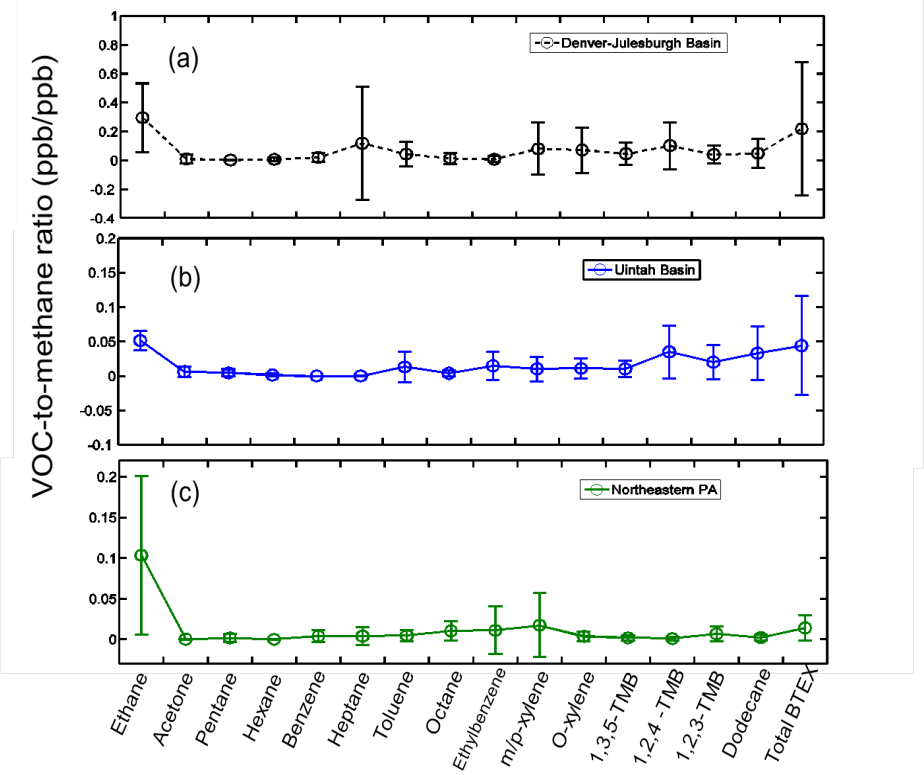

Figure 15. VOC-to-methane ratio in the (a) Denver-Julesburg Basin, (b) Uintah Basin, and (c) Northeastern PA. All VOCs and methane concentrations are background corrected. The circles in the figure represent the mean value, and the bars represent the standard deviation.

For those sites that we have both methane emission rate estimate and the VOC measurement, we estimated their VOC emission rates by scaling the methane emission rates with the background-corrected-VOC-to-methane-ratio measured at the same time, and results are presented in Figure 16. The ethane emission rates in NEPA are slightly lower and less variable compared with the Denver-Julesburg Basin and the Uintah Basin, but for all other VOC species, the emission rates in NEPA are much lower. We compared our estimated VOC emission rates in Uintah Basin with Pétron et al. (http://dx.doi.org/10.1002/2013JD021272) and our estimated median pentane emission rate is about 3 times as Pétron et al., and our estimated benzene emission rates are about 1-2 orders of magnitude higher than Pétron et al.

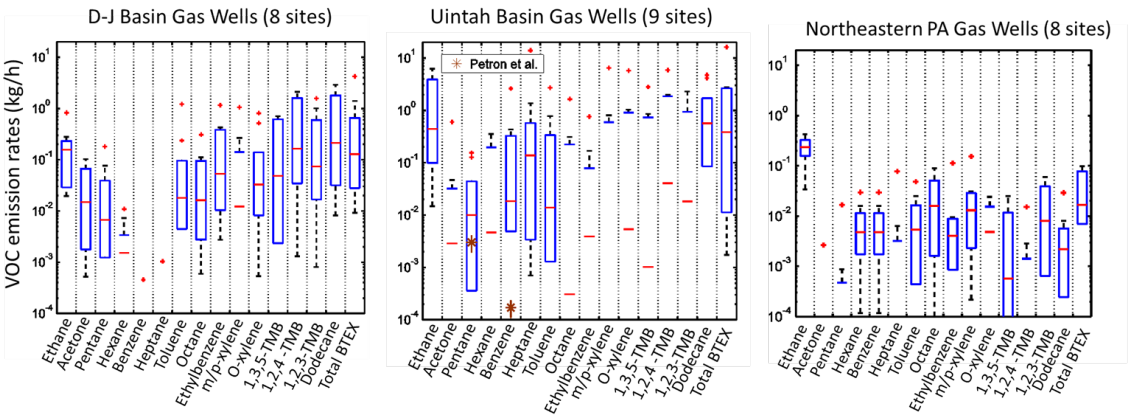

Figure 16. Estimated emission rates of all measured VOC species from the Denver-Julesburg Basin, the Uintah Basin and the Northeastern PA. The estimation was done by scaling the methane emission rates with the background corrected VOC-to-methane ratio. Methane emission rates of these sites were estimated using the tracer-release method. Results from Petron et al. (2014) are also shown in the figure to compare with our results.

3.0 Nationwide emissions estimate

3.1 Methods

3.1.1 Activity data and site characteristics**.** The 2015 monthly production data for approximately 900,000 wells were obtained through Drillinginfo's DI Desktop application. This initial database of wells were subsequently reduced to only include active onshore O&G wells with known unique API numbers and location within the contiguous U.S. states. Thus, approximately 700,000 active 2015 O&G producing wells, excluding offshore wells and other production types (e.g., storage wells, coal bed methane wells, CO_2 wells, water wells, etc.) were identified for further data processing.

The final database of wells was further processed using geospatial analysis tools in ArcGIS to identify well pad sites by aggregating production information for multiple wells originating from specific well pad sites. This procedure was previously described in detail elsewhere (Zavala-Araiza et al. (2015)). Briefly, a 50 m buffer was generated around each known well location and overlapping buffers, indicative of specific multi-well pad sites, were then merged. A spatial join of each well within the merged buffer polygon was then performed, consolidating multi-well parameters (e.g., oil, gas, and water production) to obtain site-level parameters. Well pad sites that had zero gas production in 2015 were flagged and removed from the database at this stage. In total, approximately 511,000 active O&G well pad sites (600,000 wells) were identified, with total 2015 natural gas production of approximately 27 Tcf (trillion cubic feet). Using the EPA's basin-specific estimates of CH_4 content in natural gas (EPA (2016)), we estimate total CH_4 production of approximately 23 Tcf from these wells in 2015. Well pad sites were then sorted into their respective gas production cohort, defined as: (i) <10 Mcfd, (ii) 10 to 100 Mcfd, (iii) 100 to 1000 Mcfd, and (iv) >1000 Mcfd production sites. Overall, these sites accounted for (i) 45%, (ii) 39%, (iii) 13%, and (iv) 3%, respectively, of the total number of identified well pad sites. Alternatively, gas production from these sites accounted for (i) 1%, (ii) 10%, (iii) 25%, and (iv) 64%, respectively, of the total gas production. We also examined the distribution of site age based on the reported first production date. We found that sites producing (i) <10 Mcfd, (ii) 10 to 100 Mcfd, (iii) 100 to 1000 Mcfd, and (iv) >1000 Mcfd natural gas were, on average, (i) 23, (ii) 20, (iii) 11, and (iv) 5 years old, respectively, as of

12/31/2015. Thus, two-thirds of total 2015 gas production were from 3% of total well pad sites that, on average, started production in the previous five years and produced >1000 Mcfd in 2015.

3.1.2 Site-level emissions data and sources. We consolidated a total of 870 site-level CH_4 emissions rate data from eight independent studies in six O&G production basins (Yacovitch et al. (2014); Lan et al. (2015); Goetz et al. (2015); Brantley et al. (2014); Rella et al. (2015); ERG (2011); this study) (Table 3). Site-level CH_4 emissions measurements occurred between 2010 and 2016 and involved measurement approaches that can be broadly grouped into four categories (Table 1). These categories include:

a) direct onsite measurements aided by an optical gas imaging equipment to identify leaks before quantification (ERG (2011));
b) tracer flux mobile measurements using atmospheric tracers (e.g., acetylene and nitrous oxide) as proxy for CH_4 dispersion followed by downwind mobile measurements (Goetz et al. (2015); this study);
c) EPA's Other Test Method (OTM)-33A involving downwind stationary plume measurements and Gaussian plume inverse modeling (Brantley et al. (2014); this study);
d) Downwind mobile measurements of methane plumes utilizing Gaussian plume inverse modeling for emissions rate quantification (hereafter, MM-Gaussian, e.g., Yacovitch et al. (2015); Lan et al. (2015); Rella et al. (2015); this study).

A detailed description of these methods can be found in the references cited for measurement approach (a) to (d) and in the Supplemental Information. It is important to note that the different measurement techniques have inherent limitations that make it difficult to obtain a truly representative sample of sites within each basin or region. For example, a successful direct onsite measurement scheme depends on the proper operation and performance of both the plume imaging (for leaks survey) and flow rate measurement devices (Ravikumar et al. (2016)). Additionally, downwind mobile measurements are limited by the ability to safely access roads downwind of target sites and favorable meteorological conditions at the time of sampling, e.g., good wind speeds that generate sufficient plume transport to mobile receptor platforms.

Table 3. List of studies, production region, and measurement results for the eight studies in six O&G production regions.

Ref. code	Study	Production region	Measurement period	n	Type of production facility	Technique	Absolute CH_4 (kg/h)	Normalized CH_4 (%)	Gas production (Mcfd)
(a)	Omara et al. (2016)	SW PA/Marcellus	6/2014—2/2015	18	Conventional NG production (CvNG)	Tracer flux	0.02—4.9	0.34—88	0.68—44
(a)	Omara et al. (2016)	SW PA/Marcellus	6/2014—2/2015	13	Unconventional NG production (UNG)	Tracer flux	0.85—93	0.02—1.22	450—78,000
(b)	Brantley et al. (2014)	Barnett	2010—2013	43	O&G production, mixed	OTM-33A	0.12—20	0.0001—31	0.4—39,300
(b)	Brantley et al. (2014)	Denver-Julesburg (DJB)	2010—2013	74	O&G production, mixed	OTM-33A	0.047—26	0.047—88	4.9—1,830
(b)	Brantley et al. (2014)	Pinedale	2010—2013	107	O&G production, mixed	OTM-33A	0.13—28	0.006—100	4.6—9,000
(b)	Brantley et al. (2014)	Eagle Ford	2010—2013	4	O&G production, mixed	OTM-33A	0.46—8.9	0.03—8.7	78—2,000
(c)	Rella et al. (2015)	Barnett	10/2013	185	O&G production, mixed	MM-Gaussian	0—48	0—90	2.3—6,000
(d)	ERG (2011)	Fort Worth/Barnett	2011	287	O&G production, mixed	Direct onsite measurements	0.0004—46	0.0001—31	0.4—39,300
(e)	This study	Denver-Julesburg (DJB)	04/2015	20	O&G production, mixed	Tracer flux, OTM-33A, MM-Gaussian	0.04—140	0.08—42	1—5,470
(e)	This study	Uintah	05/2015	30	O&G production, mixed	Tracer flux, OTM-33A, MM-Gaussian	0.004—44	0.01—83	4—3,580
(e)	This study	NE PA/Marcellus	05/2016	46	Unconventional NG production (UNG)	Tracer flux, MM-Gaussian	0.005—17	0.0004—28	40—25,200
(f)	Goetz et al. (2015)	SW PA/Marcellus	09/2012	3	Unconventional NG production	Tracer flux	3.4—14	0.06—0.27	4 670—8,360
(g)	Lan et al. (2015)	Barnett	10/2013	33	O&G production, mixed	MM- Gaussian	0.01—58	0.01—100	4.5—4,150
(h)	Yacovtich et al. (2015)	Barnett	2013	7	O&G production, mixed	MM-Gaussian	6—287	15—55	23—1,160

The different measurement methods also have different measurement uncertainties, which can range from approximately +/-30% to a factor of 3x the mean CH_4 emission rate (e.g., Brantley et al. (2014); Omara et al. (2016); Yacovitch et al. (2016)). We use the reported average site-level CH_4 emission rate data as is, provided site-level gas production rate at the time of measurement was reported as well. Gas production rates were required to ensure emissions data for only actively producing sites were included in the combined dataset. Furthermore, we use the gas production rates to estimate normalized CH_4 leak rates, which are defined as the fraction of site-level total CH_4 production that was emitted to the atmosphere.

Overall, the 870 sites included O&G production sites in the Barnett (555 sites), the Denver-Julesburg (94 sites), Eagle Ford (4 sites), Marcellus/southwestern PA/northern WV (34 sites), Marcellus/northeastern PA (46 sites), Uintah (30 sites), and Pinedale (107 sites) O&G production plays. These sites had site-level production rates ranging from <1 Mcfd to 78,000 Mcfd (Table 3). On the other hand, measured site-level absolute CH_4 emissions rates ranged from ≈ 0 to 300 kg/h (mean = 4.3 kg/h; median = 0.84 kg/h), while normalized CH_4 emissions rates ranged from ≈ 0 to 100% of site-level CH_4 production (mean = 3.43%; median = 0.26%).

The distributions of site-level CH_4 emission rates in each study are shown in Figure 17, in which reported emissions <0.05 kg/h CH_4 (absolute emission rates (Abs)) or <0.05% CH_4 (normalized emission rate (Norm)) are truncated to allow better visualization of the bulk of the emissions data in each study. Figure 17 shows that the majority of the sampled sites (72% of sites) exhibited emission rates between 0.1 and 10 kg/h CH_4, while just 8% of sites emitted at rates greater than 10 kg/h, with the remaining 20% of sites emitting at < 0.1 kg/h. Additionally, with the exception of the seven measurements by Yacovitch et al. (2015) that were targeted toward high emitting sites and low-producing, low-emitting conventional well pad sites in SW PA (Omara et al. (2016)), the study-specific median normalized site-level CH_4 emission rates were generally below 10% of CH_4 production.

We use these results from the combined dataset (i.e., all site-level absolute (All Abs) in Figure 17) in conjunction with their site-level production rates to scale up site-level CH_4 emissions to regional/basin-level totals.

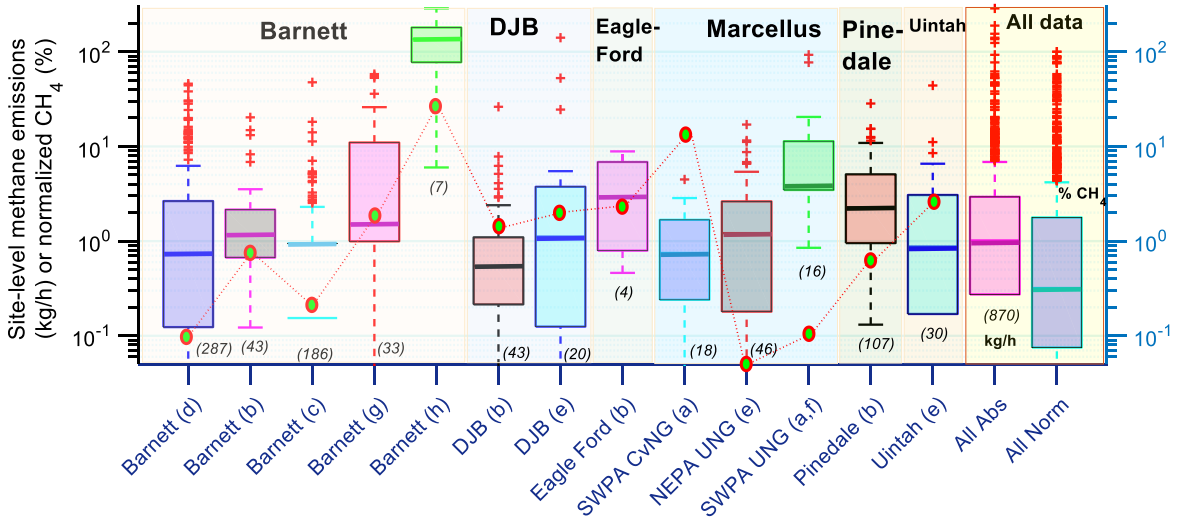

Figure 17. Box plots showing site-level absolute CH_4 emission rates (kg/h) and normalized CH_4 emission rates (% of total CH_4 production) for each of the six O&G production basins/regions. The boxes represent the 25th and 75th percentiles, while the whiskers extend to 1.5x the inter-quartile range, and values outside this range are the outliers, marked with red crosses (+). The horizontal line inside each box represents the median value. Study-specific

median normalized CH_4 emission rates are plotted in green circles with red edges connected with red dotted lines. Refer to Table 1 for reference codes (a) to (h). All Abs and All Norm represent combined dataset for all absolute and normalized CH_4 emission rates, respectively.

3.1.3 Statistical methods. Individual mean site-level CH_4 emission rates were grouped by production cohort identical to the scheme used for grouping well pad sites by production cohort (i.e., site-level CH_4 emission rates were grouped based on production rates of <10 Mcfd, 10 to 100 Mcfd, 100 to 1000 Mcfd, and >1000 Mcfd). A three-step Monte-Carlo simulation was performed to estimate the total CH_4 emissions in a given state/region. First, for each production cohort, the initial dataset of site-level CH_4 emissions was randomly resampled, with replacement, 10,000 times. This exercise produced 10,000 observations of possible CH_4 emissions distributions, each of the same length as the original dataset in each production cohort. A frequency distribution characterizing the contribution of the top 5% of the high-emitting sites to total CH_4 emissions was determined based on these observations.

Second, for each production cohort, each possible CH_4 emission distribution was randomly resampled, with replacement, each time assigning a single site-level CH_4 emission rate until each well pad (within that production cohort) in a given state/region was assigned a site-level CH_4 emission rate. Thus, for each production cohort, the contribution of the top 5% of sites to total CH_4 emissions in the state/region was constrained by the modeled frequency distribution of high emitting sites based on the empirical observations. Lastly, we use Monte-Carlo methods to sum up the results for each production cohort and to obtain resampling statistics (mean, median, 2.5^{th} percentiles and 97.5^{th} percentiles) for each state/region of interest.

3.2 Results and Discussion

3.2.1 Characteristics of site-level CH_4 emissions. We observed significant differences in key statistical parameters characterizing site-level CH_4 emission rates in each production cohort. The arithmetic mean CH_4 emission rate was lowest for sites producing <10 Mcfd (0.73 kg/h) of natural gas and highest for sites producing >1000 Mcfd (6.88 kg/h; Table 2). This result agrees with the weak, but positive correlations between absolute site-level CH_4 emissions and natural gas production rates observed in previous studies (e.g., Brantley et al. (2014); Lyon et al. (2016)). Additionally, the maximum amount of CH_4 emissions a site can emit is limited by its natural gas production rate and, as such, one would expect potentially higher CH_4 emission rates at sites with higher gas production rates, all other factors remaining constant. More importantly, large gas production sites often have comparably more auxilliary surface equipment (e.g., storage tanks, pneumatic pumps and controllers, separators, etc.) that are often significant sources of CH_4 emissions, as evidenced in southwestern Pennsylvania (Omara et al. (2016)).

Table 4. Statistical parameters characterizing site-level CH_4 emissions rates for each production cohort.

		Absolute CH_4 emission rates (kg/h)				**Normalized CH_4 emission rates (%)**			
Production cohort (Mcfd)	Total count of sites	Mean	Median	2.5th percentile	97.5th percentile	Mean	Median	2.5th percentile	97.5th percentile
<10	33	0.73	0.25	0.00	3.31	23.9	9.76	0.15	88.3
10 to 100	177	1.84	0.49	0.00	12.6	8.22	1.87	0.00	69.2
100 to 1000	384	3.83	0.68	0.00	22.1	1.54	0.28	0.00	10.9
>1000	276	6.88	1.58	0.00	52.7	0.52	0.09	0.00	2.84

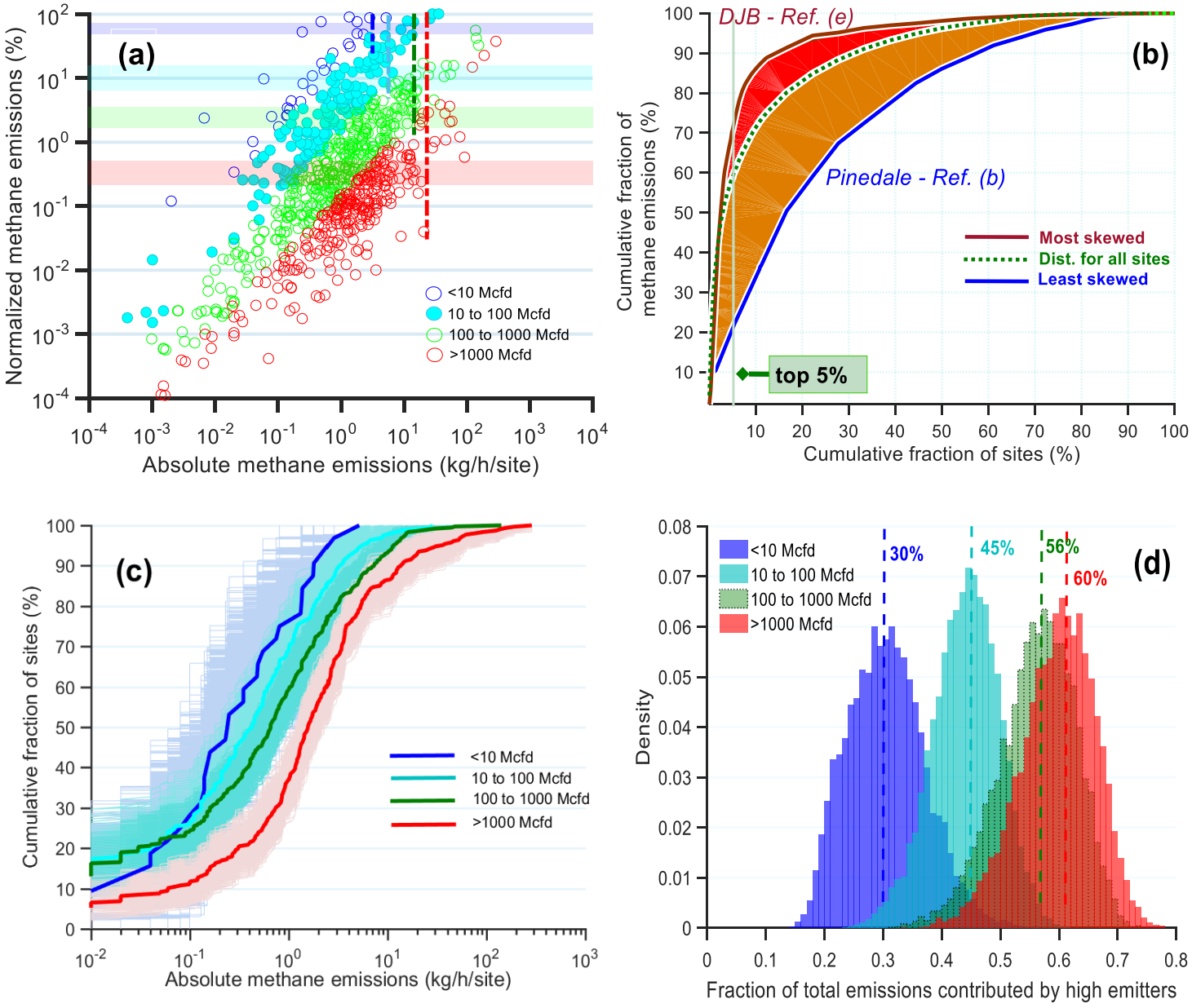

Figure 18. Characteristics of site-level CH_4 emissions data: (a) relationships between normalized and absolute CH_4 emission rates. The dotted lines delineate the absolute CH_4 emission rate threshold for the top 5% of sites in each production cohort, while the colored bands delineate the 80th to 90th percentiles of the normalized CH4 emission rates in each production cohort; (b) Empirical distribution of CH_4 emissions highlighting the most skewed (DJB; Ref. (e)), the least skewed (Pinedale; Ref. (b)) and the distributions from the combined dataset; (c) CDF of absolute CH_4 emissions, showing results of the first step of the Monte-Carlo simulation (shaded bands) for each production cohort. The solid lines show the empirical distribution; (d) The contribution of the top 5% of high emitting sites to total CH_4 emissions in each production cohort.

Conversely, the mean site-level normalized CH_4 emissions rates were highest for sites producing <10 Mcfd (23.9%) and lowest for sites producing >1000 Mcfd (0.52%, Table 4) in part because normalized CH_4 emission rates are functions of the inverse of gas production rates. This inverse correlation also reflects the random nature of abnormal process operating conditions (e.g., malfunctioning process equipment such as malfunctioning regulators, loose valves and fittings, etc.) at O&G production sites (Zavala-Araiza et al. (2015), (2017)). Such abnomal operating conditions often yield high absolute site-level CH_4 emission rates with a greater impact on normalized CH_4 emission rates at old, low producing sites that tend to have less-frequent inspection and maintenance routines (Omara et al. (2016)). Figure 18 (a) highlights these relationships between site-level normalized and absolute CH_4 emission rates for the different gas production cohorts. For example, 246 sites (28% of all sites) had normalized CH_4 emission rates > 1%, with a significant majority (93%) being sites that produced <1000 Mcfd. Similarly, 466 sites (54% of all sites) had absolute CH_4 emission rates < 1 kg/h, with more than three-quarters (77%) being sites that produced <1000 Mcfd.

Figure 18(b) demonstrates the variability in the distribution of site-level absolute CH_4 emission rates (sorted in descending order) in the different studies/regions. For simplicity, only the least-skewed and most-skewed distributions and the distribution from the combined dataset are presented. The results show that the occurrence of high emitting sites vary substantially in the different studies/regions, with the top 5% of high emitting sites accounting for between 20% and 70% of cumulative absolute CH_4 emissions. These variabilities likely result from differences in study-specific scope and measurement design (e.g., measurement technique, sample sizes, and/or site representativeness) and from regional differences in the likelihood of the occurrence of high emitting sites (Lyon et al. (2016)). These results imply that a single CH_4 emission distribution obtained from the consolidated dataset (i.e., the top 5% accounts for ≈ 60% of cumulative CH_4 emissions (Figure 2 (b))) may not adequately characterize actual CH_4 distributions in any given region at any given time (Brandt et al. (2016)).

To account for the variabilities in CH_4 emissions distributions, we use instead the results of the Monte-Carlo resampling analysis, in which total CH_4 emissions in any given region are constrained by both the distribution of sites (by production cohort) and the distribution of site-level CH_4 emission rates (by production cohort). Figure 18(c) shows the results of the first step of the Monte-Carlo resampling, producing a wide range in possible site-level CH_4 emissions distributions for each production cohort (shaded bands). We found that the top 5% of sites emit CH_4 at rates greater than 2.5, 7.0, 13, and 21 kg/h for the <10 Mcfd, 10 to 100 Mcfd, 100 to 1000 Mcfd, and >1000 Mcfd production sites, respectively. As expected, the median normalized CH_4 emission rates for these sites (i.e., the top 5% of sites in each production cohort) trended in the opposite direction, and were 88%, 77%, 11%, and 2.5% of total site-level CH_4 production, respectively. The resulting variability in the contribution of the top 5% of high emitting sites to total CH_4 emissions is shown in Figure 18(d), with mean values of 30%, 45%, 56%, and 60% for the <10 Mcfd, 10 to 100 Mcfd, 100 to 1000 Mcfd, and >1000 Mcfd sites, respectively.

We note that there are considerably fewer samples for sites <10 Mcfd in the combined dataset (Table 4), which introduces large uncertainties in the modeled CH_4 emissions distribution. Nevertheless, with each increasing level of gas production cohort, these results clearly show a pattern of increasing mean absolute CH_4 emissions and increasing impact on total/regional CH_4 emissions contributed by the top 5% of high emitting sites. One implication of this result is that relatively high absolute CH_4 emissions (and not necessarily high normalized CH_4 emissions) are more likely in regions dominated by sites with high gas production rates.

Relatedly, while high CH_4 emitting sites are present in every O&G production region (e.g., Lyon et al. (2016); Figure 18), our results suggest their cumulative impact on regional CH_4 emissions are likely greater in regions dominated by high gas production sites.

Using the results in Figure 18, we identify four characteristics of CH_4 emitting sites, taking into consideration the interplay between site-level absolute and normalized CH_4 emissions, and group sites into four quadrants shown schematically in Figure 19. The first quadrant includes O&G production sites that have low absolute CH_4 emissions and low normalized CH_4 emissions. Quadrant 1 sites may be either low or high gas producing sites and their low overall CH_4 emissions are indicative of frequent inspection and maintenance and operator's adherence to best management practices.

The second quadrant identifies sites that have low absolute CH_4 emissions, but high normalized CH_4 emissions. Quadrant 2 sites are dominated by low gas producing sites (generally <100 Mcfd) and may have less frequent inspection and maintenance. Their comparably lower gas production rates indicate they are much older compared to Quadrant 1 sites. Some CH_4 emissions at these sites may include wear and tear leaks from aging surface infrastructure.
The third quadrant identifies sites that have high absolute CH_4 emissions, but low normalized CH_4 emissions. Quadrant 3 sites are high gas producing sites (generally >100 Mcfd) with considerably more auxiliary surface equipment (e.g., wells, separators, and storage tanks) than Quadrant 2 sites. Major sources of CH_4 emissions at these sites may include emissions associated with the normal functioning of process equipment (e.g., pneumatic controller bleed and uncontrolled storage tank vents), which yield large site-level CH_4 totals because of the large number of leaking components.

The last quadrant identifies sites that have high absolute CH_4 emissions and high normalized CH_4 emissions. Quadrant 4 sites may be low or high gas producing sites. Their exceptionally high CH_4 emissions (both on an absolute and normalized basis) are indicative of the presence of one or multiple abnormal process conditions (e.g., malfunctioning gas processing equipment). The random occurrences of such high CH_4 emission events mean that any site in the other three quadrants can become a Quadrant 4 site at any point in time. These exceptionally high CH_4 emissions often persist until the emissions sources are identified and resolved.
From a CH_4 mitigation perspective, sites in Quadrants 3 and 4 with their high absolute CH_4 emission rates are clear targets for CH_4 mitigation, if a known site-level CH_4 emission threshold can be specified. In the combined dataset of site-level CH_4 emissions (Figure 17), we found median absolute and normalized CH_4 emissions of 0.84 kg/h and 0.26%, respectively. Based on these results, we propose a conservative site-level threshold of 1 kg/h and 1% CH_4 emission rate to categorize gas producing, CH_4 emitting sites using the four quadrants in Figure 19. Based on these thresholds, the results from the combined dataset indicate that 46%, 7.2%, 25%, and 21% of sites were Quadrant 1, Quadrant 2, Quadrant 3, and Quadrant 4 sites, respectively. As such, CH_4 mitigation opportunities would exist at 46% of the sampled sites (Quadrants 3 and 4) and could include frequent leak detection and repair programs (Ravikumar et al. (2017)), replacement of high-bleed pneumatic devices with low/zero-bleed devices, and/or use of high-efficiency

emission control devices (e.g., combustors and vapor recovery units).

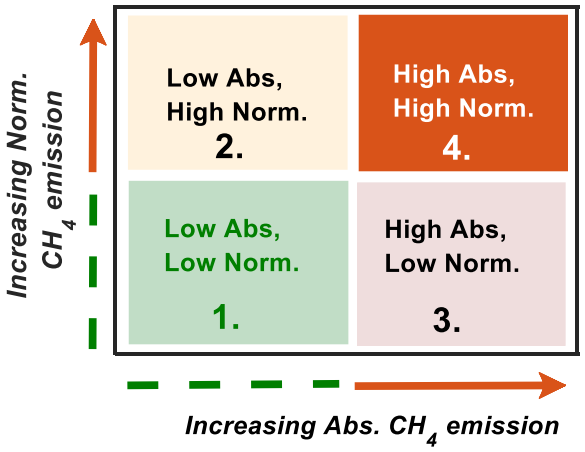

Figure 19. Schematic presentation of the characteristics of gas producing, CH_4 emitting sites, grouped into four quadrants representing (1) low absolute (Abs.) and low normalized (Norm.) CH_4 emissions (Quadrant 1); (2) Low absolute and high normalized CH_4 emissions (Quadrant 2); (3) High absolute and low normalized CH_4 emissions (Quadrant 3); and (4) High absolute and high normalized CH_4 emissions (Quadrant 4).

3.2.2 Estimates of total U.S. onshore CH_4 emissions from O&G production sites**.** We used the results of activity data (count of well pad sites in each state/region) grouped by production cohort and corresponding consolidated site-level CH_4 emissions data in a Monte-Carlo simulation to estimate total CH_4 emissions. From the 511,000 O&G production sites, we estimated total U.S. onshore CH_4 emissions of 700,000 kg/h (or 6.1 million metric tons) in 2015. These CH_4 emissions were equivalent to 1.43% of total CH_4 production in 2015, or 1.6 kg/h/site. The overall 95% confidence bound on these estimates, defined by the 97.5th and 2.5th percentiles, was estimated at +46%/-37%.

By looking at specific EIA sedimentary basins (EIA (2016)) and U.S states, interesting patterns in activity and regional CH_4 emissions emerge. As expected, the highest gas producing basin (Appalachian) and state (TX) had the highest CH_4 emissions, accounting for approximately one-fifth (19%) and one-third (32%) of total U.S. CH_4 emissions, respectively (Table 5). The Appalachian basin, encompassing gas production largely in PA, OH, WV, and KY, accounted for nearly one-third (32%) of total U.S. sites. Interestingly, 80% of all sites in this region produced <10 Mcfd of gas, with <2% of sites producing >1000 Mcfd (Supplemental Information). Thus, although the >1000 Mcfd sites in this region account for a significant fraction of total gas production (93%), the overall CH_4 emissions in this region are dominated by the emissions from the low producing sites (≈85% of CH_4 emisions come from <1000 Mcfd sites). These results highlight the importance of low gas producing sites, particularly in regions where they significantly outnumber high gas producing sites. A similar conclusion was reached in a previous study (Omara et al. (2016)).

Similarly, the dynamics between the regional distribution of gas production sites (by production cohort) and CH_4 emissions help explain the differences in overall CH_4 emissions in NM and OH (Table 5). NM and OH had similar total numbers of active O&G production sites and similar total 2015 gas production, but CH_4 emissions in NM were estimated to be approximately twice that of OH. We attribute these differences to the disparate distribution of the <10 Mcfd sites in the two states and their respective site-level CH_4 emissions characteristics; that is, sites producing <10 Mcfd accounted for 26% of the total sites in NM, but 92% of the total sites in OH. Relatedly, OK with 51,000 well pad sites had 38% more CH_4 emissions than PA, even though PA had 1.3x more well pad sites and 2.4x more gas production (Table 3).

These results imply that regional CH_4 emissions are not driven just by activity factors alone (e.g., total number of well pad sites) and are consistent with field observations of weak correlations between CH_4 emissions and quantifiable site-level or basin-level characteristics (Lyon et al. (2016)). The consolidated data in our study allow us to explore the dynamics between regional

Table 5. Summary of CH_4 emissions estimates highlighting the top ten EIA sedimentary basins and the top ten U.S states. Only active onshore O&G production sites are considered in these estimates. Overall uncertainties on CH_4 emissions estimates were +46%/-37%.

Top Ten EIA Sedimentary Basins (2015 production)						**Top Ten U.S. States**					
EIA Sedimentary Basin	Total # pads	Total gas prod (Bcf)	Mean CH_4 emissions (kg/h)	Mean CH_4 emissions (kg/h/site)	% CH_4 emissions	U.S. state	Total # pads	Total gas prod (Bcf)	Mean CH_4 emissions (kg/h)	Mean CH_4 emissions (kg/h/site)	% CH_4 emissions
Appalachian	165,000	7,200	136,000	0.83	1.1%	**TX**	124,000	8,800	224,000	1.8	1.3%
Permian	59,200	2,300	88,300	1.5	2.2%	**OK**	51,100	2,000	81,200	1.6	2.1%
Western Gulf	32,600	3,700	68,600	2.1	0.97%	**NM**	37,900	1,000	55,400	1.5	3.2%
TX-LA-MS Salt	39,700	2,800	68,200	1.7	1.3%	**PA**	67,800	4,800	58,800	0.87	0.68%
Anadarko	25,500	1,500	48,500	1.9	1.6%	**WV**	47,600	1,300	41,000	0.86	1.7%
Fort Worth	19,800	1,700	37,600	1.9	1.2%	**CO**	24,200	1,300	38,300	1.6	1.7%
Arkoma	13,200	1,400	28,400	2.2	0.99%	**LA**	19,400	1,700	30,300	1.6	0.92%
Denver	20,000	600	27,700	1.4	2.6%	**OH**	38,000	1,000	25,800	0.68	1.4%
Uinta-Piceance	12,100	1,000	23,700	2.0	1.3%	**WY**	14,400	1,800	29,400	2.0	1.0%
Williston	11,300	650	22,300	2.0	1.9%	**KS**	16,300	240	22,100	1.4	4.9%
Total U.S.	**511,000**	**27,000**	**700,000**	**1.38**	**1.43%**		**511,000**	**27,000**	**700,000**	**1.38**	**1.43%**

CH_4 emissions and the distribution of O&G production sites. However, we acknowledge that there are likely several other regional factors that explain regional-specific CH_4 emissions that our analysis cannot unravel based on the site-level information and data herein. For example, predominantly oil producing basins are likely to have different CH_4 emissions profiles than predominantly gas producing basins. Factors influencing regional CH_4 emissions may include variability in gas composition (i.e., CH_4-to-total hydrocarbon content), variability in the type, count, and configuration of onsite hydrocarbon processing equipment, and/or variability in operator practices. Furthermore, large basins are likely to have significant within-basin hetereogeneity in production characteristics that potentially yield different CH_4 emissions profiles in different fields. Additional studies are needed to assess such intra-basin variabilities. We assigned the average site-level CH_4 emissions rate (i.e., fraction of total CH_4 emissions to total count of well pad sites in each state, Table 5) to every identified well pad site in each state and generated a spatial distribution of site-level CH_4 emissions as shown in Figure 20. The gridded map highlights likely hotspots of site-level CH_4 emissions based on site count within each grid cell. For example, O&G production CH_4 emissions hotspots include areas like Weld County, CO (Denver-Julesburg), Uinta and Duchesne Counties, UT (Uinta basin), northwestern (San Juan) and southeastern (Permian) NM, and the liquids-rich fairway of the Appalachian basin (includes southwestern PA and northern WV).

A direct comparison of our CH_4 emission estimates with previously published results is limited by methodological differences. Top-down, airborne mass balance studies often provide regional CH_4 emissions estimates for the entire O&G supply chain (Supplemental Information) that are not directly comparable to our estimates for the O&G production sector. In this respect, a few studies in the Barnett Shale region (Lyon et al. (2015), Zavala-Araiza et al. (2015, 2016)) that extrapolated site-level CH_4 emissions to regional totals provide a useful point of reference. Using Barnett-specific site-level CH_4 emissions dataset (Rella et al. (2015), Lan et al. (2015), and Yacovitch et al. (2015)) and different scale-up approaches, Lyon et al. (2015) and Zavala et al. (2015, 2016) estimated 2013 Barnett CH_4 emissions of 17,000 to 38,700 kg/h and site-level CH_4 emission rates of 0.87 to 2.45 kg/h. In the present study, which consolidates the same Barnett-specific site-level measurements with other regional datasets, we found total 2015 CH_4

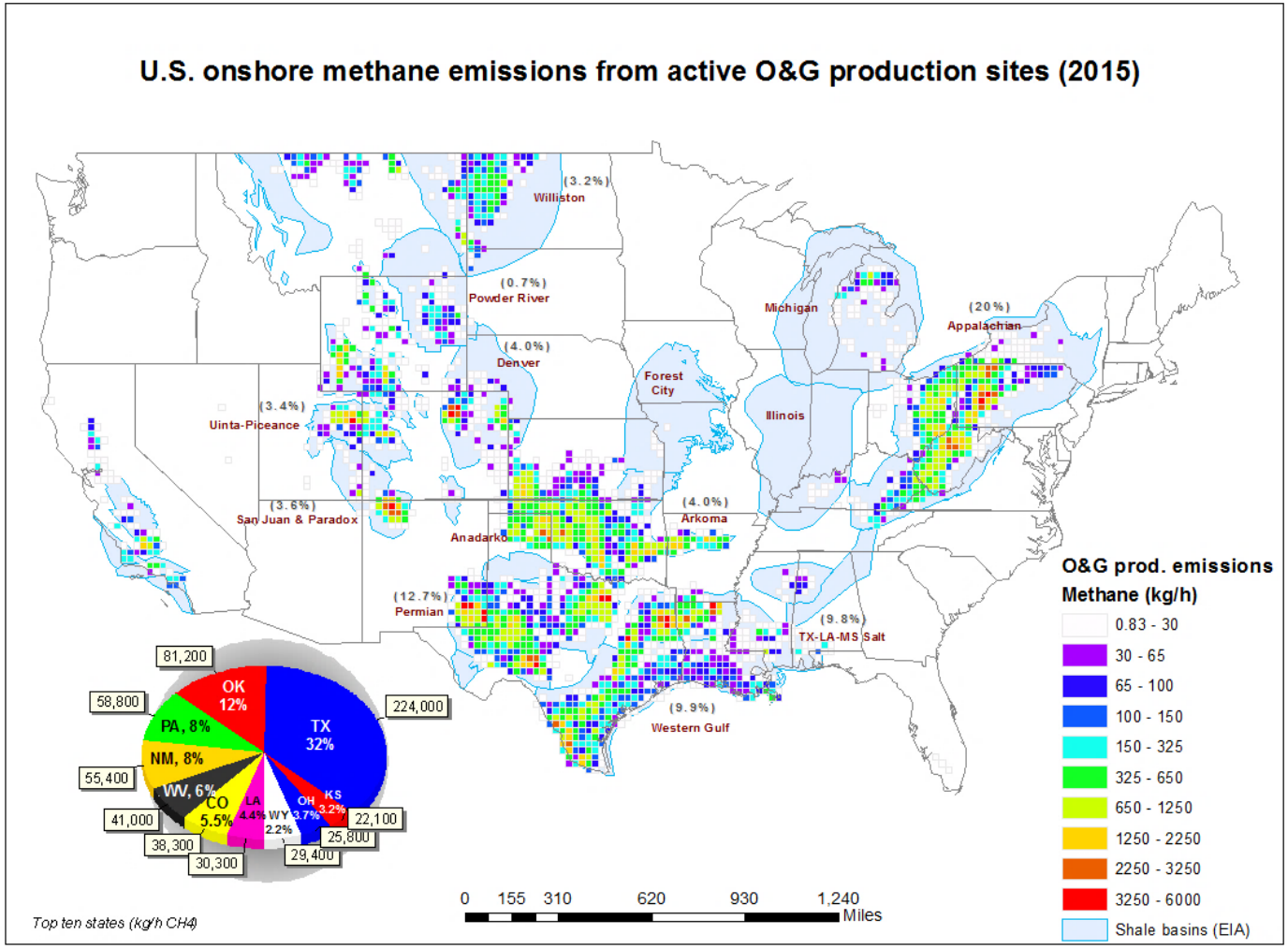

Figure 20. Spatial distribution of regional CH_4 emissions, plotted on a 35km x 35km grid cell. The numbers in parenthesis above basin names indicate the fraction of total U.S. CH_4 emissions contributed by that basin.

emissions of 37,600 kg/h or site-level CH_4 emissions of 1.90 kg/h (+33%/-28%) for the Fort Worth region, which are similar to the previous estimates. Our 2015 estimates of CH_4 emissions for the top two Appalachian Basin states (PA and WV, total CH_4 ≈ 100,000 kg/h) are ≈40% lower than the 2014 estimates by Omara et al. (2016), reflecting, in part, improved statistical power in the current study and methodological differences in the treatment of high emitting sites.

Total onshore CH_4 emissions from the natural gas and petroleum production sector (includes emissions from gathering and boosting stations) in the 2017 EPA Greenhouse Gas Inventory (EPA (2017)) was ≈644,000 kg/h or ≈1.1% (+60%/-20%), of total 2015 CH_4 production. A direct comparison with the present study is not possible given that the GHGI aggregates emissions from gathering and boosting compressor stations with that from actively producing O&G sites. However, our estimates of 700,000 kg/h or 1.4% (+46%/-37%) of total 2015 CH_4 production suggest a comparably lower CH_4 emissions estimates for the O&G production sector alone in the GHGI.

Training and professional development

The project was staffed by graduate students and a postdoc. Participating in the project contributed to their educational development. The project postdoc (Mark Omara) secured a permanent position with the Environmental Defense Fund following his work on this project. Graduate student Xiang Li leveraged data collected in this project to obtain a NASA graduate student fellowship.

Dissemination of results

Results were presented at:

- American Geophysical Union annual meetings in December 2014, 2015, and 2016
- Dr. Presto presented a seminar at the University of Pittsburgh in December 2014
- Dr. Presto presented a seminar at the Georgia Tech in November 2015
- Gordon Research Conference, Atmospheric Chemistry, August 2015
- EPA GHGI/GHGRP workshop, Pittsburgh, November 2015
- Dr. Presto presented the Mechanical Engineering department seminar at Carnegie Mellon University, March 2016
- RPSEA workshop, Denver, CO, May 2016
- Annual meeting of the American Meteorological Society, January 2017

Published Manuscripts

- Omara, M.; Sullivan, M.R.; Li, X.; Subramanian, R.; Robinson, A.L.; Presto, A.A. Methane emissions from conventional and unconventional natural gas production sites in the Marcellus Shale basin. *Environ. Sci. Technol.*, **2016**, DOI: 10.1021/acs.est.5b05503

Additionally, we prepared a report for the partner company that allowed us site access; the report summarized the measurements conducted at their sites.

Table 6. Milestone status report

Milestone Title	Planned completion date	Actual completion date	Verification method	Comments (progress towards milestone, deviations, etc.)
Completion of PMP	12/31/13	1/14/14	PMP submitted for DOE approval	
Literature search – methane emission rates from natural gas and other operations	1/1/14	1/1/14	Compile existing data on methane emissions for various sources	
Identification of Marcellus Shale sampling sites for spring/summer 2014 sampling	4/1/14	5/1/14	Identify 3-15 sampling sites, both with and without partner company buy-in, for sampling in Q3 and Q4 of FY14	
Identification of non-Marcellus Shale sampling sites for spring/summer 2014 sampling	4/1/14	9/1/14	Identify 2-10 sampling sites (land fills, abandoned wells, coal mines, etc.) for sampling in Q3 and Q4 of FY14	.
Identification sampling sites, sampling after Q4 FY14	9/1/14 and quarterly/semi-annually thereafter	9/30/15	Identify additional Marcellus and non-Marcellus sampling sites for measurements in FY15 and FY16 as per the project schedule and SOPO	
Completion of measurements	3/31/16	6/1/16	Database of methane and VOC emission rates for shale gas and non-shale gas sources	
Analysis of methane isotope ratios and ethane/methane ratios for incorporation into Monte Carlo model	12/31/16	3/31/17	PDFs and CDFs of methane isotope ratios and ethane/methane ratios, appropriately bootstrapped (if necessary) to provide coverage of the full data range	

PRODUCTS

One manuscript has been published to date: Omara, M.; Sullivan, M.R.; Li, X.; Subramanian, R.; Robinson, A.L.; Presto, A.A. Methane emissions from conventional and unconventional natural gas production sites in the Marcellus Shale basin. *Environ. Sci. Technol.*, **2016**, DOI: 10.1021/acs.est.5b05503

Results have also been presented at conferences and seminars.

PARTICIPANTS AND OTHER COLLABORATING ORGANIZATIONS

Individuals working on this project

Name:	Albert Presto
Project Role:	PI
Nearest person month worked:	3/2017
Contribution to Project:	Literature search and site selection. Contact with gas companies. Project management.
Funding Support:	This award; Heinz Endowments; CMU
Collaborated with individual in foreign country:	No
Country(ies) of foreign collaborator:	N/A
Travelled to foreign country:	N/A
If traveled to foreign country(ies), duration of stay:	N/A

Name:	Allen Robinson
Project Role:	co-PI
Nearest person month worked:	3/2017
Contribution to Project:	Project management.
Funding Support:	CMU
Collaborated with individual in foreign country:	No
Country(ies) of foreign collaborator:	N/A
Travelled to foreign country:	N/A
If traveled to foreign country(ies), duration of stay:	N/A

Name:	R Subramanian
Project Role:	co-PI
Nearest person month worked:	3/2017
Contribution to Project:	Project management, site selection.
Funding Support:	EDF; Heinz Endowments; NOAA
Collaborated with individual in foreign country:	No
Country(ies) of foreign collaborator:	N/A
Travelled to foreign country:	N/A
If traveled to foreign country(ies), duration of stay:	N/A

Name:	Eric Lipsky
Project Role:	co-PI
Nearest person month worked:	7/2015
Contribution to Project:	Project management, mobile lab maintenance.
Funding Support:	CMU, Penn St. Greater Allegheny
Collaborated with individual in foreign country:	No
Country(ies) of foreign collaborator:	N/A
Travelled to foreign country:	N/A
If traveled to foreign country(ies), duration of stay:	N/A

Name:	Mark Omara
Project Role:	post doctoral researcher
Nearest person month worked:	12/2016
Contribution to Project:	Site selection, method development, calibrations, dispersion modeling. Will conduct measurements.
Funding Support:	This grant
Collaborated with individual in foreign country:	No
Country(ies) of foreign collaborator:	N/A
Travelled to foreign country:	N/A
If traveled to foreign country(ies), duration of stay:	N/A

Name:	Naomi Zimmerman
Project Role:	post doctoral researcher
Nearest person month worked:	11/2016
Contribution to Project:	Drive-by measurements and analysis
Funding Support:	NOAA
Collaborated with individual in foreign country:	No
Country(ies) of foreign collaborator:	N/A
Travelled to foreign country:	N/A
If traveled to foreign country(ies), duration of stay:	N/A

Name:	Melissa Sullivan
Project Role:	Research Associate
Nearest person month worked:	8/2105
Contribution to Project:	Preparation of mobile laboratory for measurements, site selection

Funding Support: This grant; Heinz Endowments
Collaborated with individual in foreign country: No
Country(ies) of foreign collaborator: N/A
Travelled to foreign country: N/A
If traveled to foreign country(ies), duration of stay: N/A

Name: Xiang Li
Project Role: Graduate student
Nearest person month worked: 12/2016
Contribution to Project: Measurements.
Funding Support: This award; Heinz Endowments; CMU; NASA
Collaborated with individual in foreign country: No
Country(ies) of foreign collaborator: N/A
Travelled to foreign country: N/A
If traveled to foreign country(ies), duration of stay: N/A

Name: Hugh Li
Project Role: Graduate student (cost sharing)
Nearest person month worked: 2/2016
Contribution to Project: Measurements
Funding Support: Heinz Endowments; CMU; NSF
Collaborated with individual in foreign country: No
Country(ies) of foreign collaborator: N/A
Travelled to foreign country: N/A
If traveled to foreign country(ies), duration of stay: N/A

Name: Tim Dallmann
Project Role: Postdoc
Nearest person month worked: 4/2015
Contribution to Project: Assistance with measurements
Funding Support: Heinz Endowments; CMU
Collaborated with individual in foreign country: No
Country(ies) of foreign collaborator: N/A
Travelled to foreign country: N/A
If traveled to foreign country(ies), duration of stay: N/A

Name: Andrew Hix
Project Role: PhD student

Nearest person month worked:	2/2016
Contribution to Project:	Assistance with measurements
Funding Support:	Heinz Endowments; CMU; This award
Collaborated with individual in foreign country:	No
Country(ies) of foreign collaborator:	N/A
Travelled to foreign country:	N/A
If traveled to foreign country(ies), duration of stay:	N/A

Other organizations
RJ Lee, Inc – subcontract for use of PTR-MS to measure VOC concentrations.

IMPACT

Impact on the development of the principal discipline(s) of the project
The emissions data collected in SW PA and WV represent the largest dataset (at the time of publication) collected for conventional wells in the Marcellus region. The data collected in this project informs the contributions of Marcellus Shale sources to overall methane emissions both regionally and nationally. The national emissions estimate developed here will allow improved modeling of global methane. Also, as described above, we contributed to the methods used to quantify emissions from oil and gas facilities.

Impact on other disciplines
The nationwide methane emissions inventory has potential policy implications for implementation of LDAR (leak detection and repair) programs.

Impact on the development of human resources
The project funded one postdoctoral researcher, one research associate, PhD students. These personnel all received training on the methods used here, and had opportunities to present research results at meetings and conferences. Postdoc Mark Omara used this position to obtain employment with the Environmental Defense Fund. PhD student Xiang Li used data from this project to secure a NASA graduate fellowship.

Impact on physical, institutional, and information resources that form infrastructure
None to date

Impact on technology transfer
All results of this project have or will be disseminated in the form of presentations at conferences, scientific journal papers, and reports to DOE. We began disseminating research results in December 2014.

Impact on society beyond science and technology
This project improved our understanding of the environmental impacts of shale gas development. Data from this project and similar projects allow the public to have a more

informed opinion on the impacts of shale development, and will allow policy makers to form more effective policy and regulatory solutions. The national emissions inventory can be a useful policy tool for methane LDAR programs.

Dollar amount of the award's budget spent foreign country(ies)
None

CHANGES/PROBLEMS
Changes in approach and reasons for change
None.

Actual or anticipated problems or delays and actions or plans to resolve them
None.

Changes that have a significant impact on expenditures
None.

Significant changes in use or care of human subjects, vertebrate animals, and/or Biohazards
N/A

Change of primary performance site location from that originally proposed
N/A

BUDGETARY INFORMATION
Budget report sent separately from CMU.