Reconciling divergent estimates of oil and gas methane emissions

Supporting Information

Daniel Zavala-Araiza [a], David R. Lyon [a], Ramón A. Alvarez [a], Kenneth J. Davis [b], Robert Harriss [a], Scott C. Herndon [c], Anna Karion [d, e, 1], Eric A. Kort [f], Brian K. Lamb [g], Xin Lan [h, 2], Anthony J. Marchese [i], Stephen W. Pacala [j, 3], Allen L. Robinson [k], Paul B. Shepson [l], Colm Sweeney [d, e], Robert Talbot [h], Amy Townsend-Small [m], Tara I. Yacovitch [c], Daniel Zimmerle [i], and Steven P. Hamburg [a, 3]

a. Environmental Defense Fund, Austin, TX 78701
b. Carbon Now Cast, LLC, State College, PA 16802
c. Aerodyne Research Inc., Billerica, MA 01821
d. Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO 80309
e. Earth System Research Laboratory, National Oceanic and Atmospheric Administration, Boulder, CO 80305
f. Department of Climate and Space Sciences and Engineering, University of Michigan, Ann Arbor, MI 48109
g. Department of Civil and Environmental Engineering, Washington State University, Pullman, WA 99163
h. Department of Earth and Atmospheric Sciences, University of Houston, Houston, TX 77004
i. Department of Mechanical Engineering, Colorado State University, Fort Collins, CO 80523
j. Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544
k. Department of Mechanical Engineering, Carnegie Mellon University, Pittsburgh, PA 15213
l. Department of Chemistry, Purdue University, West Lafayette, IN 47907
m. Department of Geology, University of Cincinnati, Cincinnati, OH 45221

[1] Present address: National Institute of Standards and Technology, Gaithersburg, MD 20899.
[2] Present address: Earth System Research Laboratory, National Oceanic and Atmospheric Administration, Boulder, CO 80305
[3] To whom correspondence may be addressed. Email: pacala@princeton.edu or shamburg@edf.org.

Table of Contents

S1: Summary of published comparisons of top-down emissions to bottom-up estimates based on national data sources.

Top-down (TD) estimates of total or oil and gas CH_4 emissions regularly exceed bottom-up (BU) estimates based on emission inventories (26). Fig. S1 summarizes published comparisons between TD and BU estimates. (SC US: South-central US, DJ: Denver-Julesburg.)

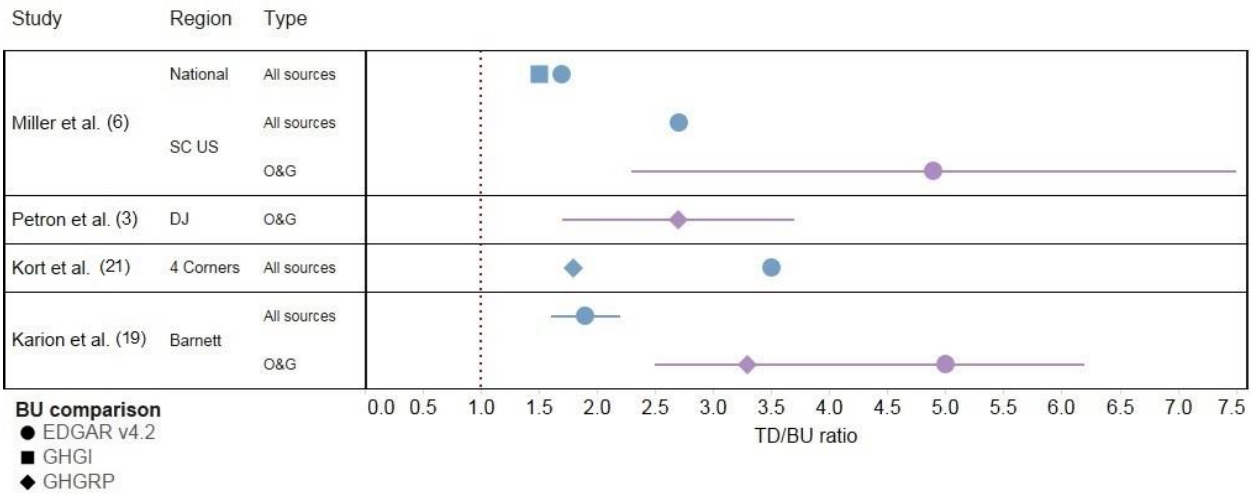

Fig. S1. Summary of published comparisons of top-down (TD) emissions to bottom-up (BU) estimates based on national data sources. Reported differences in CH_4 emissions from TD studies in U.S. gas producing areas relative to BU estimates based on one or more of the commonly cited national sources of CH_4 emissions data: the EPA Greenhouse Gas Inventory (GHGI) (36), the EPA Greenhouse Gas Reporting Program (GHGRP) (37), and the Emissions Database for Global Atmospheric Research version 4.2 (EDGAR v4.2) (38). Colors distinguish the types of sources quantified (blue = all sources; purple = oil and gas sources only, though some studies report all fossil sources, which include non-oil and gas sources like geologic seepage and coal mines). Symbol shapes denote the basis for the BU estimate used to calculate a ratio.

S2: Comparison between estimates of top-down fossil CH_4 from CH_4 and ethane measurements.

Because CH_4 is emitted by both fossil and biogenic sources, such as wetlands, feedlots and landfills, top-down (TD) CH_4 measurements estimate the combined fossil and biogenic flux. Because ethane has no biogenic sources and is the second most prevalent hydrocarbon in natural gas, the relationship of CH_4 and ethane can be used to estimate the fraction of emissions that originate from fossil sources. Two independent estimates of TD fossil CH_4 were used in the present analysis; similarity between them suggests robustness in our estimation of emissions from oil and gas infrastructure.

Fossil CH_4 emissions as reported by Karion et al. based on Smith et al.

Smith et al. (29) reported measurements of CH_4 and ethane from several flights in the study region. Measurements from one flight with light winds were used to determine the correlation of CH_4 with ethane during discrete 45-second observation windows. Smith et al. analyzed in detail the representativeness of the flight, and why it captured the methane/ethane variability of the rest of the flights. Using this data, they estimated the fraction of emissions attributed to fossil sources under a combination of different methods.

Karion et al. (19) used the results of Smith et al. (29) and reported that the fraction of CH_4 emissions from fossil sources in the Barnett Shale was is 79.5% (73.5% - 84%; 95% confidence interval (CI)). This central estimate was applied to the average TD CH_4 estimate and to each of the individual flight estimates (Fig. 1 of main paper). Following this method, the average TD fossil CH_4 emissions estimate for the Barnett Shale was 56 ± 10 Mg CH_4/h. (This result excludes an additional flight on October 25th, 2013; see SI S7 for details.)

Alternative estimate of fossil CH_4 emissions using top-down ethane fluxes.

Fig. S2A shows the remarkably consistent TD ethane estimates in Smith et al. (29). These six daily estimates ranged from 6.3 ± 3.5 to 6.8 ± 2.4 Mg C_2H_6/h and provide an alternative starting point to estimate fossil CH_4.

It is possible in principle to estimate fossil CH_4 emissions from TD ethane measurements if one has a quantitatively reliable mass-weighted mean estimate of the methane/ethane ratio of fossil emissions from the entire region. However, such a methane/ethane ratio is not well known because it varies considerably among oil and natural gas industry segments and source types (45). These estimates should be viewed with caution, given their sensitivity to the sampling strategy.

Specifically, the TD ethane estimate from each flight can be multiplied by a methane/ethane ratio (CH_4:C_2H_6, hereafter) that is representative of fossil sources in the surveyed region to give a fossil CH_4 estimate:

$$C_2H_6\left[\frac{kg}{h}\right] \times \left[\frac{CH_4}{C_2H_6}\right]_{fossil} = CH_{4\,fossil}[kg/h] \qquad \text{(S2.1)}$$

Smith et al. (29) identified three distinctive C_2H_6:CH_4 ratios in the region from examination of the series of discrete 45-second observation windows during one of their flights; one for microbial sources (0%), and two for fossil sources: a low-ethane fossil source ratio (C_2H_6:CH_4

centered around 1.8%; vol %), and a high-C_2H_6 fossil source ratio (C_2H_6:CH_4 centered around 9.6%; vol%). (Note that these ratios reported in Smith et al. are the inverse of the ratio shown in equation S2.1, and are volume based, instead of mass based.) Following a linear combination approach, their analysis suggests that for the study region, roughly 20% of the CH_4 emissions are from microbial sources, 35% from low-ethane fossil sources and 45% from high-ethane fossil sources; with all of these estimates having significant uncertainty bounds.

To determine the CH_4:C_2H_6 ratio from observations in Smith et al. (29) for use in equation S2.1, the microbial source can be ignored since there is no ethane associated with those sources. Thus 44% of the CH_4 emissions from fossil sources would have the low-ethane fossil source ratio (C_2H_6:CH_4= 1.8%; vol %) and 56% of the emissions the high-ethane fossil source ratio (C_2H_6:CH_4 = 9.6%; vol %). Using this linear combination approach the CH_4:C_2H_6 ratio (mass basis) would be 8.6. Consequently, if this CH_4:C_2H_6 ratio is applied to the average TD ethane estimate of 6.6 Mg C_2H_6/h, TD fossil CH_4 emissions would equal 57 tons CH_4/h. This estimate is not independent from the result reported by Karion et al. (19); however, it provides a CH_4:C_2H_6 ratio that can be compared to alternative independent methods and it also shows that although there is significant heterogeneity in the CH_4:C_2H_6 signals in the region (29, 45) the two distinctive signals from fossil sources are combined almost in equal parts. Thus, it is possible to determine a central estimate of the CH_4:C_2H_6 ratio if it is weighted in such a way that the weighting factor is representative of the surveyed region.

Reanalysis of the data collected by Smith et al. (29) can also provide a representative CH_4:C_2H_6 ratio (as an alternative method to the linear combination approach). Individual segments of CH_4 and ethane data from the representative flight were filtered for those cases with a significant CH_4:C_2H_6 correlation (p-value $\leq$0.05, using the 45-second observation window method, as described in Smith et al. (29)), estimating an individual CH_4:C_2H_6 ratio for each data point. For a total of 1,773 data points or windows, 83% had a significant CH_4 to ethane correlation. The integrated ethane enhancements measured for each data point were used as the weighting factor to derive an (ethane mass weighted) average CH_4:C_2H_6 ratio. Using this method, the CH_4:C_2H_6 ratio (mass basis) would be 8.4 ± 0.4. When this ratio is applied to the average ethane estimate, TD fossil CH_4 emissions would be 55 ± 4 Mg CH_4/h.

Independent fossil CH_4:C_2H_6 *ratio from ground based measurements reported by Yacovitch et al.*

Reanalysis of an independent dataset of ground-based measurements of CH_4 and ethane in the region was used to produce a characteristic fossil CH_4:C_2H_6 ratio that can be applied to the TD ethane estimates as described in equation S2.1. We caution that such a CH_4:C_2H_6 ratio is not well known because it varies considerably among oil and natural gas industry segments and source types and is sensitive to the sampling strategy used to collect the data (45).

Yacovitch et al. (32) estimated emission rates from 170 plumes of mobile ground-based atmospheric measurements. Plumes from this data set were filtered for those cases where CH_4 and ethane had a significant correlation ($R^2 \geq 0.85$; 60% of the plumes). For each plume in the final dataset an individual CH_4:C_2H_6 ratio was calculated, obtaining a distribution of CH_4:C_2H_6 ratios. Finally, the plume-specific ethane emission rates estimated by Yacovitch et al. were used as the weighting factor to produce an average CH_4:C_2H_6 ratio (mass basis) of 8.7 ± 2.8. (Yacovitch et al. reports estimated CH_4 emission rates for each plume; an ethane emission rate can be calculated using the individual CH_4:C_2H_6 ratios.) When this ratio is applied to the average

TD ethane estimate, fossil CH_4 emissions would be: 57 ± 18 Mg CH_4/h, which is in agreement with the TD fossil estimate of 56 ± 10 Mg CH_4/h (Fig. S2B).

The CH_4:C_2H_6 ratio derived from Yacovitch et al. (32) (8.7 ± 2.8; mass basis) is similar to the CH_4:C_2H_6 ratio that would be obtained from Smith et al. (29) (8.4±0.4; mass basis) and the resulting TD fossil CH_4 estimates produced from the different methods are not significantly different. Our CH_4:C_2H_6 estimate of 8.7 is also similar to those collected in ground canisters (6.5-9.8; mass basis converted from fossil CH_4 flux and Keeling plot based C_2H_6:CH_4; vol%) and within the range of natural gas composition reported in previous studies in the Barnett region (45). We reiterate that single regional CH_4:C_2H_6 ratios should be viewed with caution due to the considerable variability among oil and natural gas industry segments and source types (45). The sampling scheme in Yacovitch et al. was not designed to produce an unbiased estimate of ethane to methane and may have underrepresented areas of conventional oil production with ethane-rich emissions (45).

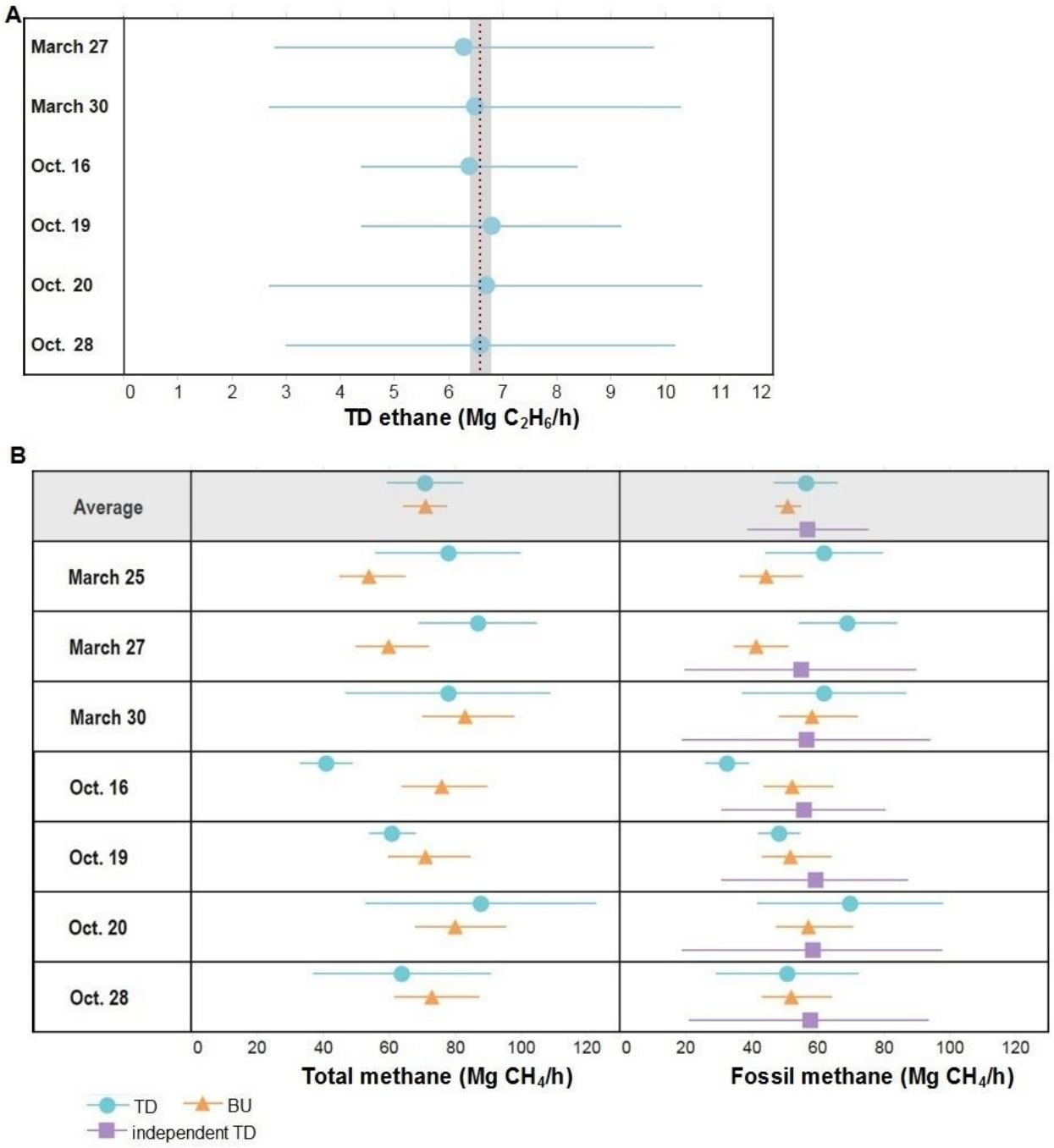

Fig. S2. Estimates of top-down (TD) fossil CH_4 from CH_4 and ethane measurements and comparison against custom bottom-up (BU) estimates (this work). (A) TD ethane estimates for the individual flights, as reported in Smith et al. (29). The red dotted line shows the average, with the shaded grey area showing the reported standard deviation of the mean. (B) TD and BU fossil CH_4 estimates. The blue (TD) and orange (BU) estimates were previously shown in Fig. 1 (main text). The purple estimates reflect the independent TD estimate that was determined in the present section; using the ethane estimates from Smith et al. and the fossil CH_4:C_2H_6 ratio (8.7 ± 2.8; mass basis) that was derived from reanalysis of Yacovitch et al. (32).

S3: Bottom-up estimates of CH_4 emissions: variations to statistical estimator.

In this section, we describe the two variations to the statistical estimator presented in the methods section of this work. We use the statistical estimator to produce emissions probability density functions (pdfs) from the systematic and high-emitter-biased samples, which are then used to derive emission factors for productions sites, compressor stations, and processing plants. These two variations produce similar results to the power law estimator that generated the results shown in the main body of this paper.

Variation 1 - Estimates using only the systematic samples of production sites, compressor stations and processing plants.

The facility-level empirical distributions are highly positively skewed, and become symmetric when plotted as the logarithm of emissions (Fig 3A and Fig 3B). This makes the lognormal distribution an obvious candidate for the pdfs. Lognormal distributions are common in problems of failure and breakage (46), and so it is not surprising to encounter them in this problem. A lognormal distribution is also expected if the leakage rate is the result of a series of independent or partially independent random numbers all multiplied together (because the logarithm of the product is the sum of the logs of the random numbers, which is a sum of random numbers and thus asymptotically normal by the Central Limit Theorem). Imagine that a valve has failed or been accidentally left open, and suppose that this connects the atmosphere, through a series of pipes, openings and chambers of different diameters, to a pressurized natural gas source (i.e. the reservoir). If *K* is the conductivity of the system, *P* is the pressure gradient from source to atmosphere, and *F* is the fraction of CH_4 in the gas, then the leak rate of CH_4 is the product *FPK*. Moreover, *K* is the harmonic mean of a series of conductivities within the system, which itself can be expressed as a product.

We assume throughout that the actual systematic emission rate distributions are lognormal. Let *x* be the logarithm of emissions and *x** be the smallest natural log emission rate that can be used with the "mobile flux plane method" of Rella et al. (30) (estimated to be -2.57 which is ln(0.08 CH_4/h)) or that of Mitchell et al. (12) (estimated to be -0.369 which is ln(0.691 kg CH_4/h)). Sites in the data sets with natural log emission rates below *x** are known only to have rates between minus infinity (zero on an arithmetic axis) and *x**.

Because *x* is normally distributed:

$$p\left(x\middle|\mu_j,\sigma_j\right) \equiv \frac{1}{\sqrt{2\pi}\sigma_j} e^{-\frac{\left(x-\mu_j\right)^2}{2{\sigma_j}^2}}, \qquad \text{(S3. 1)}$$

where the subscript *j* is 1 for productions sites, 2 for compressor stations, and 3 for processing plants.

Let $\Phi(x)$ be the cumulative standard normal:

$$\Phi\left(x\right) \equiv \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}}\, dv \qquad \text{(S3.2)}$$

Note that:

$$\int_{-\infty}^{x} p(v|\mu,\sigma)dv = \Phi\left(\frac{x-\mu}{\sigma}\right). \quad \text{(S3.3)}$$

With this notation, the log-likelihood function for the random sample of facility type *j*, with S_{0j} observations at or below the detection limit and S_{rj} observations above the detection limit, is:

$$\ell_1(\mu_j,\sigma_j) = S_{0j} ln\,\Phi\left(\frac{x_j^*-\mu_j}{\sigma_j}\right) - S_{rj} ln\sigma_j - \sum_{i=1}^{S_{rj}} \frac{(x_i-\mu_j)^2}{2\sigma_j{}^2}, \text{ j=1,2,3,} \quad \text{(S3.4)}$$

Where x_1^* is the lowest emission rate that can be used for the method of Rella et al. (30), and $x_2^*=x_3^*$ is the corresponding lowest rate for the method of Mitchell et al. (12). We obtained maximum likelihood estimates of the means and standard deviations in equation (S3.4) and their confidence limits in the usual way. The mean emissions rate on the arithmetic scale, M_j, is equal to: $\exp(\mu_j + \frac{\sigma_j^2}{2})$. This is the quantity we are after for the bottom-up (BU) estimates. Although maximum likelihood estimates of M_j are known to be biased, the bias is small in our study. Using the formula in Wilson et al. (47) the expected bias is ~3% for production sites and processing plants and ~1% for compressor stations. These values should be smaller for the larger samples available with the power law estimator and the second variation (below). Also, although measurement errors inflate the estimated variance, this effect is also small. Measurement errors for the systematic samples are on the order of 25% for Rella et al. (30) and 50% for Mitchell et al. (12) on the arithmetic scale. At 50% the variance of the logarithms would be inflated by ~0.2, while the estimates of the variances are between 2 and 5 (Fig. 3A and Fig. 3B, main text). While the estimated measurement error in Yacovitch et al. (32) is significantly higher than in the others (0.33x - 3.3x), which would be one unit on the logarithmic scale, it appears by the agreement between this first variation (which does not use Yacovitch et al.) and the main statistical estimator (which does) that the actual measurement error in Yacovitch et al. does not affect our estimations to a significant extent.

Variation 2 - Estimates that integrate the systematic samples (Mitchell et al., Rella et al.), and the high-emitter biased samples of Lan et al.and Yacovitch et al. by estimating the bias mechanistically based on Gaussian plume theory.

Each observation in the high-emitter-biased samples corresponds to a measured set of meteorological conditions (wind speed and stability class). Given a lognormal distribution of emissions rates, what is the likelihood of an observation of natural log emissions rate x_i conditional on the observed wind speed W_i and stability class s_i? Again, the bias in the high emitters sample occurs because large emission sources can be detected from farther away than small emission sources. For example, if the detectable length of a plume of strength x_{Large} is twice that of a plume of strength x_{Small}, then this doubles the probability density of sites in the high-emitter biased sample emitting at x_{Large} relative to those emitting at x_{Small}.

We used the Gaussian dispersion theory (43), to calculate the bias conditional on W_i and s_i. The functions for the vertical and horizontal standard deviations (σ_y and σ_z) were calculated using power-law expressions based on coefficients determined by Gifford (44). These parameters depend on stability class and distance from the source: $\sigma_y(s,D)$ and $\sigma_z(\mathrm{s},\mathrm{D})$. We can use these

formulae to calculate the maximum distance over which a source of strength $E=e^x$ could be detected along a plume's centerline as in the previous section, but here conditional on stability class and wind speed, by solving $\phi_0 = \frac{e^x B}{W\sigma_y(s,D^*)\sigma_z(s,D^*))}$ numerically for the function *D*(s,W,x)*. This gives maximum detection distance as a function of stability class, wind speed and the natural logarithm of the emissions strength. Thus, the likelihood of observing an emissions source of type *j* and strength x_k under observed conditions s_k and W_k is:

$$\mathrm{q}_j(\mathrm{x}_k|\mathrm{s}_k,\mathrm{W}_k) = \frac{D^*(\mathrm{s}_k,\mathrm{W}_k,\mathrm{x}_k)\frac{1}{\sqrt{2\pi}\sigma_j}e^{-\frac{(\mathrm{x}_k-\mu_j)^2}{2\sigma_j^2}}}{\int_0^\infty D^*(\mathrm{s}_k,\mathrm{W}_k,\nu)\frac{1}{\sqrt{2\pi}\sigma_j}e^{-\frac{(\nu-\mu_j)^2}{2\sigma_j^2}}\,d\nu}, \quad \text{(S3.5)}$$

and so the log likelihood functions are:

$$\ell_3(\mu_j,\sigma_j) = S_{0j}ln\Phi\left(\frac{x_j^*-\mu_j}{\sigma_j}\right) - S_{rj}ln\sigma_j - \sum_{i=1}^{S_{rj}}\frac{(x_i-\mu_j)^2}{2\sigma_j{}^2} + \sum_{k=1}^{S_{bj}}\ln(\mathrm{q}_j(\mathrm{x}_k|\mathrm{s}_k,\mathrm{W}_k)),\ j=1,2,3. \quad \text{(S3.6)}$$

The only unknowns in these equations are the means and standard deviations for each of the three types of facilities (production pads, compressor stations and processing plants). We solved for the values of these that maximize the above functions using a direct search algorithm and calculated 95% limits by inverting the Likelihood Ratio Test.

The results for each of the three estimation methods (Table S1) show that the estimates are robust to changes in method. Table S2 summarizes the different datasets used for each method. In summary, with the (main) statistical estimator we were able to integrate systematic samples with samples biased towards high emitters, producing emission factors representative of the entire population of sites.

Agreement between the power law estimator and variation 1 illustrate that even though dataset (iv) and data set (v) are of systematic samples, they did not under sample high emitters. Therefore, it is possible to use this first variation of the statistical estimator to produce representative emission factors. Agreement between the power law estimator and variation 2 shows that it is possible to estimate the bias of the high emitters mechanistically- by incorporating information related to the dispersion of emissions (under Gaussian dispersion theory) into the statistical estimator. This second variation produces an estimator that is internally consistent with the measurement approaches, since similar information about the dispersion of emissions (e.g., stability classes and meteorological conditions) is used by the sampling teams to estimate emission strengths from the measured concentrations (31, 32).

Table S1. Emission factors (EF) derived from the three statistical estimators. *Power law estimator*: Integrates systematic samples and high-emitter biased samples, using a power law to estimate the bias statistically. *Variation 1*: Estimates using only the systematic samples. *Variation 2*: Integrates systematic samples and high-emitter biased samples, estimating the bias mechanistically with a Gaussian plume model. EFs are determined for (A) productions sites, (B) compressor stations, and (C) processing plants. The 95% confidence interval (CI) of each parameter is shown between parentheses.

(A) Production sites

Method	μ	σ	EF kg CH_4/h
Power law estimator	-1.79 (-2.13, -1.45)	2.17 (1.96, 2.39)	1.76 (1.27, 2.45)
Variation 1	-1.84 (-2.20, -1.49)	2.21 (1.89, 2.52)	1.81 (1.73, 1.89)
Variation 2	-1.83 (-2.19, -1.47)	2.23 (1.91, 2.55)	1.94 (1.36, 2.75)

(B) Compressor stations

Method	μ	σ	EF kg CH_4/h
Power law estimator	3.05 (2.77, 3.32)	1.49 (1.32, 1.67)	64.2 (48.8, 84.4)
Variation 1	3.13 (2.86, 3.40)	1.38 (1.19, 1.57)	59.6 (45.9, 77.3)
Variation 2	3.13 (2.86,3.40)	1.38 (1.19,1.57)	59.4 (45.5, 77.6)

(C) Processing plants

Method	μ	σ	EF kg CH_4/h
Power law estimator	4.41 (3.92, 4.91)	1.31 (1.00, 1.62)	195 (121, 315)
Variation 1	4.34 (3.83, 4.85)	1.27 (0.912, 1.63)	173 (104, 285)
Variation 2	4.36 (3.86, 4.86)	1.26 (0.910, 1.62)	173 (106, 283)

Table S2. The table shows whether the sampling schemes behind each dataset were designed to produce representative samples of emission rates (systematic), or biased toward high emitters (high-emitter-biased); also shown for each dataset is N, the number of samples, as well as the range and mean of the reported emission rates (see Methods). This work's systematic sample of processing plants includes the combined set of processing plants and facilities classified as compression, dehydration, and treatment (C/D/T) sites in Mitchell et al. (12).

Source	Type of samples	Dataset, reference	N	Range (kg/h)	Mean (kg/h)	Used in method: (1)	(2)	(3)
Production sites	systematic	(iv), Rella et al. (30)	186	0.0 – 48	1.2	✓	✓	✓
	high-emitter-biased	(vi), Yacovitch et al. (32)	48	0.15 - 287	34		✓	✓
	high-emitter-biased	(vi), Lan et al. (31)	33	0.01 - 58	10		✓	
Compressor stations	systematic	(v), Mitchell et al. (12)	101	0.69 – 696	54	✓	✓	✓
	high-emitter-biased	(vi), Yacovitch et al. (32)	10	0.60 – 1,360	194		✓	✓
	high-emitter biased	(vi), Lan et al. (31)	7	21 – 2,119	698		✓	
Processing plants	systematic	(v), Mitchell et al. (12)	24	3.3 – 604	141	✓	✓	✓
	high-emitter-biased	(vi), Yacovitch et al. (32)	2	162 - 163	163		✓	✓
	high-emitter-biased	(vi), Lan et al. (31)	2	746 – 1,723	1,234		✓	

S4: Comparison between results using statistical estimators (this work) to those from Lyon et al.

Lyon et al. (27) used a two-step Monte Carlo simulation process to estimate emission factors (EF) for two of the major oil and gas emission sources (compressor stations and processing plants). The first step involved drawing emission rates from a "sampled distribution" (systematic, quasi-random sample), and the second step involved drawing emission rates from a "fat tail site-distribution" (biased towards high-emitters). A key step in this process is selecting the probability at which emission rates are drawn from the fat-tail distribution. Lyon et al. estimated this probability based on the number of observed fat-tail sites compared to the total sites in the region. Lyon et al. reported sensitivity tests showing the effect of alternative assumptions for the fat–tail probability within a range of 0% - 5%. The use of empirical data to set the probability of fat-tail sites—as done by Lyon et al.—requires arguably arbitrary choices of what constitutes the emission threshold for fat-tail sites and the frequency of fat-tail sites.

For natural gas producing sites, Lyon et al. (27) used a more complex approach than the two-step Monte Carlo process used for compressor stations and processing plants. This method, described in detail in Zavala-Araiza et al. (42), is based on characterizing the skewed distribution of production site emissions using the concept of functional super-emitters: sites with an excess of emissions resulting from abnormal or otherwise avoidable operating conditions, such as malfunctioning equipment. This excess of emissions can be characterized by looking at the proportional loss rate (CH_4 emissions relative to CH_4 produced), thus providing an alternative means of estimating the contribution of the fat-tail sites to the emissions from the entire population of sites.

Zavala-Araiza et al. (42) classified samples into gas production cohorts and then into three categories based on the percentile of proportional loss rates for each gas production cohort: α sites (below the 85^{th} percentile of proportional loss rates) and β and γ sites (equal or above the 85^{th} percentile of proportional loss rate), with β and γ sites being considered functional super-emitters. This method is sensitive to the selection of the threshold that distinguishes α-sites from functional super-emitters, as well as the probability of the highest functional super-emitters (γ sites). Both parameters were chosen based on empirical observations; however, the selection of such thresholds still requires a somewhat arbitrary choice to be made. Zavala-Araiza et al. report sensitivity analyses on the selection of both parameters.

The three statistical estimators that are presented in this work (see Methods and Section S2 in SI Appendix) provide an alternative way to derive emission distributions (pdfs) for the entire population of different oil and gas sources without having to choose specific thresholds or probabilities of fat-tail sites. Table S3 compares the emission factors reported by Lyon et al. (27) against the emission factors that were estimated in the present work. For the case of compressor stations the mean emission factors differ by 12%, with overlapping 95% confidence intervals (CI). This result suggests that the fat-tail frequency selected by Lyon et al. (1%) is similar to the frequency at which high emitters are represented in the pdf produced in this work (Fig. 3C and Fig. 3D, main text).

For the case of processing plants, Lyon et al. (27) divided this source into small processing plants (EF = 84 kg CH_4/h; 95% CI: 45 - 133 kg CH_4/h) and large processing plants (EF = 190 kg

CH_4/h; 95% CI: 112 - 301 kg CH_4/h). The average emission factor (weighted in terms of regional population of small and large processing plants) of 145 kg CH_4/h is 26% lower than the one presented in this work. The emission factors are within the CI of each other. Results from this work suggest that the frequency of the fat-tail processing plants should be greater than the 2% used by Lyon et al. If the frequency of fat-tail processing plants was set to roughly 9% (both for small and large processing plants), the method used by Lyon et al. would yield an emission factor that aligns with the one presented in this work.

The largest difference in emission factors was for production. Looking at the results of the statistical estimators in the context of the functional super-emitter framework presented by Zavala-Araiza et al. (42) would imply that the population of functional super-emitters in the Barnett Shale is bigger than the one assumed for that analysis (> 85th percentile of each gas production cohort) and 0.25% probability for γ sites. For a threshold that identifies functional super-emitters in the range of the 70th – 85th percentile, and a probability of γ-sites of roughly 1.5% (compared to a quarter of a percent used in Zavala-Araiza et al.), the average emission factor for production sites would be within a 10% difference from the one presented in this work.

Table S3. Comparison between emission factors (EF) from Lyon et al. (27) and the ones calculated in this work, following the statistical estimator methods described in the Methods section. Values shown are central estimates and 95% confidence interval (CI).

	EF (kg CH_4/h)	
Source	As reported in Lyon et al.(27)	This work [a]
Production sites	1.0 (0.97, 1.1) [b]	1.8 (1.3, 2.5)
Compressor stations	72 (50, 100)	64 (49, 84)
Processing plants	145 (84, 213) [c]	195 (121, 315)

[a] Emission factors shown for "this work" represent results from power law estimator (see Methods).
[b] Emission factors for production sites reported in Lyon et al. (27) were divided into gas production cohorts. The value reported here represents the average across all cohorts, excluding sites with no gas production (which had an emission factor of 0.01 kg CH_4/h). If those sites were considered, the average from Lyon et al. would be 0.87 (0.82, 0.92) kg CH_4/h and the average of the present work would be: 1.5 (1.1, 2.1) kg CH_4/h.
[c] Processing plants in Lyon et al. (27) were divided into small and large. The value reported here represents the average of both categories.

If the emission factors presented by Lyon et al. (27) were used to produce spatially resolved bottom-up (BU) estimates for each of the seven flights shown in Fig. 1 (main text), the average BU total CH_4 estimate would be 62 ± 7 Mg CH_4/h and the average BU fossil CH_4 estimate would be 42 ± 3 Mg CH_4/h, and these estimates would be within a 13% and 26% difference, respectively, from the top-down (TD) CH_4 estimates presented in this work (Table S10). When estimates reported by Lyon et al. are used, the importance of an accurate account of activity factors as well as the presence of high emitters is still highlighted (see Fig. S3, for spatial distribution of activity factors). Nonetheless, the estimation methods developed in the present work provide a more integrative way to obtain representative emission factors without the need to make arbitrary choices about the fat tail of the distributions and the frequency of high emitters.

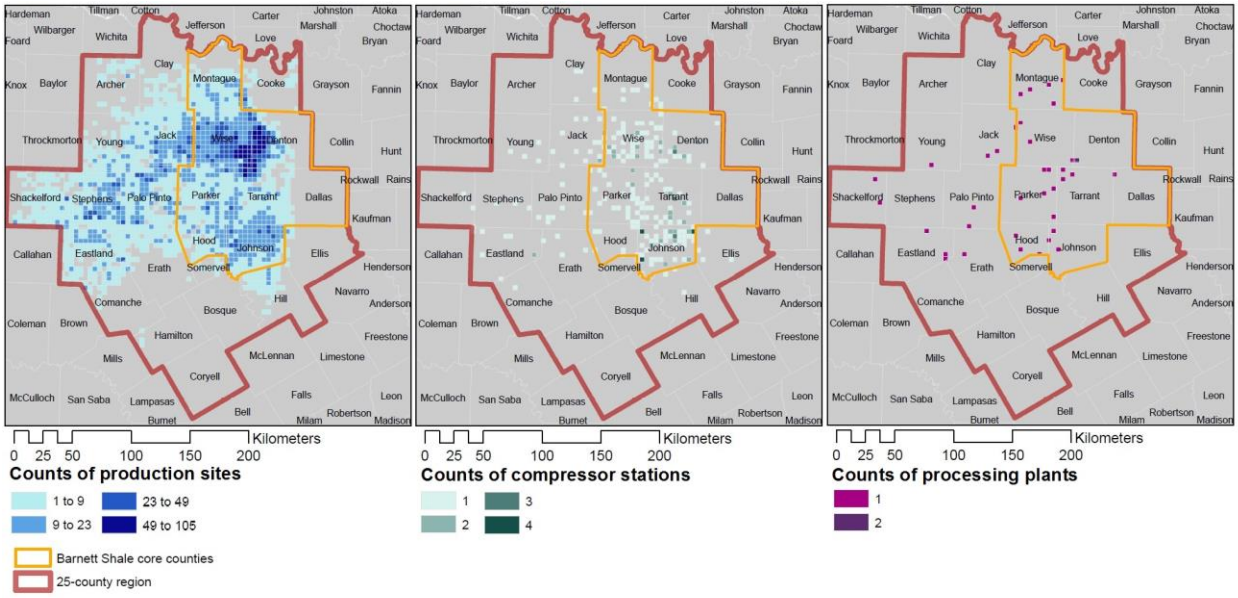

Fig. S3. Activity factors for the oil and gas major sources of CH_4 emissions aggregated into 4 km by 4 km grid cells in the 25-county Barnett region (total area is 58,000 km^2). Maps show counts of production sites, compressor stations, and processing plants per grid cell (see Table S4 for total counts). As a reference, the maps also show the 8 core county region (area is 17,400 km^2); this area encapsulates where most of the oil and gas sites and facilities are located (27) and was used as a point of comparison in previous studies (19).

Table S4 compares results reported in Lyon et al. (27) against the present work, when emissions are spatially resolved for the 25-county Barnett region (instead of the individual flights). The activity factors are the same as the ones reported in Lyon et al. For the present work, the only changes apply to the emission factors from production sites, compressor stations, and processing plants. Lyon et al.'s BU estimate of total CH_4 oil and gas emissions is 20% than the BU presented in this work.

Table S4. Comparison of bottom-up (BU) estimate of emissions for the spatially resolved 25-county Barnett region, for the same source categories previously estimated in Lyon et al. (27). Categories shown in bold reflect sources whose emission factors were updated in the present work using the estimation methods described in the Methods section (Table 1, main text).

Source	Activity factor reported in Lyon et al. (29)	Emissions reported in Lyon et al. (27) (kg CH_4 /h)	Emissions from this work (kg CH_4 /h)
Production sites	15,044 well pads 5,842 well pads	Gas production sites 16,400 (15,400 – 17,300) Oil production sites 1,800 (1,700 – 1,900)	**30,900 (22,100 – 43,200)**
Compressor stations	259 facilities 17 facilities	Gathering compressor stations 18,700 (12,900 – 26,000) Transmission & storage compressor stations 1,600 (850 – 1,700)	**18,100 (13,700 – 23,800)**
Processing plants	38 plants	5,500 (3,700 – 8,100)	**7,400 (4,600 – 12,000)**
Gathering pipelines	20,100 miles	940 (760 – 1,200)	940 (760 – 1,200)
Well completions	38 gas wells | 36 oil wells	150 (30 – 290)	150 (30 - 290)
Transmission pipelines	3,300 miles	230 (190 – 300)	230 (190 - 300)
Local distribution	5,730,000 inhabitants	930 (750 – 1,600)	930 (750 – 1,700)
O&G subtotal		46,200 (40,000 – 54,100)	**58,600 (48,400 – 73,021)**
Fossil subtotal		48,400 (42,100 – 56,400)	**60,700 (50,500 – 75,200)**
Biogenic subtotal		24,000 (17,200 – 30,100)	24,000 (17,600 – 29,800)
Emissions total		**72,300 (63,400 – 82,400)**	**84,700 (72,600 – 100,300)**

In summary, the estimation methods presented in this work provide additional confidence on the estimation of BU emissions, providing a method that seamlessly derives a continuous pdf that is representative of the entire population of sites (for each source), eliminating the necessity of dealing with the fat-tail of sites in a somewhat *ad hoc* manner.

The framework of functional super-emitters developed by Zavala-Araiza et al. (42) also provides a way to operationally identify sites with emission reduction opportunities. Future work will adapt the estimation methods developed in the present work to the functional super-emitter framework in such a way that there is no need for arbitrary assumptions about the thresholds. As a first step, section S5 analyzes the relationship between emissions and produced or processed gas. The potential policy implications in terms of emission reduction opportunities of accurately identifying functional-super-emitters are described in detail in Zavala-Araiza et al.

Because the source regions for the different flights differ from the 25-county Barnett source region (see Fig. S3 and Fig. S4), the oil and gas BU estimate presented in Table S4 (59 Mg CH_4/h) is different from the mean BU derived from the flight envelopes (50 Mg CH_4/h). Table

S5 and Table S6 compare BU estimates spatially resolved for the 25-county region as well as for each of the source regions from the individual flights. Table S6 also compares the 25-county BU estimate against publically available inventories (spatially resolved for the same source region).

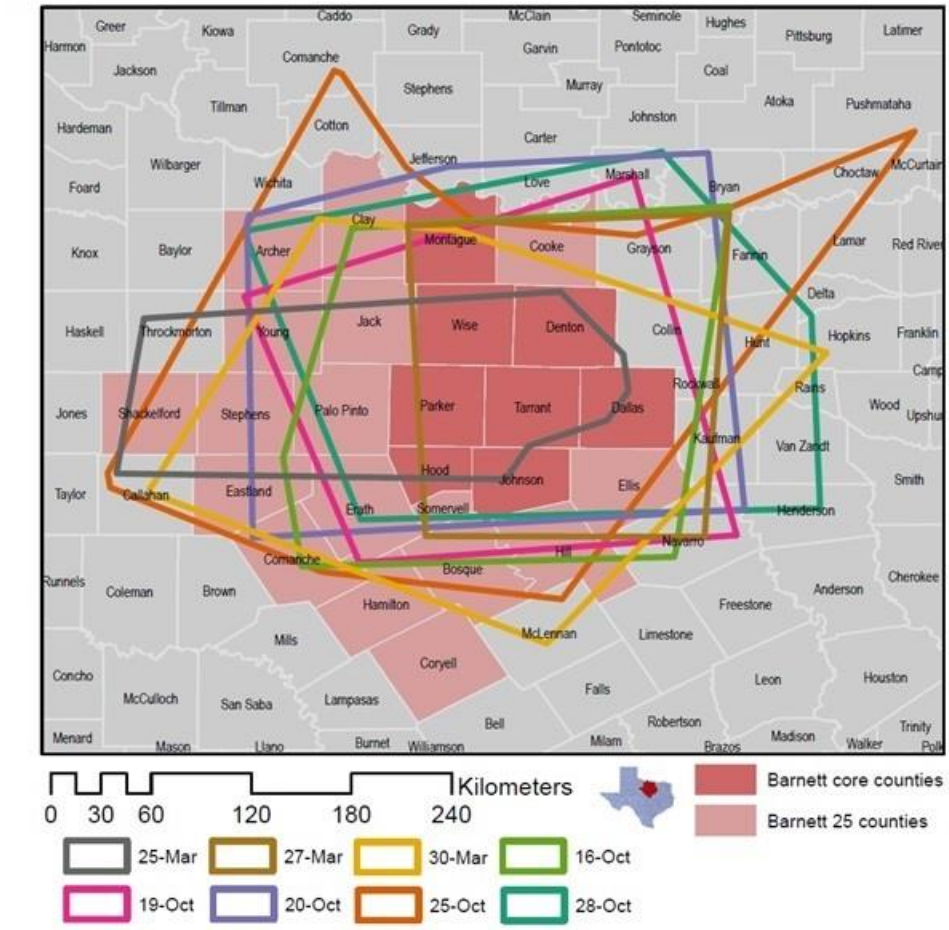

Fig. S4. Boundaries of the source regions estimated to be sampled by mass balance flights and for which bottom-up (BU) inventories were developed. Baseline flight envelopes provided by the authors of Karion et al. (19).

Table S5. Comparison of BU estimates for the source regions of each of the flights and the Barnett 25-county region (with 95% confidence intervals (CI)).

BU estimate	O&G emissions (Mg CH_4/h)	Fossil emissions (Mg CH_4/h)	Biogenic emissions (Mg CH_4/h)	Total emissions (Mg CH_4/h)
March 25, 2013	44 (36, 55)	44 (36, 55)	9 (6, 9)	54 (45, 65)
March 27, 2013	41 (34, 50)	41 (35, 51)	18 (11, 25)	60 (50, 72)
March 30, 2013	57 (47, 71)	58 (48, 72)	23 (16, 30)	83 (70, 98)
October 16, 2013	52 (43, 64)	52 (44, 65)	22 (15, 29)	76 (64, 90)
October 19, 2013	51 (42, 63)	52 (43, 64)	18 (11, 25)	71 (60, 85)
October 20, 2013	56 (46, 70)	57 (47, 71)	22 (14, 29)	80 (68, 96)
October 28, 2013	51 (42, 63)	52 (43 64)	20 (13, 27)	73 (62, 87)
Average from flights	50 (46, 54)	51 (47,55)	19 (16, 22)	71 (64,78)
Barnett 25-county	59 (48, 73)	61 (50, 75)	24 (18, 30)	85 (73, 100)

Table S6 Comparison of top-down (TD) and bottom-up (BU) CH_4 emission estimates and leak rates for the 25-county and flight envelope Barnett region with alternative emission estimates from Lyon et al. (27) based on the United States Environmental Protection Agency (EPA) Greenhouse Gas Inventory (GHGI) (36), EPA Greenhouse Gas Reporting Program (GHGRP) (37), and Emissions Database for Global Atmospheric Research v4.2 (EDGAR v4.2) (38) (with 95% confidence intervals (CI)). Calculation of CI is discussed in S7.

	BU 25-county	BU flight envelopes	TD flight envelopes
This work O&G CH_4 emissions (Mg CH_4/h)	59 (48 – 73)	50 (46 – 54)	56 (47 – 66) [b]
EPA GHGI O&G CH_4 emissions (Mg CH_4/h)	31.0	NA	NA
EPA GHGRP O&G CH_4 emissions (Mg CH_4/h)	17.0	NA	NA
EDGAR v4.2 O&G CH_4 emissions (Mg CH_4/h)	10.8	NA	NA
Gas production (Mg CH_4/h)[a]	3,945	3,782	3,782
Leak rate	1.5% (1.2 – 1.9%)	1.3% (1.2 – 1.4%)	1.5% (1.2 – 1.7%)

[a] October 2013 daily average gas production (standard cubic feet / day) was converted to a CH_4 flux using a regional production weighted-average gas production of 88.5% and a CH_4 density of 19.2 g / standard cubic feet.
[b] For the TD estimate, this row shows fossil CH_4 emissions, due to the fact that a TD estimate cannot differentiate between fossil emissions and oil and gas emissions. However, as shown in Table S5, there is only a small difference between the average fossil and oil and gas BU estimates for the flights.

S5: Analysis of high emitters as a function of produced or processed gas.

Here, we analyze the relationship between emission rates and produced or processed gas for each source. Even though facility emission rates depend only weakly on the total amount of gas produced or processed, facility level emissions as a fraction of the total CH_4 produced or processed (hereafter, proportional loss rate) is a more effective metric than absolute emissions to identify sites with avoidable emissions (e.g., malfunctioning equipment) (42).

We assume that the distribution of emission rates, as well as the distribution of produced or processed gas, are lognormal (see section S3). Let x_1 be the logarithm of emission rates and x_2 the logarithm of total CH_4 produced (for production sites) or total CH_4 processed (for compressor stations and processing plants). If we define the proportional loss rate as the ratio of CH_4 emitted to CH_4 produced/processed, x_3, then the logarithm of the proportional loss rate can be written as:

$$x_3 = x_1 - x_2 \tag{S5.1}$$

Using variation 2 of the statistical estimator (see section S3), it is possible to estimate the probability density function (pdf) of emissions conditional on production/throughput $P(x_1|x_2)$, when μ_1 is expressed as a linear regression of production:

$$\mu_1 = A + Bx_2 \tag{S5.2}$$

We estimate the pdf of x_1 (logarithm of emission rates conditional on production/throughput) with parameters (μ_1, σ_1), for each of the sources.

Similarly we can estimate the pdf of x_2 (logarithm of total produced or processed CH_4), with parameters (μ_2, σ_2) for each of the sources. Since both x_1 and x_2 are normally distributed, we can express the joint density as a bivariate normal distribution $D(x_1, x_2)$ (Fig. S5A):

$$D(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{z}{2(1-\rho^2)}} \tag{S5.3}$$

Where:

$$z \equiv \frac{(x_1-\mu_1)}{{\sigma_1}^2} - \frac{2\rho(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)}{{\sigma_2}^2}$$

And:

$$\rho \equiv \frac{V_{12}}{\sigma_1\sigma_2}.$$

Using equation S5.1, it is possible to express x_2 as a function of x_1 and x_3, thus expressing the joint density (equation S5.3) as a function of emission rates and proportional loss rates, $D(x_1,x_3)$. Consequently, the marginal probability of proportional loss rates $P(x_3)$ can be calculated as:

$$P(x_3) = \int_{-\infty}^{\infty} D(x_1, x_3)\, dx_1 \tag{S5.4}$$

Fig. S5B and Fig. S5C show the marginal probability of the proportional loss rates and the cumulative distribution of proportional loss rates. From this analysis it is possible to infer that 10% of the processing plants emit at least 0.6% of the gas they process. A higher proportion of compressor stations and production sites have higher proportional loss rates: 10% of the

compressor stations emit at least 3.6% of the gas they process and 10% of the production sites emit at least 5.2% of the gas they produce.

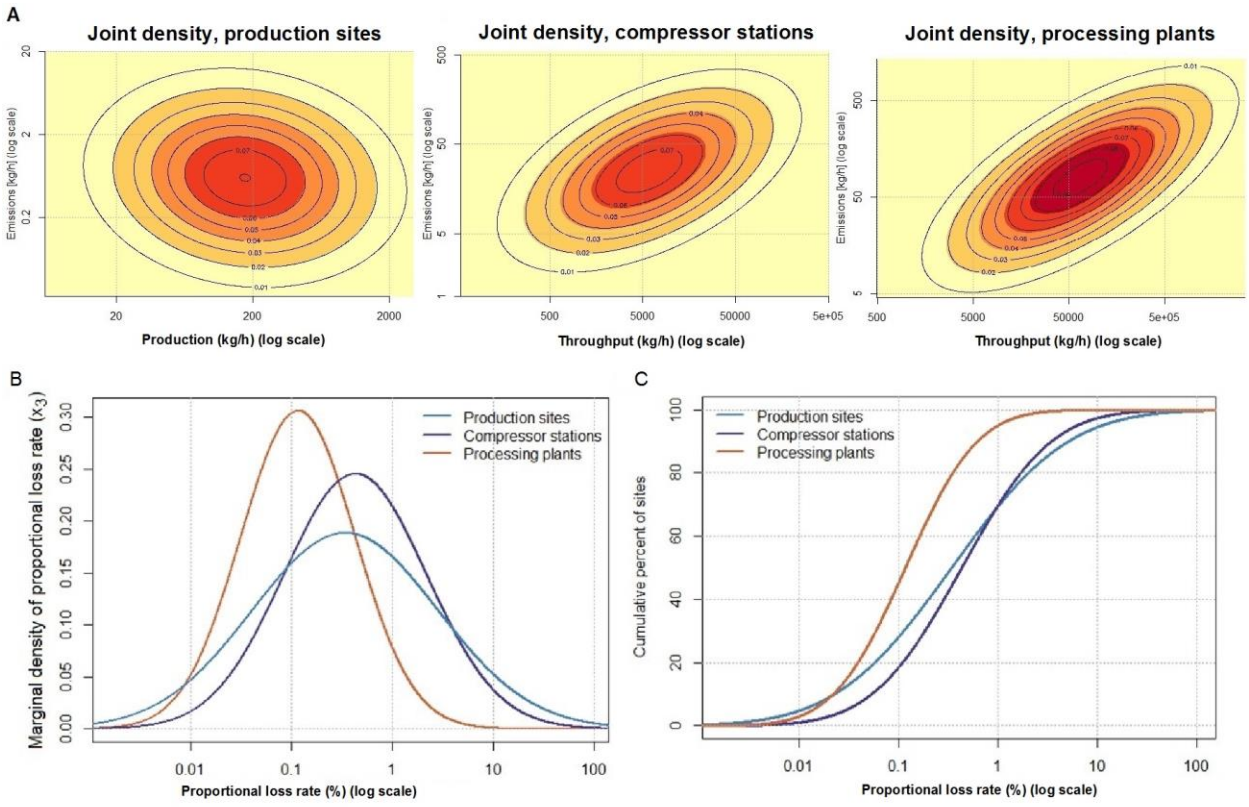

Fig. S5. High emitters as a function of production or throughput. (A) Contour plots of the joint density of emission rates and production/throughput for each of the sources. The color shows the increase in the joint density (darker red, greater density). (B) Marginal probability of the proportional loss rate of each of the sources. (C) Cumulative distribution of the marginal probability of proportional loss rates for each source.

S6. Climate implications of electricity generated with Barnett natural gas vs. coal.

All data and calculations used in this analysis are available in the file: Zavala et al TWP.xls.

Materials and methods.

Using the Technology Warming Potential (TWP) framework of Alvarez et al. (1), we compare the climate implications of electricity generated from a power plant using natural gas produced in the Barnett Shale region relative to electricity generated at typical coal-fired power plants in Texas. The TWP is the ratio of the total cumulative radiative forcing, up to a time horizon TH, caused by the cumulative emissions of CO_2 from the gas plant (from the combustion of the gas, along with CO_2 from production and transportation of the fuel) plus the cumulative CH_4 emissions associated with the production, processing and distribution of the gas, and the corresponding quantities for the coal plant (coal production also results in CH_4 emissions). Values of the TWP greater than one mean that the gas plant causes more radiative forcing than the coal plant over the first TH years of operation. Values less than one mean the reverse. TWPs are functions of TH, because CH_4 has a much shorter residence time in the atmosphere than CO_2. We make the comparison for Dallas-Fort Worth electricity to avoid the complication of long-distance gas transport and seasonal storage. We consider the fleet conversion TWP, which assumes a continuous stream of emissions starting at the beginning of an analytical time horizon (pulse and service life TWPs are provided in the electronic spreadsheet Zavala et al TWP.xls, tab 'TWP calcs').

We examined a range of heat rates for natural gas electric generating units operating in the 25-county Barnett Shale region or contiguous counties. These units account for about 18% of total 2013 generation from natural gas fired power plants in Texas, and exhibit a similar distribution of efficiencies as the rest of the Texas natural gas power plant fleet. Since no coal plants are located within the region, we considered heat rates typical of the Texas fleet of coal plants. Heat rates were derived from 2013 operating data reported to the EPA Clean Air Markets Program (48). The type of coal used at Texas plants was obtained from EIA form 923 data for 2013 (49).

Table S7. Input for radiative forcing calculations.

Power plant	Emission case	Heat rate (Btu/kWh)	Emissions (kg/mmBtu)			Emissions (kg/mmBtu)		
			Upstream CH_4	Upstream CO_2 ł	Combustion CO_2*	Upstream CH_4	Upstream CO_2 ł	Combustion CO_2*
Average Dallas-Fort Worth gas	base Barnett leakage	8,200	0.285	5.00	53.1	2.34	41.0	435
Best Dallas-Fort Worth gas	base Barnett leakage	6,600	0.285	5.00	53.1	1.88	32.9	349
Worst Dallas-Fort Worth gas	base Barnett leakage	14,300	0.285	5.00	53.1	4.08	71.5	759
Average Texas coal (LIG)	low gassy (LIG)	9,800	0.058	0.00	97.7	0.57	0.00	959
Average Texas coal (SUB)	low gassy (SUB)	9,800	0.011	1.10	97.2	0.11	10.8	954
Average Texas coal (BIT)	high gassy (BIT)	9,800	0.296	0.81	93.3	2.91	7.90	916
Average Texas coal (SUB)	Low gassy (SUB)	8,600	0.011	1.10	97.2	0.1	9.50	839
Average Texas coal (BIT)	high gassy (BIT)	10,900	0.296	0.81	93.3	3.23	8.80	1019

*CO_2 EFs from coal and gas are from EIA, Voluntary Reporting of Greenhouse Gases Program Table 1 (downloaded 2/2011).
ł Upstream CO_2 for coal plants was calculated from Table 1 of Alvarez et al.(1), adjusted for heat rates applicable to scenarios examined and heat content of coal (lignite transport was assumed to be negligible given proximity to power plant).

The CH_4 emission rate associated with Barnett Shale gas used in a natural gas power plant in the Dallas-Fort Worth area was determined by normalizing the emissions from all natural gas infrastructure exclusive of local distribution as reported in this work (55.8 Mg CH_4/h) by the amount of natural gas produced in the Barnett Shale counties delivered into the transmission pipeline system. Following Alvarez et al. (1), we allocated 56% of estimated CH_4 emissions from oil production sites to the natural gas supply based on the share of total energy produced by oil production sites due to natural gas (this is a small correction as oil wells produce less than 10% of total oil and gas CH_4 emissions in the 25-county Barnett region). The amount of gas delivered was estimated to be roughly 16% lower than gross gas production after accounting for the volumes of gas lost due to removal of natural gas liquids and non-hydrocarbon gases, emissions, and consumptive use at production sites, compressor stations, and processing plants.

Table S8 Data used to determine allocation of CH_4 emissions from oil production sites to the natural gas supply. (Source: Lyon et al. (27), SI.)

		Production	Heat Content*	Energy	Percent
Gas production sites	oil	460,000	5.8	2,660,000	1%
	gas	164,290,000	1.1	180,720,000	99%
Oil production sites	oil	1,230,000	5.8	7,140,000	44%
	gas	8,390,000	1.1	9,230,000	56%

*Source: (50, 51).

We make the comparison for Dallas-Fort Worth area electricity to avoid the complication of long-distance gas transport and seasonal storage. To apply the analysis to Barnett gas used in other areas, transmission and storage emissions would have to be added. Here we provide a rough estimate of the likely magnitude of these emissions using existing data, though we note that ongoing work may provide updated estimates (13, 41). The 2015 EPA Greenhouse Gas

Inventory (GHGI) estimates 2013 emissions from the natural gas industry's Transmission and Storage segment to be 2,176 Gg/y. When normalized by the amount of natural gas delivered to consumers in 2013, 32.8 trillion cubic feet (52), the average natural gas leak rate from Transmission and Storage is approximately 0.5% of gas delivered, or about one-third of the 1.6% loss rate we determined for Barnett natural gas.

Upstream CO_2 emissions for natural gas used in power plants include direct fugitive and vented emissions of natural gas stream as well as CO_2 from combustion of natural gas used as fuel along the upstream supply chain (production sites, compressor stations, and processing plants). Fugitive and vented emissions were estimated by scaling national emissions from production and processing in the EPA GHGI by the pro-rata share of national gas produced in the Barnett Shale (7.7%). Upstream combustion emissions were estimated based on the amount of natural gas used as a fuel at production sites; these account for about two-thirds of the total upstream CO_2 emissions.

Table S9. Coal mine CH_4 and CO_2 emissions from combustion used in this work (53, 54).

Coal type	Coal HHV as received (Btu/lb)	Mine CH_4 (scf/ton)	Combustion CO_2 (kg/mmBtu)
Texas lignite (LIG)	6,550	39.9	97.7
PRB- low gassy (SUB)	8,560	10	97.2
PRB - high gassy (SUB)	8,560	40	97.2
Bituminous - high gassy	11,670	360	93.3

We assumed that any CH_4 emissions from the operation of natural gas or coal power plants are insignificant, but this assumption deserves further scrutiny.

Results.

Based on the TWP approach of Alvarez et al. (1), each percent of natural gas lost to the atmosphere prior to end-use combustion adds 33% to the 20-year radiative forcing of the CO_2 produced from complete combustion of the gas without any losses (15% on a 100-year time horizon) (Fig. S6A). These estimates account for the effect of all fuel-cycle CO_2 emissions (i.e., vented, fugitive, and combustion CO_2 emissions from the production, processing, transportation and consumption of natural gas (Table S7)), and assume a scenario where natural gas is consumed continuously, which is appropriate when considering the climate implications of individual and economy-wide energy choices. Using an alternative assumption of a one-time pulse of emissions results in slightly smaller increases in radiative forcing (29% and 10%, on 20-year and 100-year time horizons, respectively).

CH_4 emissions from the production, gathering, processing and delivery of Barnett Shale gas (~1.6% of gas consumed) substantially increase the cumulative radiative forcing from complete combustion of the gas when compared to emissions of CO_2 alone: roughly 50% over a 20-year basis (~20% on a 100-year basis). Reducing CH_4 leakage leads to a proportional decrease in this effect (Fig. S6A).

Our results indicate that, on average, natural gas power plants in the greater Barnett Shale region produce climate benefits on all time frames, compared to coal-fired plants in Texas (Fig. S5B).

The reference case in Fig. S6B assumes thermal efficiencies for the two plants that are typical of the Dallas-Fort Worth area, and the low fugitive CH_4 emissions characteristic of the surface mined, sub-bituminous coal burned in the area's coal plants. We also include cases with the highest possible efficiency coal plant in Texas, and with CH_4 emissions in the Barnett region reduced by 45% and 90%.

However, power plants with heat rates greater than ~10,750 Btu/kWh produce climate damage for some period of time (Fig. S6B). This category includes about half of the generating units in the region, responsible for about 20% of the electric generation (Fig. S6D). These plants are likely to be operated only during peak periods, and are responsible for a relatively small fraction of the electricity from all natural gas power plants; nevertheless, this result suggests both natural gas leakage and natural gas plant efficiency warrant further attention.

All but one of the coal plants in Texas use sub-bituminous coal (from surface mines in the Powder River Basin) or lignite (from mines in Texas in close proximity to the power plant). Both of these coals produce relatively little CH_4 compared to bituminous coal from underground mines (used in only one Texas plant). Our main analysis assumes sub-bituminous coal whose upstream CH_4 is less than 20% the lignite value (Table S9). However, TWP results are reasonably similar whether the coal plant is assumed to use lignite or sub-bituminous coal (Fig. S7A). By contrast, the use of high-gassy bituminous coal significantly changes the TWP result. Fig. S7B indicates that for sufficiently low CH_4 leakage in the natural gas supply, the high CH_4 emissions from gassy coal cause the TWP curve to increase instead of decrease over time. Fig. S7C shows that TWP values can vary by more than a factor of 3 when comparing combinations of the best and worst gas and coal plants.

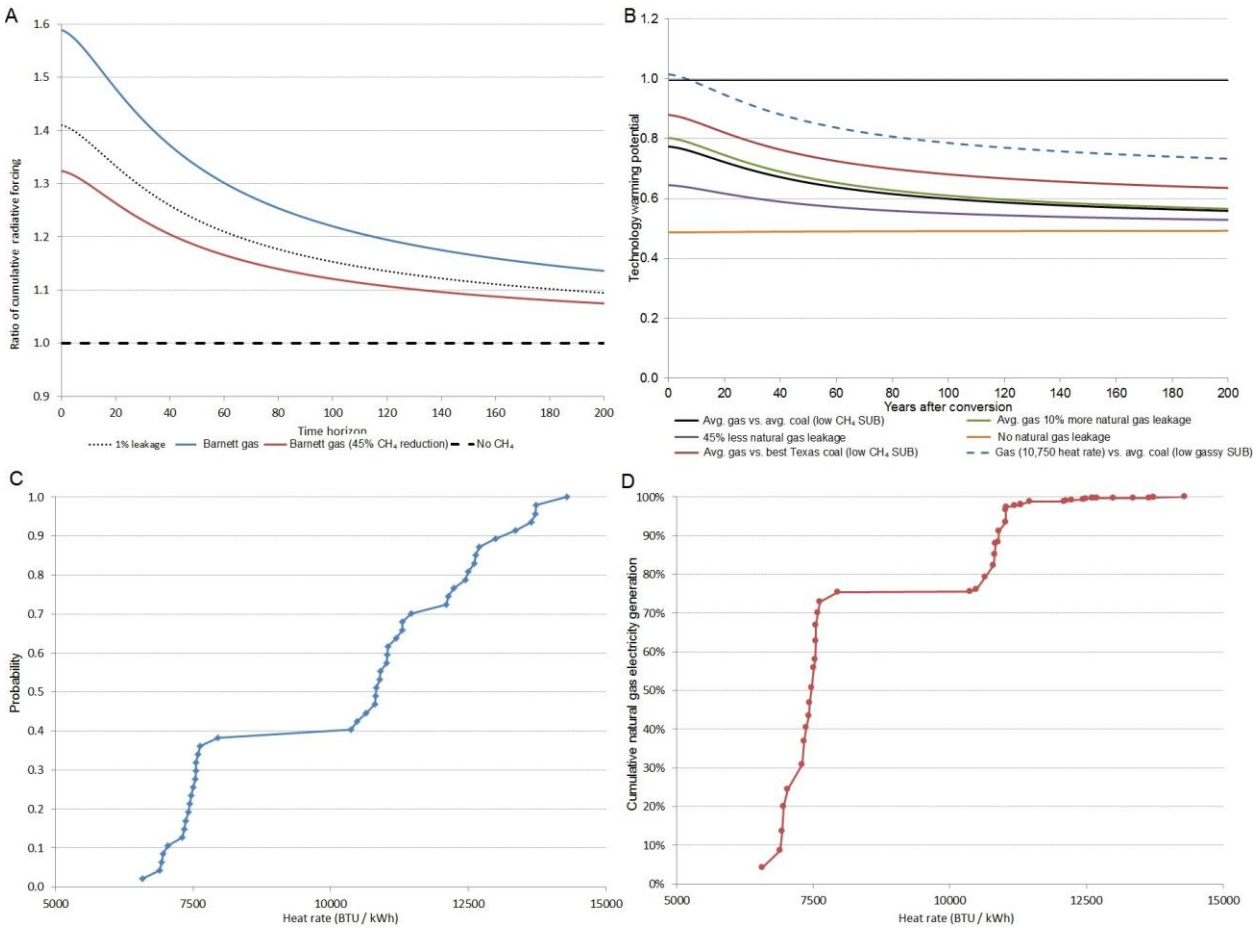

Fig. S6. Effect of CH_4 leakage on fuel-cycle radiative forcing and summary of climate influence of electric power plants fueled with natural gas from the Barnett Shale compared to coal. (A) Generic effect of each percent of natural gas loss on the cumulative radiative forcing of natural gas use (dotted line): comparison is to the forcing due to CO_2 emissions from the natural gas supply chain and combustion alone. The relative increase in cumulative radiative forcing from the continuous operation of an average natural gas fired power plant in the Barnett Shale region due to CH_4 leakage along the supply chain equivalent to 1.6% of gas consumed is greater than the forcing due to CO_2 emissions alone by approximately 50% and 20% over 20 and 100 years, respectively (blue line). Also shown is the effect of a hypothetical 45% reduction in supply chain CH_4 leakage in the Barnett Shale (red line). (B) Technology Warming Potential (TWP) of the average natural gas power plant in the greater Barnett Shale region using locally-sourced gas, relative to the average coal-fired power plant in Texas fueled with low CH_4 sub-bituminous coal (black curve). Other curves show effects of alternative values of CH_4 leakage and power plant heat rates applied to the scenario in black curve: 45% and 100% reduction in upstream natural gas leakage (purple and orange, respectively); 10% increase in natural gas leakage, simulating the top-down estimate of Barnett Shale CH_4 emissions (green); 30% less efficient natural gas power plant than the regional generation-weighted average (10,750 Btu/kWh) (dashed blue); and coal-fired generating unit that represents the most efficient unit in Texas (8,640 Btu/kWh) (red). (C) Distribution of effective heat rates in 2013 for natural gas power plants in the Barnett Shale region. (D) Cumulative electric generation in 2013 from the same population of plants. About half of the generating units in the region, responsible for about 20% of the electric generation, have a heat rate above 10,750 Btu/kWh— the value above which TWP of natural gas plants using Barnett gas is > 1 for some period of time (dashed blue line in panel B).

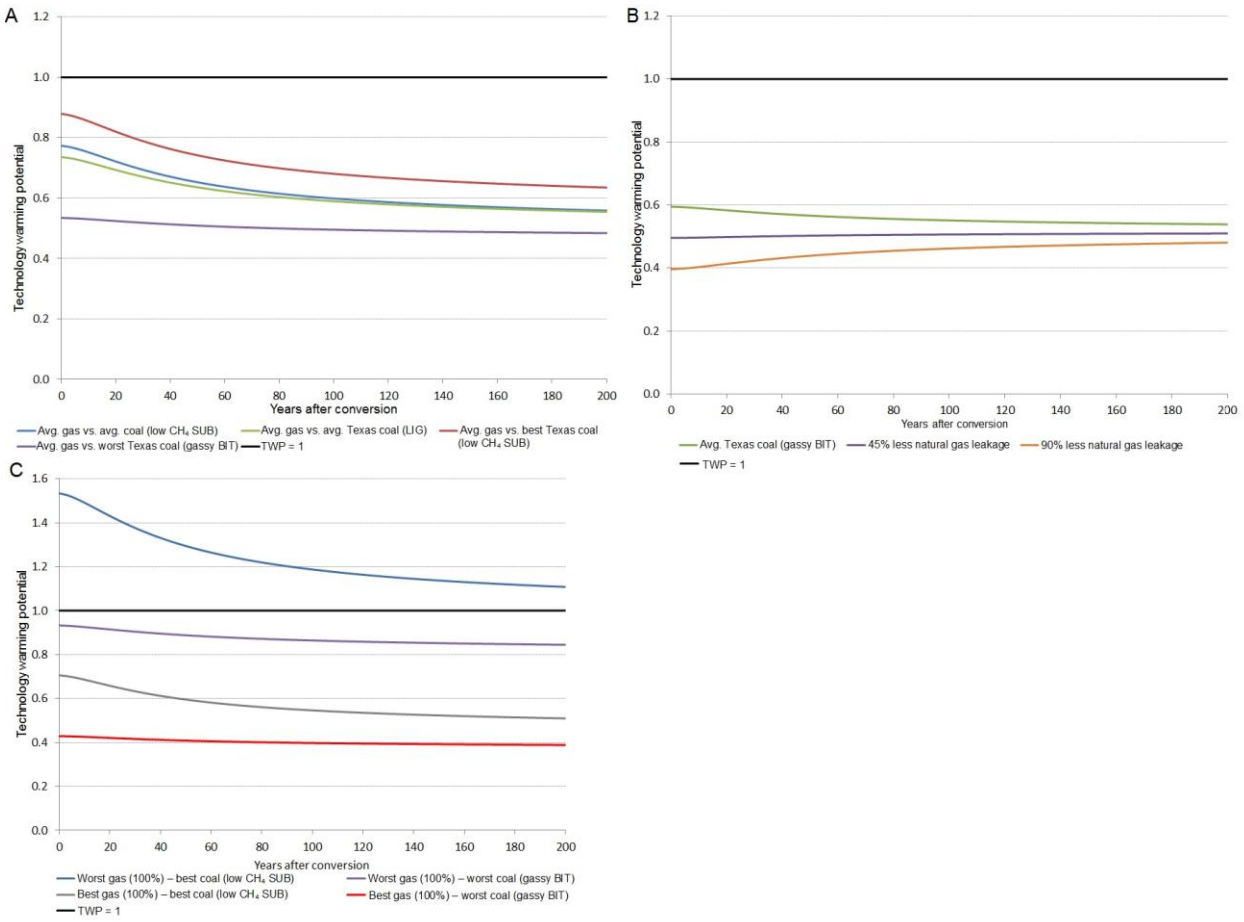

Fig. S7. Sensitivity of Technology Warming Potential (TWP) results to a broad range of alternative assumptions. (A) TWP of average Barnett gas plant relative to coal plants under alternative assumptions about the type of coal and coal plant heat rate: average Texas coal plant using low-CH_4 sub-bituminous coal (blue) and lignite (green); least efficient coal plant using high-CH_4 bituminous coal (purple); and the most efficient coal plant using low-CH_4 sub-bituminous coal (red). (B) TWP of the average Barnett gas plant with alternative natural gas leakage assumptions relative to the average Texas coal plant using high-CH_4 bituminous coal: baseline Barnett natural gas leakage (green); 45% less leakage (purple); and 90% less leakage (orange). Note that for sufficiently low natural gas leakage relative to the upstream coal CH_4 emissions, the shape of the TWP curve can change from monotonically decreasing (green) to monotonically increasing (orange). (C) TWP results comparing best and worst Barnett area natural gas power plants using local gas to the best and worst Texas coal plant scenarios: worst natural gas unit compared to best coal plant using low-CH_4 subbituminous coal (blue); worst natural gas unit compared to the worst coal plant using high-CH_4 bituminous coal; best gas unit compared to best coal unit using low-CH_4 sub-bituminous coal (gray); and best gas unit compared to the worst coal plant using high-CH_4 bituminous coal (red).

S7: Analysis of the uncertainty of top-down and bottom-up estimates.

Uncertainty in the reported top-down mean.

Karion et al. (19) reported the average top-down (TD) CH_4 estimate to be the arithmetic mean of the estimates from the individual flights; with the 95% confidence interval (CI) determined using a bootstrapping method (resampling, with replacement, the estimates from the individual flights). As described in Smith et al. (29), the October 25th flight was considered an outlier (analysis of this flight and estimates including it are shown in S7). Using this approach, for the seven flights considered in this work, the total CH_4 estimate is 71 Mg CH_4/h, with a 95% CI of 59 – 83 Mg CH_4/h (a relative uncertainty of ± 17%) (Table S10).

Table S10 Summary of TD and BU estimates for each of the flights, with 95% confidence intervals (CI). Individual BU estimates are for source regions corresponding to areas sampled by the aircraft on each flight (with 95% CI).

Flight	Total TD CH_4 emissions (Mg CH_4/h) [a]	Total BU CH_4 emissions (this work) (Mg CH_4/h)	Fossil TD CH_4 emissions (Mg CH_4/h)	Fossil BU CH_4 emissions (this work) (Mg CH_4/h)
March 25, 2013	78 (56, 100)	54 (45, 65)	62(44, 80)	44 (36, 55)
March 27, 2013	87 (69, 105)	60 (50, 72)	69(54, 84)	41 (35, 51)
March 30, 2013	78 (47, 109)	83 (70, 98)	62(37, 87)	58 (48, 72)
October 16, 2013	41 (33, 49)	76 (64, 90)	33(26, 39)	52 (44, 65)
October 19, 2013	61 (54, 68)	71 (60, 85)	48(42, 55)	52 (43, 64)
October 20, 2013	88 (53, 123)	80 (68, 96)	70(48, 98)	57 (47, 71)
October 28, 2013	64 (37, 91)	73 (62, 87)	51(29, 73)	52 (43, 64)
Average	71 (59, 83)	71 (64,78)	56 (47, 66)	51 (47, 55)

[a] As reported in Karion et al. (19).

As discussed in the main paper, we used the result reported by Karion et al. (19) (based on Smith et al. (29)) that 79.5% (73.5% – 84%; 95% CI) of Barnett CH_4 emissions comes from fossil sources, to estimate fossil CH_4 emissions from total CH_4 emissions. After propagating the additional uncertainty of the fossil attribution factor (in quadrature) to each of the individual flights and to the average TD estimate, the average TD fossil CH_4 estimate is 56 Mg CH_4/h, with a 95% CI of 47 – 66 Mg CH_4/h.

Uncertainty in the bottom-up mean.

For the bottom-up (BU) estimates, the CI of the CH_4 estimates for each of the source regions sampled by individual flights reflects the uncertainty in the emission factors for source types analyzed with the statistical estimators in this work (production sites, compressor stations, and processing plants (Methods, Table 1, main text) plus the uncertainty in the emission factors for the rest of the emission categories considered in the inventory (as reported in Lyon et al. (27)). Our reported CI for the BU estimates is thus based exclusively on the uncertainty in emission factors. Lyon et al. uses published uncertainty estimates for some sources (e.g., gathering pipelines, landfills) and excludes uncertainty for minor sources (e.g., geologic seepage) lacking reputable uncertainty estimates. These sources account for 7% and 33% of the fossil and total emissions and therefore should have minor impact on the final uncertainty. Additional uncertainty due to activity factors is expected, but was not included in the absence of empirical data; however, we expect this uncertainty should be minimal due to the comprehensive effort to

identify sources from multiple data sources. There may be additional uncertainty from source categories excluded from the BU inventory (e.g., wetlands), but the close agreement of the TD and BU estimates suggests the contribution of these sources is small. Finally, another source of uncertainty in BU estimates is the spatial extent of flight-specific source regions sampled (which determines the total sources aggregated for each individual flight); this uncertainty is also expected to be less than 10 – 20% for each flight (see below).

In accordance with the method for estimating the 95% CI from the TD estimate, a bootstrapping method was applied to the BU estimates from the individual flights. Thus, the total BU CH_4 estimate is 71 Mg CH_4/h, with a 95% CI of 64 – 78 Mg CH_4/h. Similarly, the fossil BU CH_4 estimate is 51 Mg CH_4/h, with a 95% CI of 47 – 55 Mg CH_4/h.

Sensitivity to boundary of flight envelopes.

Day-to-day differences in the BU CH_4 estimates are due almost entirely to changes in the assumed source areas for each flight because emissions from virtually all sources are assumed to be constant (new well completions, accounting for less than 1% of total CH_4, were allowed to vary). Table S11 summarizes the differences between the source regions estimated to be sampled by the mass balance flights. BU estimates were relatively insensitive to modest changes in the assumed upwind and lateral boundaries of TD flight envelopes (Fig. S4).

BU emission estimates are very sensitive to the boundaries of flight envelopes when they are near the 8 core counties; this is the section of the Barnett Shale production region where most of the oil and gas sites and facilities are located (27) (Fig. S3). When one of the boundaries is very close to the 8 core county boundaries, changes can have a significant impact. As illustrated in Fig. S3, the east boundary of the March 25th source region does not cover the core counties entirely, consequently, emission estimates are very sensitive to changes in the location of the east boundary (a 25 km reduction/extension causes a ± 21% change in the BU estimate) (Fig. S8). A similar situation is observed with the west boundary of the March 27th source region, which cuts through Parker and Hood counties (both counties are part of the Barnett 8 core counties). A reduction of 25 km of that west boundary causes a reduction of 16% in the BU estimate.

The significant changes in BU estimates (> 20%) observed for the 50 km reductions (e.g. March 25th, October 20th, and October 28th) imply that the sampled source region may have captured the 8 core county region in its entirety.

With the exception of the cases stated above, where the reductions or expansions affected the overlap with the 8 core county region, expanding or reducing individual boundaries by 25 km or less did not change BU emissions estimates by more than ~10%. This suggests that on most days the mass balance flights effectively sampled the core emissions area (Fig. S8).

Table S11: Summary of source regions estimated to be sampled by mass balance flights and from which bottom-up (BU) inventories were developed (results also shown in Fig. 1 (main text) and Table S10) (with 95% confidence intervals (CI)).

Flight	Total area (km^2)	Average wind direction [a]	Total TD CH_4 emissions (Mg CH_4/h) [a]	Total BU CH_4 emissions (this work) (Mg CH_4/h)
March 25, 2013	24,700	NNW	78 (56, 100)	54 (45, 65)
March 27, 2013	28,700	S	87 (69, 105)	60 (50, 72)
March 30, 2013	50,200	SW	78 (47, 109)	83 (70, 98)
October 16, 2013	40,500	N	41 (33, 49)	76 (64, 90)
October 19, 2013	38,600	N	61 (54, 68)	71 (60, 85)
October 20, 2013	50,300	S	88 (53, 123)	80 (68, 96)
October 25, 2013[b]	63,300	SE	109 (79, 139)	87 (74, 104)
October 28, 2013	48,500	SSE	64 (37, 91)	73 (62, 87)

[a] As reported in Karion et al. (19).
[b] Flight for October 25th was excluded from the analysis; as discussed in Smith et al. (29). See section S7 for more details.

Uncertainty in the comparison of top-down and bottom-up estimates.

Let x_{TDj} and x_{BUj} be the pair of TD and BU estimates, respectively, for each individual flight *j*.

The difference between TD and BU for each *j* can be expressed as:

$$x_{TDj} - x_{BUj} \quad \text{(S6.1)}$$

We treat each difference as an independent observation and estimate the mean and standard error of the total CH_4 estimates. For the case of the fossil CH_4 estimates we need to account for the additional uncertainty of the attribution factor reported by Karion et al. (19) (based on Smith et al. (29)) that 79.5% (73.5% – 84%; 95% CI) of Barnett CH_4 emissions comes from fossil sources. Thus, the mean fossil difference (5.4 Mg CH_4/h with standard deviation of 14 Mg CH_4/h) and the fossil attribution factor (79.5% with a standard deviation of 6%) are treated as independent variables; the variance of the product of those two variables is used to determine the standard error and 95% CI of the mean difference of fossil CH_4 (Table S12).

Table S12. Summary of comparison between top-down (TD) and bottom-up (BU) estimates.

Estimate	Mean difference (Mg CH_4/h)	Standard error (Mg CH_4/h)
Total CH_4	0.10	7.5
Fossil CH_4	5.4	9.1

Thus, for total CH_4, the difference between TD and BU, with a 95% CI, would be 0.1% ± 21% (expressed as a percent of the average TD estimate). In absolute units, the difference (with 95% CI) would be 0.1 ± 15 Mg CH_4/h.

For the case of fossil CH_4, the difference between TD and BU; with 95% CI, would be 9.6% ± 32% (expressed as a percent of the average TD estimate). In absolute units, the difference (with 95% CI) would be 5.4 ± 18 Mg CH_4/h.

Thought experiment: Effect of number of flights on top-down (TD) estimates.

The uncertainty due to the daily variability in TD CH_4 estimates using aircraft mass balance methods can be mitigated by conducting multiple flights. Similar studies published to date have relied on only one or two flights to estimate emissions (2–4). As noted in Karion et al. (19), taking the average of 8 mass balance flights was an important strength of their study. Using a bootstrap approach that resampled (with replacement) the distribution of 8 flight estimates in Karion et al., we estimate a 38% probability that a single flight could produce a result outside the reported 95% CI, with a maximum potential difference of more than 40% from the reported 8-flight mean estimate. See Table S13, which also shows results for subsets of 2–7 flights.

Table S13. Simulation of the effect of the number of flights on the probability that the resulting average falls outside Karion et al.'s (19) 95% confidence interval (CI) for total CH_4 emissions in the Barnett Shale region (76 ± 13 Mg CH_4/h).

Number of flights	Probability that average falls outside of Karion et al.'s 95% CI* (below lower CI, above upper CI)	Maximum relative difference of possible N-flight averages relative to Karion et al.'s reported 8-flight mean value	
		Minimum value	Maximum value
1	38% (25%, 13%)	-46%	44%
2	32% (18%, 14%)	-32%	30%
3	23% (12%, 11%)	-27%	25%
4	18% (9%, 8%)	-19%	19%
5	13% (7%, 6%)	-15%	16%
6	10% (6%, 4%)	-10%	11%
7	7% (4%, 3%)	-6%	7%

*The reported 95% CI is based on the mean of eight estimates, exclusive of any systematic error.

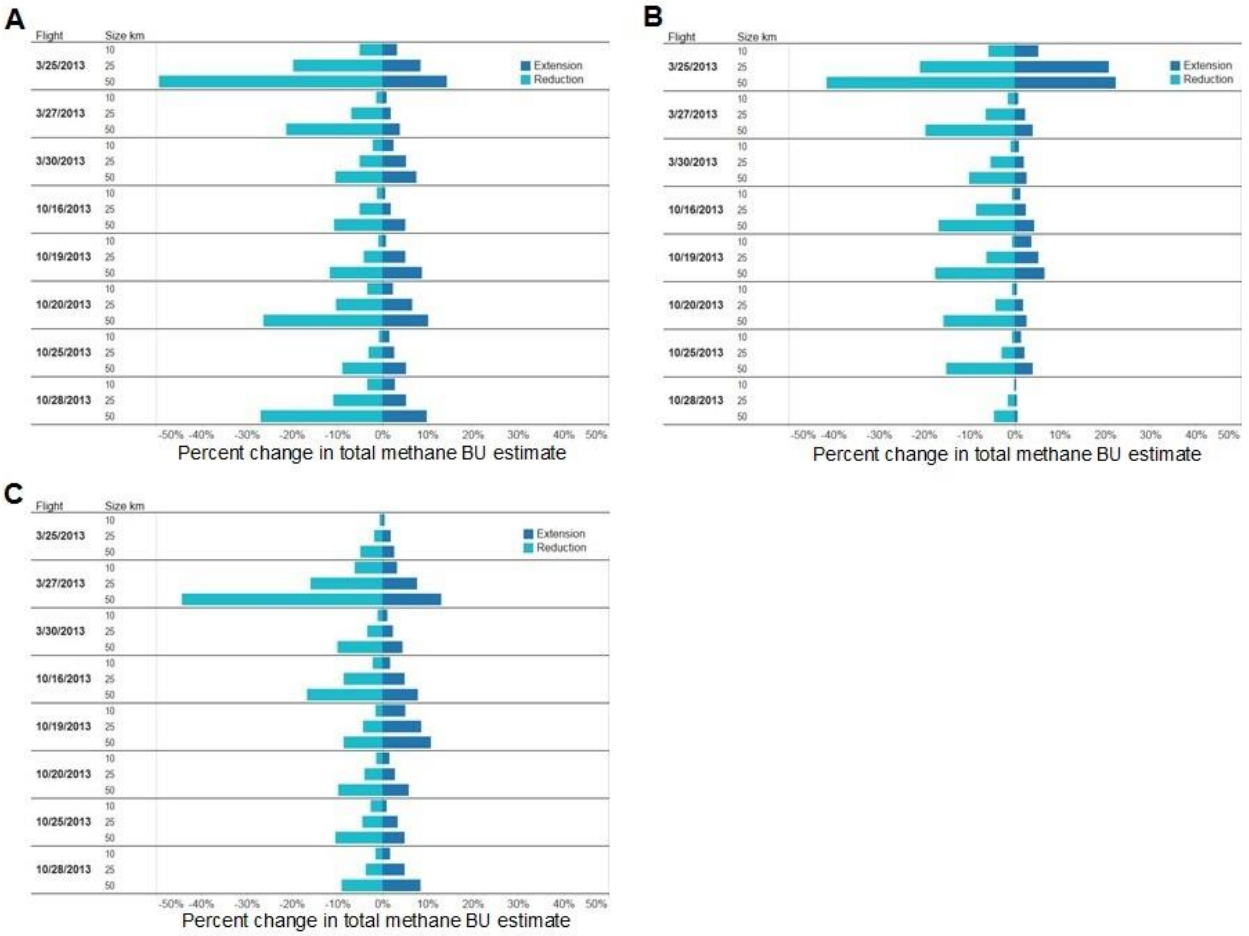

Fig. S8. Sensitivity to boundary of flight envelopes. Relative changes in daily bottom-up (BU) CH_4 emissions due to 10, 25, and 50 km changes to the boundaries of source areas sampled by individual flights (Fig. S3): (A) upwind boundary; (B) east boundary; (C) west boundary. The significant changes in BU estimates (> 20%) observed for the 50 km reductions (e.g., March 25th, October 20th, and October 28th) imply that the sampled source region is not covering the 8 core county region in its entirety. With the exception of the cases stated above, where the reductions or expansions affected the 8 core county region, expanding or reducing individual boundaries by 25 km or less did not change BU emissions estimates by more than ~10%. This suggests that on most days the mass balance flights effectively sampled the core emissions area.

S8: Discussion about the October 25th, 2013 flight as an outlier.

As mentioned in Karion et al. (19), the region of influence for the October 25th flight includes a larger area than the rest of the flights; as shown in Fig. S4 and Table S11, the source region for this flight is roughly 58% bigger than the average area of the rest of the flights. This source region has 14% more gas production and four times more liquids production than the total for the Barnett 8 core counties (19). This produces a top-down (TD) CH_4 total estimate of 109 ± 30 Mg CH_4/h, which is roughly 54% higher than the average of the rest of the flights.

In a similar way, Smith et al. (29) produced an ethane estimate of 16.4 ± 7.0 Mg C_2H_6/h, which is 2.5 times higher than the average of the rest of the flights (see Fig. S2A). Due to the difference between the October 25th flight and the rest of the flights Smith et al. considered this flight as an outlier and excluded it from their analysis. Consequently, the estimates presented in this work's main text also excluded the October 25th flight, leaving seven CH_4 TD estimates and six ethane TD estimates.

Table S14 summarizes the TD and bottom-up (BU) estimates if the October 25th, 2013 flight is included in the analysis.

Table S14. Summary of top down (TD) and bottom-up (BU) estimates when the October 25th, 2013 flight is included (with 95% confidence intervals (CI)). TD independent fossil reflects the independent TD estimate that was determined in section S2; using the ethane estimates from Smith et al. (29) and the fossil CH_4:C_2H_6 ratio (8.7 ± 2.8; mass basis) that was derived from reanalysis of Yacovitch et al. (32).

Estimate	Average (Mg CH_4/h)	95% CI (Mg CH_4/h)
TD total CH_4	76	(63,88)
BU total CH_4	73	(66,80)
TD fossil CH_4	60	(49, 71)
TD independent fossil CH_4	69	(38,101)
BU fossil CH_4	52	(47, 55)

Thus, for total CH_4, the difference between TD and BU, with a 95% confidence interval (CI), would be 3.7% ± 18% (expressed as a percent of the average TD estimate). In absolute units, the difference (with 95% CI) would be 2.8 ± 14 Mg CH_4/h.

For the case of fossil CH_4, the difference between TD and BU, with a 95% CI, would be 13% ± 31% (expressed as a percent of the average TD estimate). In absolute units, the difference (with 95% CI) would be 8.0 ± 19 Mg CH_4/h.

The main difference is observed with the independent TD estimate, which uses an independent CH_4:C_2H_6 ratio (mass basis) of 8.7 ± 2.8 applied to the individual TD ethane estimates reported in (29) (see S3). When the October 25th, 2013 flight is incorporated to obtain the independent TD estimate for fossil CH_4, the estimate would be 69 ± 31 Mg CH_4/h (instead of 57 ± 18 Mg CH_4/h; when October 25th, 2013 flight was excluded).

References.

(1) Alvarez RA, Pacala SW, Winebrake JJ, Chameides WL, Hamburg SP (2012) Greater focus needed on methane leakage from natural gas infrastructure. *Proc. Natl. Acad. Sci. U.S.A.* 109 (17): 6435-6440.

(2) Karion A, et al. (2013) Methane emissions estimate from airborne measurements over a western United States natural gas field. *Geophys. Res. Lett.* 40 (16): 4393-4397.

(3) Pétron G et al. (2014) A new look at methane and non-methane hydrocarbon emissions from oil and natural gas operations in the Colorado Denver-Julesburg Basin. *J. Geophys. Res. Atmos.* 119 (11): 6836-6852.

(4) Peischl J. et al. (2015) Quantifying atmospheric methane emissions from the Haynesville, Fayetteville, and northeastern Marcellus shale gas production regions. *J. Geophys. Res. Atmos.* 120 (5): 2119–2139.

(5) Jeong S, et al. (2013) A multitower measurement network estimate of California's methane emissions. *J. Geophys. Res. Atmos*. 118 (19): 11,339–11,351, doi:10.1002/jgrd.50854.

(6) Miller SM, et al. (2013) Anthropogenic emissions of methane in the United States. *Proc. Natl. Acad. Sci. U.S.A.* 110 (50): 20018-20022.

(7) Schwietzke S, Griffin SWM, Matthews HS, Bruhwiler LMP (2014) Natural gas fugitive emissions rates constrained by global atmospheric methane and ethane. *Environ. Sci. Technol.* 48 (14): 7714–7722.

(8) Caulton DR, et al. (2014) Toward a better understanding and quantification of methane emissions from shale gas development. *Proc. Natl. Acad. Sci. U.S.A.* 111 (17): 6237-6242.

(9) Allen DT, et al. (2013) Measurements of methane emissions at natural gas production sites in the United States. *Proc. Natl. Acad. Sci.* 110 (44): 17768-17773.

(10) Allen DT, et al. (2015) Methane emissions from process equipment at natural gas production sites in the United States: Pneumatic controllers. *Environ. Sci. Technol.* 49 (1): 633-640.

(11) Allen DT, 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.

(12) Mitchell AL, et al. (2015) Measurements of methane emissions from natural gas gathering facilities and processing plants: Measurement results. *Environ. Sci. Technol.* 49 (5): 3219–3227.

(13) Subramanian R, et al. (2015) Methane emissions from natural gas compressor stations in the transmission and storage sector: Measurements and comparisons with the EPA Greenhouse Gas Reporting Program Protocol. *Environ. Sci. Technol*. 49 (5): 3252–3261.

(14) Lamb BK, et al. (2015) Direct measurements show decreasing methane emissions from natural gas local distribution systems in the United States. *Environ. Sci. Technol*. 49 (8): 5161-5169.

(15) Turner, AJ, et al. (2015). Estimating global and North American methane emissions with high spatial resolution using GOSAT satellite data. *Atmos. Chem. Phys. Discuss.* 15(4), 4495-4536.
(16) Townsend-Small A, Tyler SC, Pataki DE, Xu X, Christensen LE (2012) Isotopic measurements of atmospheric methane in Los Angeles, California, USA: Influence of "fugitive" fossil fuel emissions. *J. Geophys. Res.* 117 (D7).
(17) Wennberg PO, et al. (2012) On the sources of methane to the Los Angeles atmosphere. *Environ. Sci. Technol.* 46 (17): 9282-9289.
(18) Peischl J, et al. (2013) Quantifying sources of methane using light alkanes in the Los Angeles basin, California. *J. Geophys. Res. Atmos*. 118 (10): 4974-4990.
(19) Karion A, et al. (2015) Aircraft-based estimate of total methane emissions from the Barnett Shale region. *Environ. Sci. Technol.* 49 (13): 8124-8131.
(20) Petron G, et al. (2012) Hydrocarbon emissions characterization in the Colorado Front Range: A pilot study. *J. Geophys. Res.: Atmos.* 117 (D4).
(21) Kort EA, et al. (2014) Four Corners: The largest US methane anomaly viewed from space. *Geophys. Res. Lett.* 41 (19):6898-6903.
(22) McKain K, et al. (2009) Methane emissions from natural gas infrastructure and use in the urban region of Boston, Massachusetts. *Proc. Natl. Acad. Sci. U.S.A.* 112 (7): 1941-1946.
(23) Katzenstein AS, et al. (2003) Extensive regional atmospheric hydrocarbon pollution in the southwestern United States. *Proc. Natl. Acad. Sci. U.S.A.* 100 (21): 11975-11979.
(24) Wunch D, et al. (2009) Emissions of greenhouse gases from a North American megacity. *Geophys. Res. Lett.* 36 (15).
(25) Hsu Y, et al. (2010) Methane emissions inventory verification in southern California. *Atmos. Environ*. 44 (1): 1-7.
(26) Brandt AR, et al. (2014) Methane leaks from North American natural gas systems. *Science* 343 (6172): 733–735.
(27) Lyon DR, et al. (2015) Constructing a spatially resolved methane emission inventory for the Barnett Shale region. *Environ. Sci. Technol.* 49 (13): 8147-8157.
(28) Harriss R, et al. (2015) Using multi-scale measurements to improve methane emission estimates from oil and gas operations in the Barnett Shale region, Texas. *Environ. Sci. Technol.* 49 (13): 7524-7526.
(29) Smith ML, et al. (2015) Airborne ethane observations in the Barnett shale: Quantification of ethane flux and attribution of methane emissions. *Environ. Sci. Technol.* 49 (13): 8158-8166.
(30) Rella CW, et al. (2015) Measuring emissions from oil and natural gas well pads using the mobile flux plane technique. *Environ. Sci. Technol.* 49 (7): 4742-4748.
(31) Lan X, Talbot R, Laine P, Torres A (2015) Characterizing fugitive methane emissions in the Barnett Shale area using a mobile laboratory. *Environ. Sci. Technol.* 49 (13): 8139-8146.

(32) Yacovitch TI, et al. (2015) Mobile laboratory observations of methane emissions in the Barnett Shale region. *Environ. Sci. Technol.* 49 (13): 7889-7895.
(33) Xu L, Lin X, Amen J, Welding K, McDermitt D (2014) Impact of changes in barometric pressure on landfill methane emission, *Global Biogeochem. Cycles* 28 (7): 679–695.
(34) Czepiel, PW, et al. (2003) The influence of atmospheric pressure on landfill methane emissions. *Waste Manage*. 23 (7): 593–598.
(35) Lavoie TN, et al. (2015) Aircraft-based measurements of point source methane emissions in the Barnett Shale basin. *Environ. Sci. Technol.* 49 (13): 7904-7913.
(36) United States Environmental Protection Agency. Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2012, 2014; http://www.epa.gov/climatechange/ghgemissions/usinventoryreport.html
(37) United States Environmental Protection Agency. Greenhouse Gas Reporting Program, 2014; http://ghgdata.epa.gov/ghgp/main.do
(38) European Commission Joint Research Centre /Netherlands Environmental Assessment Agency. Emission Database for Global Atmospheric Research, release version 4.2, 2013; http://edgar.jrc.ec.europa.eu
(39) Camuzeaux JR, Alvarez RA, Brooks SA, Browne JB, Sterner T. (2015) Influence of Methane Emissions and Vehicle Efficiency on the Climate Implications of Heavy-Duty Natural Gas Trucks. *Environ. Sci. Technol.* 49(11):6402–10.
(40) Tong F, Jaramillo P, Azevedo IML. (2015) Comparison of Life Cycle Greenhouse Gases from Natural Gas Pathways for Medium and Heavy-Duty Vehicles. *Environ. Sci. Technol.* 49(12):7123–33.
(41) Zimmerle DJ, et al. (2015) Methane emissions from the natural gas transmission and storage system in the United States. *Environ. Sci. Technol.* 49(15), 9374-9383.
(42) Zavala-Araiza D, et al. (2015) Toward a functional definition of methane super-emitters: Application to natural gas production sites. *Environ. Sci. Technol.* 49 (13): 8167-8174.
(43) Seinfeld JH, Pandis SN (2012) *Atmospheric chemistry and physics: from air pollution to* climate change. (Wiley, Hoboken, NJ) pp 749-752, 862-866.
(44) Gifford F (1959) Statistical properties of a fluctuating plume dispersion model. *Adv Geophys.* 6: 117-137.
(45) Townsend-Small A, et al. (2015) Integrating source apportionment tracers into a bottom-up inventory of methane emissions in the Barnett Shale hydraulic fracturing region. *Environ. Sci. Technol.* 49 (13), 8175-8182.
(46) Dacey MF, Krumbein WC (1979) Models of breakage and selection for particle size distributions. *J. Int. Ass. Math. Geol.* 11(2): 193-222.
(47) Wilson BG, Adams BJ, Karney BW (1990) Bias in log-transformed frequency distributions. *J. Hydrol.* 118 (1): 19-37.
(48) United States Environmental Protection Agency. Air Markets Program Data, 2015; http://ampd.epa.gov/ampd/QueryToolie.html

(49) United States Energy Information Administration. Form EIA-923 detailed data, 2015; http://www.eia.gov/electricity/data/eia923/
(50) United States Energy Information Administration. Table A4. Approximate heat content of natural gas, 2015; http://www.eia.gov/totalenergy/data/monthly/pdf/sec13_4.pdf
(51) United States Energy Information Administration. Table A2. Approximate heat content of petroleum production, imports, and exports, 2015; http://www.eia.gov/totalenergy/data/monthly/pdf/sec13_2.pdf
(52) United States Energy Information Administration. Natural gas consumption by end use, 2015; http://www.eia.gov/dnav/ng/ng_cons_sum_a_EPG0_vgt_mmcf_a.htm
(53) National Energy Technology Laboratory (2012) QGESS: Specifications for Selected Feedstocks. DOE/NETL-341/011812 (US Department of Energy, Washington, DC). Available at http://www.netl.doe.gov/File%20Library/research/energy%20analysis/publications/QGESSSec1.pdf.
(54) National Energy Technology Laboratory (2011) Process Name: Surface Mine, Gulf Lignite Coal, Operations. Life Cycle Inventory Data Process Documentation File (US Department of Energy, Washington, DC).