Supplementary Materials for

Assessment of Methane Emissions from the U.S. Oil and Gas Supply Chain

Ramón A. Alvarez*, Daniel Zavala-Araiza, David R. Lyon, David T. Allen, Zachary R. Barkley, Adam R. Brandt, Kenneth J. Davis, Scott C. Herndon, Daniel J. Jacob, Anna Karion, Eric A. Kort, Brian K. Lamb, Thomas Lauvaux, Joannes D. Maasakkers, Anthony J. Marchese, Mark Omara, Stephen W. Pacala, Jeff Peischl, Allen L. Robinson, Paul B. Shepson, Colm Sweeney, Amy Townsend-Small, Steven C. Wofsy, Steven P. Hamburg

Correspondence to: ralvarez@edf.org

This PDF file includes:

Other Supplementary Materials for this manuscript include the following:

Database S1 (supporting datasets to understand and assess the paper's conclusions)
Database S2 (shapefiles of top-down area boundaries used in this work)

S1. Materials and Methods

S1.0. Overview of oil and natural gas supply chain

The highly integrated U.S. oil and natural gas (O/NG) industry comprises extensive, dispersed and heterogeneous infrastructure used to extract, process, and transport hydrocarbons. Sharp growth in U.S. O/NG production after 2008 was made possible by exploitation of shale and other low-permeability geologic formations through improved horizontal drilling and hydraulic fracturing techniques (*3*). Production occurs at ~1 million wells across 30 states with many wells co-producing natural gas, oil and other liquid hydrocarbons (*13*). One-quarter of gas producing wells account for 85-90% of US natural gas production; activity data for active O/NG wells relevant to the calculations in Sections S1.2 and S1.9 are shown in Table S4 (*13*). Natural gas infrastructure also includes approximately: 5,000 gathering compressor stations and 400,000 miles of gathering pipelines (*17*); ~700 gas processing plants clustered in producing regions (*17*); an interstate gas pipeline network of ~300,000 miles interspersed with ~1,800 transmission compressor stations; ~400 underground natural gas storage fields; ~80 liquefied natural gas (LNG) storage facilities and LNG import terminals; and 2 million miles of distribution pipelines contained in more than 1,500 local gas distribution systems (*17*, *37*). CH_4 emissions associated with oil production, transportation and refining are dominated by production facilities. Offshore oil and gas production facilities account for 16% and 5% of US production, respectively, (*13*) and have not been the subject of recent research on CH_4 emissions. Table S3 presents emission sources across the O/NG supply chain, their estimated emissions, and identifies sources for which recent emission measurements have been reported.

S1.1. Technical approach.

Our contribution consists of three elements, summarized briefly here and described in detail in Sections S1.2 to S1.4. Sections S1.5 to S1.9 describe additional methods or data used to derive results in the main paper, or uncertainties therein.

1) We developed a bottom-up (BU) estimate of CH_4 emissions from the U.S. oil and natural gas (O/NG) supply chain based largely on recently reported facility-scale measurements specific to each industry segment (Section S1.2). The estimate for natural gas producing sites is derived from datasets of production site emissions in 6 U.S. production areas (reported emissions were based on downwind measurements of CH_4 enhancements that capture all emissions at a facility) (Database S1). We used these datasets to estimate emission probability density functions (pdfs) using a non-linear model that accounts for a weak dependence of site-level emissions on natural gas production levels. We use those pdfs to assign emissions to the populations of gas producing sites in each of the 62 U.S. gas producing basins, and then aggregated those into a national emission total. We also evaluated an alternative approach that did not assume any dependence of emissions on production, with both methods producing similar results (Section S1.9). For gathering stations and processing plants, our BU estimate relies on previously reported emission factors based on national datasets of facility-level measurements. For transmission and storage compressor stations, our estimate is based on Zimmerle et al. (*26*), which relies on a combination of component and facility-level emission measurements. Estimates for abandoned wells are based on well-specific measurements. For other sources representing <15% of supply chain

emissions, our BU estimate relied on reported values in the U.S. Greenhouse Gas Inventory (GHGI) or engineering estimates. Table S1 lists recent measurement studies.

2) We compared results of top-down (TD) studies in nine U.S. production areas to our BU results, spatially aggregated to match TD survey areas, as an independent assessment of our BU methodology (Section S1.3).

3) We constructed a national inventory of production segment emissions using recent datasets of component-based measurements combined with industry-reported emissions and activity data (Section S1.4). This inventory is compared against the GHGI estimate to assess the sensitivity to alternative inventory methodologies and underlying data inputs.

S1.2 National bottom-up estimate of O/NG CH_4 emissions.

a. Natural gas producing sites. Our BU estimate of CH_4 emissions from natural gas producing sites is based on reported site-based measurements at 433 sites in six production areas (Barnett Shale, Fayetteville Shale, Marcellus Shale [Southwest PA/WV], Uintah County, Upper Green River Basin, and Weld County) (*19–21*, *38*) (Fig. S2, Database S1). The 23 counties where these measurements were made account for ~19% of natural gas production, ~5% of oil production, and ~8% of total U.S. wells (*13*). Measurement methods included the mobile flux plane technique (*19*), dual tracer flux approach (*20*), and OTM-33A, an inverse Gaussian method (*21*, *38*). All three methods capture a snapshot of site-level emissions, with reported duration of individual plume captures of ~50 s (*18*), 30 s to a few minutes (*19*) and 15-20 minutes (*21, 38*). Sites were reported to be sampled on a quasi-random basis without advance operator knowledge. All measurement results were used as reported with two exceptions. First, for the Fayetteville dataset (*21*) we employed the mean emission value for four production sites with multiple observations (whereas the published analysis had randomly assigned one of the measurements to such sites). And second, values reported as zero or below the detection limit (0.08 kg/h, 0.036 kg/h and 0.01 kg/h in Rella et al. (*19*), Robertson et al. (*21*), and Omara et al. (*20*), respectively) were treated as censored data points (see below). Such censoring applied to 78 (40%) and 18 (35%) measurements in the Barnett and Fayetteville, respectively.

We estimate emissions probability density functions (pdfs) conditional on production by using a non-linear model with basin-specific intercepts (i.e., fixed effects for basins). We assume our underlying emissions pdfs are lognormal, which is expected in a system where many independent random and multiplicative events can contribute to the occurrence and magnitude of emissions (*18*).

Previous work by Brandt et al. (*22*) has shown that emissions associated with O/NG operations could follow more extreme distributions than lognormal (*22*). However, Brandt et al. focused primarily on component-level emissions estimates; in that work, a goodness-of-fit test was performed on only one site-level dataset (see Brandt et al. (*22*), Table S21 where the Kolmogorov-Smirnov (KS) test was applied to the Rella dataset and supported a lognormal distribution). There is good reason to expect that site-level emissions, being a sum of component-level leaks, will be less extremely distributed than the component-level leaks. We therefore revisit the datasets considered in Brandt et al. (*22*) to test other datasets of site-level

measurements (see SI of Ref. *22* for datafile), as well as evaluate the additional site-level datasets used in this work but not considered in Brandt et al (*22*).

The datasets examined for lognormality include: for production sites, Rella (*19*), Omara (*20*), Robertson (*21*) (four distinct datasets), and Lan (*39*); for gathering stations, Mitchell (*25*) and Lan (*39*); for transmission and storage stations, Subramanian (*40*) (two distinct datasets). All the datasets except Lan (*39*) and Subramanian (*40*) are used directly in this work's facility-level bottom-up calculation. Brandt (*22*) did not consider Omara (*20*) and Robertson (*21*). For each dataset, we first log-transform the reported site-level emissions. We then test if the log-transformed data are normally distributed by using the Lilliefors test and Shapiro-Wilk test for normality (*41*, *42*). The Lilliefors test is similar to a Kolmogorov-Smirnov (KS) test, but is more appropriate than KS when the analyst does not know the parameters of the null distribution. If the analyst wants to test if a dataset comes from a normal distribution with mu and sigma known, then a one-sample KS test is appropriate. If instead one wants to test if the data come from a normal distribution with unknown parameters, the Lilliefors test is preferred. Similarly, the Shapiro-Wilk test is a general test to determine whether one can reject the null hypothesis that data are distributed normally in cases when the underlying distribution parameters (μ and σ) are unknown (*41*). The null hypothesis for the Lilliefors and Shapiro-Wilk tests is that values are drawn from a normal distribution, with critical P-value of 0.05.

Results from both tests applied to all of the datasets used directly in this work indicate that one cannot reject the null hypothesis that the site-level sample data arise from a lognormal population distribution (see summary table in Database S1). For five of seven site-level datasets considered in Brandt, the null hypothesis of log-normality cannot be rejected by the Shapiro-Wilk test (six of seven for the Lilliefors test). (We emphasize that the Rella and Mitchell datasets are common to this work and Brandt et al. (*22*)) While results of these tests are not a guarantee that the data are generated from an underlying lognormal distribution, we conclude that assuming site-level lognormal emissions in this work is a reasonable assumption and that the assumption is not generally contradicted by the work of Brandt et al. (*22*).

Thus, the problem is one dominated by a multitude of stochastic effects rather than one major and deterministic influence as in a smooth regression. Nonetheless, we do not want to bias the analysis because the sampled population of sites oversampled higher production sites compared to the national population (Fig. S3). The measurements of production site emissions plotted against a site's natural gas production exhibit large variability but little systematic trend except at the highest production levels (Fig. S2). We take account of the weak dependence of the mean on production to account for inter-basin differences in site-level distributions of gas production. We also present an alternative method (Section S1.9), without assuming any production dependence (i.e., a unique lognormal for each basin), which produces a similar result.

Following the approach in Zavala-Araiza (*18*) we assume that the emission rate distributions as well as the natural gas production distributions of the site-based measurements are lognormal. Let x_E be the natural logarithm of CH_4 emissions (in kg/h) measured at a production site, and x_P the natural logarithm of total natural gas production in thousand standard cubic feet per year (Mcf/y; 1 standard cubic foot (scf) natural gas = 0.028 m^3) at each production site. We use the systematic samples collected at each basin *j* to estimate the pdf of emissions

conditional on production, $p(x_E|x_P)$, with lognormal parameters μ_j and σ_j, where μ_j is expressed as a non-linear regression of production:

$$\mu_j = a_j + b{x_P}^{\theta_1} + c{x_P}^{\theta_2} \quad (1)$$

We selected a two-term power law to characterize the relationship between emission and production to capture the apparent features of the data in Fig. S2. The first power law term is intended to characterize the relatively constant behavior of the first part of the gas production range, and the second power law term is intended to characterize the observed increase in emissions from the highest producing sites.

The log likelihood function is:

$$l(\mu_j, \sigma_j) = \sum_{j=1}^{J}\left[S_{0_j} ln\, \Phi\left(\frac{{x_E}^*{}_j - \mu_j}{\sigma_j}\right) - S_{r_j} ln\sigma_j - \sum_{i=1}^{S_{rj}} \frac{\left(x_{E_{ij}} - \mu_j\right)^2}{2{\sigma_j}^2}\right] \quad (2)$$

where J is 6, the total number of basins with site-level data; S_{0_j} is the number of samples at or below the detection limit ${x_E}^*{}_j$; S_r is the number of samples above the detection limit; μ_j is given by Equation 1, and $\Phi\left(\frac{{x_E}^*{}_j - \mu_j}{\sigma_j}\right)$ is the cumulative normal.

We estimate the 16 parameters ($a_j, b, c, \theta_1, \theta_2, \sigma_j$) by solving for the values that maximize Eq. 2 and use a direct search algorithm to calculate 95% confidence limits by inverting the Likelihood Ratio Test. Table S5 summarizes parameters that describe $p(x_E|x_P)$.

We discretize 520,000 natural gas producing sites in the U.S. into 600 unique cohorts (where cohorts are distinguished by their reported natural gas production at the level of two significant figures). This significantly reduces computing time. We can use the parameters shown in Table S5 to estimate an emission factor (EF) for each natural gas production cohort where:

$$EF = e^{\mu + \frac{1}{2}\sigma^2} \quad (3)$$

We estimate EF and its 95% interval by using a direct search algorithm applied to equations 1 to 3 and by inverting the Likelihood Ratio Test. For each natural gas production cohort, we end up with six different EF with 95% intervals (based on the six basin-specific sets of parameters) (Database S1). The results of our pdf of emissions conditional on production are shown in Figs. S4, S5 and S6. The mean EF for the national population of natural gas-producing sites determined using the methodology described in this section is 1.5 kg CH_4 h^{-1} $site^{-1}$ (95% confidence interval (CI): 1.1–1.9 kg/h); this EF is lower than the arithmetic mean of the 433 site-based emission measurements [1.9 kg CH_4 h^{-1} $site^{-1}$ (95% CI: 1.3–2.6 kg CH_4 h^{-1} $site^{-1}$)] as well as the EF obtained from the parameters of a lognormal fit to the distribution of the 433 measurements using Eq. 3 [2.4 kg CH_4 h^{-1} $site^{-1}$ (95% CI: 1.4–2.9 kg CH_4 h^{-1} $site^{-1}$), where μ = -1.0 and σ = 1.9]. These comparisons show the effect of incorporating the non-linear relationship between emissions and production to scale reported measurements from the sampled population of sites to the national population of sites, because the latter is skewed toward lower production levels than the sampled population (Fig. S3). The estimated mean EF for the population of

Barnett Shale production sites using this methodology (1.6 kg CH_4 h^{-1} $site^{-1}$ [95% CI: 1.5–1.8]) is consistent with the value reported in previous work that was based exclusively on measurements made in the Barnett Shale (1.8 kg CH_4 h^{-1} $site^{-1}$ [95% CI: 1.3–2.5]) (*18*).

b. National scale-up of emissions from routine operations at natural gas producing sites. For each of the 62 active U.S. O/NG production basins, and for each of the 600 unique natural gas producing cohorts in the national population of gas producing sites (at 2 significant figures), we randomly sample one of the six basin EFs and draw a random mean from its distribution (fitting the EFs with 95% intervals to a Gaussian distribution) and apply it to all sites within the basin having that amount of gas production. We repeat this process 10,000 times. Total emissions from U.S. sites with non-zero natural gas production are 6,700 Gg CH_4/y (95% CI: 5,100–8,600 Gg/y).

Approximately 10% of U.S. gas producing sites have production levels that fall outside of the range of sampled sites. The minimum gas production in the reported measurement datasets is 0.68 Mcf/d. To estimate emissions for the 9% of national production sites with gas production below 0.68 Mcf/d, we sampled the basin EFs for sites at 0.68 Mcf/d; this yielded a mean EF of 0.92 kg CH_4 h^{-1} $site^{-1}$ (95% CI: 0.60–1.3 kg CH_4 h^{-1} $site^{-1}$). Total emissions for these sites are thus estimated to be 370 Gg/y (5.5% of national emissions from gas producing sites); this contribution is included in our national estimate of total emissions from natural gas producing sites (6,700 Gg/y, see above). This emission estimate for low production sites is robust to alternative EFs: assigning an emission factor corresponding to sites with ~10 times greater gas production (6 Mcf/d vs. 0.68 Mcf/d) changes emissions from these sites by ~3%.

There are 33 sites (<0.01% of the national population) above the maximum gas production of sampled sites. Emissions from these sites are 33 Gg/y using the regression parameters in Table S5 (0.5% of national production emissions). Changing the EF from these very high-production sites doesn't significantly affect our overall results. For example, if we used the EF corresponding to the gas production of the highest-producing site in the measurement dataset (instead of the regression), total emissions would be 15 Gg/y, reducing our national estimate by <0.3%.

The distribution of national well site activity data (well counts, site counts, and O/NG production) and estimated emissions are summarized in Table S4. The population of gas producing sites (91%) and their emissions (89%) are dominated by sites producing between ~1–5,000 Mcf/d, with two-thirds of total emissions coming from sites producing 10–5,000 Mcf/d. These two production cohorts substantially overlap with the distribution of sampled production sites in the 6 datasets used in this work.

c. Other production sources. Several emission sources associated with O/NG production are not captured in the site-level measurements used in this work. Their emissions were estimated as follows, and summarized in Table S6.

Abandoned O/NG well emissions, which are not included in the 2017 EPA GHGI (*17*) but are proposed to be added in the draft of the 2018 GHGI (*43*), were estimated from activity data in Drillinginfo (*13*), a proprietary database compiling well-level O/NG production data from

state agencies, and EFs from recent measurement studies. We classified wells as abandoned if their last production date was before January 2015 and plugged or unplugged based on their well status code. For our central estimate EFs, we used the national EFs for plugged and unplugged wells from Townsend-Small et al. (*44*). For our lower bound EFs, we used 50% of the Townsend-Small et al. EFs. For our upper bound EFs, we used the Kang et al. (*45*) EFs specific to Pennsylvania wells. EPA estimates that there are 1.2 million abandoned wells not included in the Drillinginfo database. For our upper bound national estimate, we added these wells to our Drillinginfo-based activity data, assuming they are unplugged since effective plugging was uncommon in early wells.

Hydraulically fractured well completion emissions were based on Drillinginfo activity data on the number of gas producing wells with either a completion date or first production date in 2015 (*13*) and basin-level EFs derived from US EPA Greenhouse Gas Reporting Program (GHGRP) (*46*) data for hydraulically fractured completions and workovers. Well workover emissions were estimated from gas producing well counts and, the national ratio of GHGRP reported workover events to Drillinginfo gas producing well counts for reporters (0.079), and basin-level EFs derived from GHGRP data for non-hydraulically fractured completions and workovers. Well testing emissions, which account for less than 0.1% of total emissions, were estimated from GHGRP data as described in Section S1.4. Emissions from offshore O/NG production are based on the reported value in the US GHGI (*17*).

We augmented our basin-level BU estimates of production site emissions in the Fayetteville and Bakken (for TD comparisons) as well as the national annual BU estimate to account for additional emissions that were not sampled in the datasets of site-level measurements (manual unloadings and gas flaring in the Fayetteville and Bakken, respectively). In the Fayetteville BU estimate we included 10 ± 2 Mg/h from manual liquids unloadings as reported by Schwietzke et al. during two TD flights (*47*); in the annual BU estimate for the Fayetteville we include 20 Gg/y for liquids unloading as determined from US EPA GHGRP data. (The hourly emission rate from manual unloadings observed on the TD flights is much higher than the hourly rate derived from the total annual emissions estimated in this work, because those emissions are concentrated in daytime hours (Section S1.6).) For the Bakken, we augmented production site emissions by 5.9 Mg CH_4/h to account for the high level of associated gas that was flared at the time of TD sampling (27% of gas produced in May 2014 (*13*)), based on a study of incomplete combustion in 37 unique Bakken flares (*48*). The flare study (*48*) estimated that 21 ± 4% of the CH_4 emissions reported in the TD study of Peischl et al. (*49*) were the result of incomplete combustion in flares. For the 2015 annual inventory estimate of incomplete combustion in the Bakken, we reduced the 2014 estimate in (*49*) by 20% to account for lower flaring in 2015 (42 Gg CH_4/y). We included these idiosyncratic emissions observed in two basins in our estimates in order to enable the most direct comparison of BU and TD emissions in those two basins. Such idiosyncratic behaviors may occur in other basins; however, in the absence of observations we make no adjustments in any other basin. The effect on annual national emissions is <0.1%.

We estimated emission from the approximately 200,000 oil wells with zero reported gas production (oil-only wells, responsible for 7% of U.S. oil production) by assuming that these wells co-produce some amount of associated gas even if no marketed gas production is reported. Oil-only wells therefore emit CH_4 due to small quantities of dissolved gas that are vented when

produced oil is brought to atmospheric pressure. We assume that the gas production of oil-only wells can be estimated from their reported oil production and the gas-to-oil ratio (GOR) of oil wells with low reported gas production. Our logic is that these two groups of wells likely have similar characteristics except for the availability of infrastructure for capturing gas production. GOR is sensitive to the selected maximum gas production of co-producing oil wells. As gas production increases, the weighted average GOR increases, which is likely due to a greater number of wells purposely drilled for producing both oil and gas. Our GOR selection is based on co-producing wells with gas production <1000 scf/day. We believe that these oil wells, which account for the lowest 15th percentile in gas production, are reasonably representative of oil wells with zero reported gas production, but this assumption deserves further scrutiny. The weighted average GOR of 31,291 gas-producing oil wells with reported gas production <1,000 scf natural gas day^{-1} is 220 scf gas bbl^{-1} (1 barrel (bbl) oil = 0.16 m^3; corresponding GOR values for wells producing <10,000, <100, and <10 scf/day are 900, 21, and 2.7 scf/bbl, respectively (*13*)). Assuming a GOR of 220 scf/bbl and a CH_4 content of 50% in the vented gas, we derive an emission factor of 2.1 kg CH_4 bbl^{-1} for the co-produced gas from oil-only wells. We assume a ±50% uncertainty in this estimate due to the uncertainty in GOR and CH_4 content, although we acknowledge this is only a zeroth-order assumption in the absence of any emission measurements or data on GOR and gas composition specific to oil-only wells. In reality, the uncertainty is unknown and also warrants further research. We thus estimate national emissions from oil-only wells of 430 Gg/y (95% CI 210-640 Gg/y), or average per-well emissions of 0.25 kg well^{-1} h^{-1} (approximately one-fourth the value of the lowest gas producing well sites in our analysis above). Oil production of the co-producing wells with very low gas production is about half of that for oil-only wells (1.6 vs 2.9 bbl d^{-1}, respectively); it is unclear what effect this difference in oil production would have on our results. Based on our results, emissions from sites with very low gas production (370 Gg/y from sites producing <0.68 mcf/d) and oil-only wells (430 Gg/y) contribute only ~10% of national O/NG production emissions, but both deserve further attention.

d. Natural gas gathering. Gathering station emissions were estimated from Drillinginfo gas production (*13*) and state-specific emission rates reported in Marchese et al. (*29*). Zavala-Araiza et al. (*18*) analyzed the underlying Mitchell et al. dataset (*25*) using a lognormal fit to calculate a site-based emission factor of 59.6 (45.9 – 77.3) kg CH_4 h^{-1} site^{-1}; this lognormal fit accounts for the effect of high-emitting facilities with emissions above the sampled range reported by Mitchell et al. In sum, we adjusted the Marchese et al. central estimate loss rates by the ratio of the Zavala-Araiza et al. and Mitchell et al. EFs (59.6/54) to better account for heavy-tail emissions. Upper bound loss rates were adjusted upward by 17% to account for the reported potential measurement bias due to incomplete plume capture in the Mitchell et al dataset. Gathering station blowdown emissions were estimated from operational station emissions by multiplying by the ratio of estimated U.S. blowdown to station emissions from Marchese et al., 0.10 (0.003 – 0.42). Gathering pipeline emissions were estimated with the 2017 EPA GHGI activity and emission factors for gathering pipeline leaks (*17*). In contrast to EPA's approach of estimating gathering pipeline emissions from gas wells only, we apply the factors to both gas wells and gas-producing oil wells.

e. Other segments. Natural gas processing, transmission and storage (T&S), and local distribution emissions were estimated from recent measurement and modeling studies (Marchese

et al. (*29*), Zimmerle et al. (*26*), Lamb et al. (*50*)). For processing plants, we start with a reported value of normal operation emissions from Marchese et al. of 506 Gg (+10.8%/-10.3%), which was based on 16 site-level tracer flux correlation measurements from Mitchell et al. (*25*). As described above for gathering stations, we account for under-sampling of the heavy-tail by adjusting the Marchese et al. reported central estimate by the ratio of the log-normal fit EF from Zavala-Araiza et al. (*18*) [173 (104 – 285) kg CH_4 h^{-1} $site^{-1}$] to the Mitchell et al. EF (141 kg CH_4 h^{-1} $site^{-1}$). We also adjust emissions by the 2015:2012 ratio of processing plants from the 2017 GHGI (667/606). For processing plant blowdowns, we use the EPA 2017 GHGI estimate of 35.5 Gg CH_4 (*17*). Loss rate uncertainty was estimated by propagating the uncertainty from Marchese et al.'s reported national emission estimates and our fat-tail adjustment. Upper bound loss rates were adjusted upward by 15% to account for potential measurement bias due to incomplete plume capture. For T&S, we start with the 2017 GHGI estimate of 1,349 Gg, which includes 1,060 Gg for T&S station emissions, which EPA calculated by adjusting the Zimmerle et al. estimate of 2012 T&S station emissions (exclusive of uncategorized/superemitter emissions) for the 2015 station count (*17*, *26*). We increase the GHGI estimate by 200 Mg $station^{-1}$ y^{-1} (437 Gg/y) to account for uncategorized/super-emitter emissions that are estimated in Zimmerle et al. but excluded from the 2017 GHGI. For T&S uncertainty, we use the values reported in Zimmerle et al. for their national estimate (+29.7%/-18.8%). The 2017 EPA GHGI estimate for T&S of 1,349 Gg CH_4 also includes emissions from LNG facilities (70 Gg) and transmission pipelines (220 Gg), (*17*) both of which were used as is in this work. For local distribution, we used the 2017 EPA GHGI estimate of 444 Gg, which adjusted Lamb et al. emissions by 2015 activity data for pipelines and metering and regulating stations. We used the Lamb et al. upper bound uncertainty (+117%) and assumed a lower bound uncertainty of -50% in the absence of a reported value in the original work. We believe the GHGI estimate of local distribution emissions is likely to be biased low and requires greater attention (Section S1.5). For oil refining and transportation, we use the 2017 EPA GHGI estimate of 34 Gg CH_4 (+147%/-24%) (*17*).

S1.3 Comparison of this work's bottom-up emissions estimates to top-down results.

We spatially aggregated our bottom-up (BU) emission estimates to enable comparisons to top-down (TD) results in each of the nine source regions sampled by airborne-based teams (Table S2; Fig. S1). Definitions for TD and BU follow those in main text: TD studies quantify ambient methane enhancements using aircraft, satellites or tower networks and infer aggregate emissions from all contributing sources across large geographies; BU studies generate regional, state, or national emission estimates by aggregating and extrapolating measured emissions from individual pieces of equipment, operations, or facilities, using measurements made directly at the emission point or, in the case of facilities, directly downwind. The boundaries of regions sampled in TD studies were determined from information in the original papers and/or in consultation with the corresponding authors (boundaries are provided as shapefiles in Database S2). For production sites, we use spatially explicit activity data for all sites inside the sampled source region. For all production sites with a given gas production cohort in each basin, we apply a mean value drawn randomly from the EF distribution (fitting the EFs with 95% intervals to a Gaussian distribution) resulting from Eqs. 1 to 3. For the four TD basins where production site-based measurements were also available, we directly apply the basin-specific EF generated for each different gas production cohort. For the other five TD basins, we randomly sample one of the six EFs. For other sectors (e.g., gathering, processing, transmission and storage) we sample a

random mean of total emissions from the distribution derived using the sector-specific estimation methods above (assumed to be normal). We sum up all emissions and repeat the process 10,000 times, combining emissions from the 9 basins. Fig. 1B (main text) compares the results from this Monte Carlo iteration process for the bottom-up estimate and the top-down results.

The Empirical BU estimates in Fig. 1 were adjusted to account for independently reported source contributions that would have been observed by TD measurements and affected their reported estimate of O/NG emissions. As described above, we include independently reported contributions of 10 Mg/h for manual liquids unloading in the Fayetteville and 5.9 Mg/h for incomplete combustion in flares in the Bakken that were not captured in the production site-level measurement datasets. In West Arkoma, we added 5.1±1.7 Mg/h to account for an unidentified source in Fort Smith (*51*). Even though the source is not believed to be O/NG-related, we add it to the empirical BU estimate because the original paper did not adjust their absolute O/NG emission estimate to account for it (they did, however, expand the lower confidence interval of the area's production normalized emission rate). Finally, in the San Juan, we accounted for geologic seepage (a source not included in our bottom up estimate) which was reported to contribute as much as ~20% of observed TD emissions in the region. Emissions from geologic seeps measured during 2007-2015 exhibit significant interannual variability with a reported range of 0.02-0.12 Tg CH_4/y, or 8±6 Mg CH_4/h (*52*), which is the value we use in this work.

To aggregate emissions within the spatial domain covered by each top-down flight (Fig. S1) we created shapefiles for each domain that characterized the activity data specific to the year of each flight using Drillinginfo (*13*) data (i.e., number of well sites, site production, etc.). Production and gathering emissions were adjusted by the ratio of activity data from the portion of the county covered by the top-down flights during the year of the flights to 2015 total county activity data to better align the spatiotemporal domains. The number of hydraulically fractured well completions during TD flights was estimated by calculating the average number of new completions per day during the flight month based on Drillinginfo activity data.

Maasakkers et al. 2016 developed a 0.1º x 0.1º gridded inventory of 2012 CH_4 emissions from the EPA 2016 GHGI by disaggregating national emissions proportional to relevant activity data (*53*). We used the Maasakkers et al. emission estimates to determine emissions in the TD spatial domains by spatially joining our domain-specific shapefiles with their gridded inventory using ArcGIS. For processing and transmission and storage, we estimated emissions in the TD areas by adjusting the Maasakkers et al. emissions by the ratio of our national estimates (as described above) and 2012 national emission estimates for those sources (as reported in the 2016 GHGI). Smith et al 2017 did not find a statistically significant difference between their 4-day mean CH_4 flux in the San Juan basin of 0.54±0.20 Tg y^{-1} and the bottom-up estimate based on the gridded 2012 EPA Greenhouse Gas Inventory of Maasakkers et al (*52*). The inventory estimate encompassed a larger spatial domain than was covered by the flight envelopes (rectangle bounded by 36.0º N, -109.0º W and 37.5º N, -106.5º W). In this work, we defined the source area of the San Juan basin that was sampled by aircraft mass balance flights to match the area covered by the largest flight envelope (rectangle bounded by 36.01º N, -108.6º W and 37.5º N, -106.8º W). Because of large non-O/NG sources located between the boundaries of these two domains, we find that the Maasakkers et al inventory would predict total emissions in the smaller domain that are 25% lower than the original estimate (0.30 vs. 0.40 Tg y^{-1}, respectively).

The comparison of our mean BU estimates to TD results in the four basins where production site measurements were also made is shown in Fig. S8. The comparisons for all nine TD basins are shown in Fig. 1 and Table S7.

The TD estimates of O/NG CH_4 emissions for discrete geographies can be used for validation of our BU methodology (Fig. 1, Fig. S8, and Table S7). By contrast, the national production normalized BU emission rate reported in the main text (2.3% (+0.4/-0.3%)) is not directly comparable to the mean production normalized emission rate from the results of TD studies in nine O/NG production areas (1.8% ± 0.5%) (Table S2). The reason they are not directly comparable is because TD studies, which sampled active O/NG production areas, under-sample the long-distance transmission and local distribution of natural gas. Transmission and local distribution infrastructure contribute only a small fraction of emissions in most production areas surveyed in TD studies because they are widely distributed throughout the country and in urbanized areas, respectively. If we excluded emissions from transmission and local distribution from our national BU estimate, we obtain a national production normalized emission rate of ~2.0% ± 0.3% – a metric that would more closely represent emissions from O/NG production areas. It is worth noting that although absolute emissions from TD studies are 11% larger than our corresponding BU emissions (Fig. 1, a 9-basin total), the production normalized emission rate from the 9 TD studies (1.8% ± 0.5%) is ~10% smaller than our national BU rate even after adjusting for transmission and distribution (~2.0% ± 0.3%). (We note that neither of the two TD-BU differences is statistically significant.) The remaining difference is likely due to the emphasis that TD studies placed on surveying high-production areas like Northeast Pennsylvania (with lower than average production normalized emission rates) relative to low-production areas like West Arkoma with higher normalized emission rates (Table S2). While the 9 sampled TD basins are skewed toward higher producing areas relative to the national population (Fig. S3), there is no indication that the differences indicate any methodological bias (Section S1.8). As discussed in Section S1.2, our emission estimation methodology accounts for the differences in gas production levels at individual sites and, by extension, across basins with differing production characteristics (see also Fig. S7, where normalized production site emissions are shown to increase as site gas production decreases).

S1.4 Alternative inventory estimate incorporating component-based measurements

We used recent component-based measurement data (Table S1) to develop an alternative emission inventory of oil and gas emissions that is otherwise analogous to US EPA GHGI (*17*). The estimates from our alternative inventory method are reported by source in Table S3, which includes a comparison to GHGI and our site-based estimates (Sections S1.2 and S1.3). Table S8 summarizes the emission factors and sources of activity and emissions data used to develop each component/source estimate.

Our approach estimates emissions for oil and natural gas (O/NG) in individual counties by integrating results of recent measurement studies with other data sources including Drillinginfo (DI) (*13*), US EPA GHGRP (*46*), and US EPA GHGI (*17*). GHGRP is a mandatory reporting program for U.S. facilities with annual greenhouse gas emissions ≥25,000 metric tons (MT) carbon dioxide equivalents. Onshore production facilities, which are defined as a company's

aggregate well pad assets in a geologic basin (following the classification of the American Association of Petroleum Geologists), report under Subpart W (CFR Title 40 Part 98 Subpart W). Reporting facilities are required to provide source-specific activity data and emission estimates based on prescribed calculation methodologies that include emission factors, engineering equations, and in limited instances direct measurements. Data are reported at the sub-basin level (county/reservoir type) for associated gas venting and flaring, completions and workovers, hydrocarbon tanks, and liquids unloading; data are reported at the basin-level for dehydrators, flares, reciprocating compressors, centrifugal compressors, well testing, combustion exhaust, pneumatic controllers, pneumatic pumps, and equipment leaks. We downloaded GHGRP data for emissions in 2015 from the U.S. EPA Envirofacts website on October 25, 2016 (https://www3.epa.gov/enviro/) (*46*).

Drillinginfo (*13*) was used to determine 2015 production volumes and actively producing well counts for all operators including those that do not report to the GHGRP. Well-level data were aggregated to the company/basin level to create groups equivalent to GHGRP facilities. Although GHGRP facilities began reporting well counts, gas production, and oil production in the 2015 reporting year, there are several advantages to using DI data in our model. First, the GHGRP activity data has extensive quality assurance limitations such as incorrectly reported units. Second, well-level DI data was used to separate well counts and production by well type, which is not possible with the aggregated GHGRP data. We defined three well categories based on the gas-to-oil ratio of 2015 production: gas wells (GOR >100,000 scf bbl^{-1}), gas producing oil wells (0<GOR<100,000 scf bbl^{-1}), and oil-only wells (zero reported gas production). Finally, DI data can be used consistently for both reporters and non-reporters. We were able to assign DI data to 507 of 524 GHGRP facilities by matching operator names. The remaining GHGRP facilities were excluded from our analysis to avoid double counting their activity data with non-reporters. Our approach includes some uncertainty due to potential changes in well pad ownership between the 2015 reporting year and download date of DI data (January 2017). Starting in the 2016 GHGRP reporting year, onshore production facilities are required to report well API numbers, which will allow future analyses to more accurately match the GHGRP and DI data.

Measurement data from recent studies were used to estimate emissions from equipment leaks (*54*), pneumatic controllers (*55*), pneumatic pumps (*54*), abandoned wells (*44*, *45*), gathering stations (*25*, *29*), processing plants (*25*, *29*), transmission and storage stations (*26*, *40*), and local distribution (*50*). We did not incorporate liquids unloading data from Allen et al 2014b (*56*) because the GHGRP provided more detailed data on event counts and emission rates; the Allen et al estimate of 2012 national emissions from liquids unloading was within a few percent of the GHGRP estimate.

For most O/NG onshore production sources (associated gas venting, associated gas flaring, dehydrators, flares, liquids unloading, reciprocating compressors, hydrocarbon tanks, well testing), GHGRP emissions and/or activity data were used to estimate emissions from both reporters and non-reporters. Basin-level, source-specific emission factors (EFs) with 95% confidence intervals were developed from Monte Carlo simulations of GHGRP reporter data. First, we developed linear models for each emission source using regressions of GHGRP-reported emissions and DI activity data parameters (well count, gas production, and oil

production from gas wells, gas producing oil wells, oil only producing wells, and total wells). Second, we calculated multiple sets of EFs by normalizing reported emissions to each activity data parameter with a statistically significant linear model ($\alpha = 0.05$). For each combination of basin, source, and activity data parameter, we generated 1,000 datasets from reporter EFs. The sample size of each dataset was equal to the number of reporters in the basin, or one if there were no reporters. Reporter EFs were weighted-random-sampled from a mixture of basin-specific and national data. The fraction of reporter EFs sampled from the same basin was equal to the total reporters' fraction of the basin's activity data; EFs were weighted by individual reporter's activity data fraction in the basin. The remaining fraction were selected from the full national dataset of reporter EFs weighted by individual reporter's fraction of national activity data. For each activity-data-parameter-based EF, central estimates were calculated as the mean of the 1,000 datasets; lower and upper bound EFs were defined as the 2.5 and 97.5 percentile, respectively. We estimated county-level, non-reporter emissions by multiplying these EFs by the relevant non-reporter activity data parameter. For sources reported at the basin level, we estimated county-level reporter emissions by disaggregating reported emissions proportional to the same activity data parameter. To account for the uncertainty of extrapolating emissions from different parameters, a final central estimate was calculated by weighting individual central estimates by the inverse relative confidence interval of the linear model slope. Final 2.5% and 97.5% confidence intervals were calculated as the minimum and maximum lower bound and upper bound, respectively, of individual estimates.

Eight production sources were estimated with variations on the GHGRP approach. For centrifugal compressors, the emission estimate is based on a combination of all activity data parameters because no linear model yielded a statistically significant relationship. For hydrocarbon tanks, we excluded reported emissions from stuck separator dump valves to limit our estimates to operational, tank flashing emissions. For CH_4 in combustion exhaust, we used GHGRP CO_2 emissions instead of the reported CH_4 emissions because the GHGRP natural gas combustion EF of 1 g CH_4 $MMBTU^{-1}$ is up to 500X too low for O&G combustion sources (*52*); we recalculated CH_4 emissions from reported CO_2 emissions using the U.S. EPA AP-42 EF CH_4:CO_2 ratio for 4-stroke, rich-burn natural gas reciprocating compressors (*57*). For pneumatic pumps, we combined GHGRP pump counts with an EF, 1.9 (1.2 – 2.9) MT CH_4 yr^{-1}, derived by bootstrapping 62 chemical injection pump measurements from Allen et al. 2013 (*54*).

For pneumatic controllers, we combined GHGRP type-specific controller counts (high-bleed, low-bleed, and intermittent-bleed) with EFs based on a bootstrap analysis of 377 pneumatic controller measurements from Allen et al. 2014 (*55*): 2.0 (1.5 – 2.5), 0.42 (0.56 – 0.70), 0.15 (0.09 – 0.23), and 7.3 (5.5 – 9.4) MT CH_4 yr^{-1} $controller^{-1}$ for high-bleed, low-bleed, intermittent-bleed, and malfunctioning controllers, respectively. We classified controllers sampled in Allen et al. (*55*) based on their measured emission rate and temporal profile; we define malfunctioning controllers as high-emitting controllers with assessed equipment issues in Allen et al. We assume that the national population includes 7% malfunctioning controllers, the same proportion as the Allen et al. dataset. This is a conservatively low estimate of the fraction of malfunctioning controllers since Allen et al. only assessed the top 40 highest emitting controllers for equipment issues.

Equipment leak emissions were estimated using the same approach as Zavala-Araiza et al. (*23*). Briefly, the Allen et al. 2013 (*54*) dataset of 279 leaks at 150 well pads was used to create two distributions: 1) individual leak emission rates, and 2) the number of leaks per site based on the number of gas-producing wells per site. We used ArcGIS to cluster wells within 50 m as sites and then used a Monte Carlo model to assign number of leaks (dependent on well count) and emission rates. For oil wells with no gas production, we estimated emissions separately using GHGRP reported emissions from heavy crude service equipment, which has a statistically significant correlation with oil-only well counts ($R^2 = 0.32$). Bootstrap sampling with replacement of the GHGRP data was used to calculate oil-only EFs with 95% CIs. Produced water tank flashing emissions were estimated using a similar approach as Zavala-Araiza et al. (*23*). When available, DI data was used to determine county-level water production for gas wells and oil wells. For states without reported well-level water production, we used state-level water production and well counts to estimate water production from county-level well counts (*58*). Emissions were estimated from water production volumes using oil and gas well EFs from the EPA O&G Estimation Tool (*59*); we assume that all produced water tanks are uncontrolled.

Emissions from well completions and workovers were estimated using the same approach that we used for the bottom-up estimate in Section S1.2.

S1.5. Local Distribution and end-use emissions.

Our emission estimate for local distribution (LD) systems is taken directly from the EPA GHGI (*17*), which incorporates recent emission measurements reported for pipeline leaks and metering and regulating (M&R) stations (*50*). These new measurements showed a decrease in emissions from M&R stations which was supported by widespread indications that many M&R stations have been upgraded or rebuilt since the early 1990s. For pipeline leaks, new emission factors were derived from 230 leak measurements across 13 cities and combined with leak counts reported from routine gas company surveys to estimate U.S. emissions. At the same time, top-down (TD) studies of urban CH_4 emissions in Boston (*14*), Indianapolis (*15*), and Los Angeles (*16*) suggest that the loss rates for pipeline quality gas are ~2-3% of total gas delivered. The loss rates from the TD studies are more than an order of magnitude larger than the GHGI national estimate (0.16% [95% CI: 0.081%–0.35%] of 270 Tg CH_4 delivered in 2015 for uses other than electric power [15 tcf natural gas, assuming CH_4 content of 95 vol% (*60*)]).

However, a direct comparison of CH_4 emissions rates in the GHGI and in the urban areas sampled in TD studies is complicated because some of the difference is expected to be caused by leaks that occur downstream of customer meters, as well as end-use consumption of natural gas by residential, commercial and industrial consumers (*61*, *62*). Such methane emissions downstream of customer meters would be observed in TD studies but are not included in the LD inventory, whose boundary is the customer meter. In addition, some of the difference between the GHGI and TD studies could be caused by an underestimation of the number of pipeline leaks and/or measurement or sampling bias that result in underestimated emission factors. However, we know of no published measurements that partition urban CH_4 enhancements between emissions from LD infrastructure and emissions that occur downstream of customer meters. On the balance of this evidence, we believe that the GHGI estimate for CH_4 emissions from the LD system in Table 1 should be considered to be a lower bound.

More attention is needed to quantify and apportion the urban CH_4 budget because emissions from LD and end use may materially affect the overall life-cycle emissions of certain uses of natural gas. In particular, the presence of diffuse low-level sources needs to be much better quantified including undetected leaks and losses downstream of customer meters. Improved TD estimates including more constrained uncertainty bounds are also needed along with better ways to apportion TD CH_4 emissions between natural gas distribution, sources downstream of customer meters, and other sources unrelated to natural gas.

If the TD results (*14-16*) are representative of other U.S. cities served by natural gas distribution systems, then EPA's estimate of LD emissions may be low, potentially by a substantial amount depending on the magnitude of emission downstream of customer meters. For example, consider that 2.7% of the methane in delivered natural gas is emitted along the supply chain up to "city gates" (Section S1.7), the transfer points where LD systems receive natural gas from the interstate transmission pipeline system (the estimate of 2.7% excludes the GHGI estimate of ~0.2% for LD, see Table 1). If losses from LD and end use were 2-3% of gas delivered, then the life-cycle CH_4 emission rate of natural gas sourced from LD systems could be as much as double the amount of CH_4 emitted upstream of the city gate. Large increases in life-cycle CH_4 emissions would have a material effect on the climate implications of consumptive end use of natural gas that is delivered by LD pipelines. By contrast, losses downstream of city gates would not affect life-cycle implications of coal-to-gas fuel switching for electrical generation in the U.S. for any natural gas power plant that obtains natural gas directly from transmission pipelines (this distinction would also apply to other end uses upstream of city gates).

Moreover, the contribution of LD systems and natural gas end uses to total U.S. CH_4 emissions could be substantial if CH_4 emission rates from TD studies of 2-3% (*14–16*) are representative of other U.S. cities. For example, an additional 5 Tg CH_4 would be emitted relative to the 0.4 Tg CH_4 reported in the GHGI for LD systems if 2% of the natural gas delivered to residential, commercial, and industrial consumers in the U.S. were, in fact, emitted – an increase of more than 35% in the national supply chain emissions estimated in this work (Table 1).

As new measurements of LD and end use emissions become available, our estimate of national emissions can be updated by replacing the LD estimate in Table 1 with new values.

S1.6. Alternative hypothesis: emissions higher in daytime than nighttime.

We evaluated an alternative hypothesis that could explain the large differences between inventories of annual emissions and independent TD measurements of production area emissions or measurement-based BU estimates: O/NG emissions are systematically higher during daytime hours when TD and BU measurements have been made, and lower at night. This situation was reported for the Fayetteville Shale but appears to be unique because the effect is caused by manual liquids unloadings, which represent a much higher fraction of total production emissions than in any other basin. Manual liquid unloadings are a process where a production well is vented to the atmosphere to remove accumulated liquids from the wellbore. Because the

procedure requires the presence of an operator at the site, it is preferentially conducted during working hours. In the Fayetteville Shale, manual liquid unloadings have been reported to account for 30-50% of the total O/NG emissions estimated by two TD flights (*47*). These manual liquid unloadings emissions in the Fayetteville are ~5x the value expected from the hourly average of annual unloading emissions in our inventory and are almost as large as the total emissions from other production sources in our BU estimate. Our BU estimate for production sites is based on measurement data from six production areas and is unlikely to be significantly affected by liquids unloadings (Table S9). The measurements in the Fayetteville explicitly exclude manual unloadings (Section S1.3), and the percent of wells with emissions from liquids unloadings at any one time in the other sampled basins is more than an order of magnitude lower than the Fayetteville (Table S9).

In addition, there is no reason to expect daytime bias in the kinds of abnormal operating conditions that are thought to characterize high-emitting production (and gathering) sites, which operate continuously. In fact, it is plausible that abnormal emissions could actually be higher at night because they are less likely to be found and corrected in the absence of operators.

For other segments, enhancements in daytime emissions could also result from episodic activities that are more likely to occur during working hours, such as maintenance-related blowdowns at compressor stations or processing plants, where pressurized equipment is vented to atmosphere; such events have been observed to result in emissions on the order of 10^2–10^4 kg/h (*39*, *63–65*). We conducted a probabilistic analysis of the range of potential emissions from such episodic events by combining reported blowdown frequencies and volumes in the GHGRP with the number of transmission/storage and processing plants in each TD basin (assuming the duration of reported blowdown events is 1 hour). Estimated basin-level, aggregate blowdown emissions are highly variable with median rates aligning well with our BU estimate (about 10-20% of each segment's total emissions). Although basin-level blowdown emission rates can exceed 10^4 kg/h, emissions this large should occur less than ~1–10% of daytime hours.

We conclude that an alternative hypothesis of higher daytime emissions is unlikely to explain the discrepancy in TD and inventory estimates (or this work's BU and inventory estimates) in basins other than the Fayetteville.

S1.7. Climate impacts of O/NG CH_4 emissions

The main text contextualizes the climate impacts of CH_4 emissions from the O/NG supply chain in two ways:

1) [O]ur estimate of CH_4 emissions across the supply chain, per unit of gas consumed, results in roughly the same radiative forcing as does the CO_2 from combustion of natural gas over a 20-year time horizon (31% over 100 years).

2) [T]he climate impact of 13 Tg CH_4/y over a 20-year time horizon roughly equals that from the annual CO_2 emissions from all U.S. coal-fired power plants operating in 2015 (31% of the impact over a 100-year time horizon)

Here we summarize how those comparisons were made.

The equivalent CO2 emissions (CO_2e) of a pulse of CH_4 emissions can be estimated using the global warming potential (GWP) of methane, referenced to a specific time horizon of 20 years (GWP20) or 100 years (GWP100). In this work we use a GWP20 of 96 and a GWP100 of 32, which are 14% larger than the corresponding IPCC AR5 values (84 and 28, respectively, exclusive of the effects of climate-carbon feedbacks and CO_2 from CH_4 oxidation), based on recently published calculations that include shortwave forcing (*66*). Using the GWP20/100 values of 96/32, we determine that 13 Tg CH_4 is equivalent to 1,300/420 Tg $CO_2e(20/100)$.

The EPA GHGI reports 2015 emissions of CO_2 from coal-fired electric generation and from gasoline vehicles are 1,350 Tg and 1,070 Tg, respectively (EPA GHGI 2017 (*17*), Tables 3-9 and 3-12 respectively). Therefore, the $CO_2e(20/100)$ corresponding to the 13 Tg CH_4 emitted from the O/NG supply chain is 94%/31% of the emissions from U.S. coal-fired electric generation and 120%/39% of the emissions from U.S. gasoline vehicles.

We use Eq. 4 to estimate the contribution of the CH_4 emitted from the natural gas (NG) supply chain prior to combustion to the total radiative forcing (RF, a surrogate for climate warming or climate damage) of NG production and use:

$$\text{Fraction of total RF from } CH_4 \text{ losses} = \frac{RF_{CH_4}}{RF_{CO_2} + RF_{CH_4}} = \frac{19.2 \cdot L \cdot GWP_{20/100}}{58.1 + (19.2 \cdot L \cdot GWP_{20/100})}, \quad (4)$$

where L represents the supply chain loss rate (CH_4 emitted per unit of CH_4 consumed) and GWP20/100 are the 20- and 100-year Global Warming Potentials for CH_4 (96 and 32, respectively). For this calculation, we assume that each unit of natural gas consumed (one unit of natural gas = 1 standard cubic foot, or scf) contains 100% CH_4; each scf of CH_4 contains 19 g CH_4; each scf of gas combusted produces 53 g CO_2; and an additional amount of CO_2 equal to 10% of the CO_2 from end-use combustion comes from leaks, venting and NG combustion along the supply chain (*5*). Results under a range of supply chain loss rates are shown in Table S10. Our estimate of supply chain emissions, 13 Tg CH_4/y, represents 2.9% of total methane delivered (25 tcf/y NG delivered, assuming an average CH_4 content in NG of 95% by volume). At a loss rate of 2.9% of gas delivered, CH_4 emissions across the supply chain, per unit of gas consumed, result in 92% of the radiative forcing caused by the CO_2 from combustion of natural gas, over a 20-year time horizon (31% over 100 years).

S1.8. Uncertainties in top-down emission estimates using the aircraft mass balance technique

Uncertainties in the TD aircraft mass balance method used to estimate O/NG emissions using measured enhancements of CH_4 mixing ratios downwind of a source region are described here. Briefly, the primary uncertainties include: 1) the determination of upwind boundary conditions; 2) adherence to the assumptions of steady-state, homogeneous winds; 3) the apportioning of total CH_4 emissions between O/NG and non-O/NG sources (the latter include other thermogenic sources like coal mines, as well as agricultural, urban, and natural sources), and 4) accurate quantification of atmospheric winds and mixing depth. These factors are considered to varying degrees and with varying methodology in the studies in Table S1; their

relative effects on each study region (especially boundary conditions and partitioning of sources) will vary significantly. Because the conclusions of this work are based on comparisons of bottom-up estimates to an ensemble of estimates from 9 regions, the uncertainties discussed in this section are very unlikely to change the overall conclusions of this work. Most of these sources of uncertainty were considered in most of the TD studies. Methods and magnitudes of uncertainty estimates differ across these studies, but no likely mechanism for large systematic errors is evident. A comparison the mass-balance approach used in most studies with an approach using a numerical reanalysis of atmospheric transport that does not require the steady-state, homogeneous winds assumption to hold did not reveal any significant systematic bias between these methods in Northeast Pennsylvania (*67*). Similarly, recent surface CH_4 observations in the Uinta basin coupled with an atmospheric transport model (*68*) corroborate the aircraft mass balance estimate of Karion et al 2013 (*69*).

Variability in background. Multiple approaches have been used to estimate CH_4 background concentration. In most studies (e.g., (*51*, *70–72*)) multiple methods have been explored to identify emission sources upstream of the region of interest. In large basins, the time that it takes for an air mass to travel from one side of the basin to the other makes it difficult to measure the true upwind boundary condition because the atmospheric boundary layer in the early morning is stratified, and the lowest layers are too low to be sampled by aircraft. Instead, the boundary condition is extrapolated from the edges of the regional plume assuming spatially homogeneous contributions upwind of the region of interest. In many studies this assumption is tested with upwind transects which are used to evaluate emission sources that exist upstream of the region of interest. This approach can lead to biases if there are upwind emissions (from outside the study area) that affect downwind measurements but are not picked up at the time of the upwind background measurement, or vice versa. Poor characterization of background CH_4 in the upwind boundary can result in a high bias in TD results if plumes from outside the study region are not accounted for (*67*).

Meteorological conditions. Aircraft mass-balance studies rely on the basic assumption that winds are approximately constant in space and time over the region of interest. Temporal and spatial heterogeneity of boundary layer winds can result in both over and under estimates of emissions depending on whether the variability in wind speed is coincident with the spatial heterogeneity of emitters characteristic of oil and gas fields. A second source of uncertainty is the height of the planetary boundary layer and accounting for daytime entrainment of free tropospheric air into the planetary boundary layer. During cloud free, midday conditions over land, thermodynamically dense air in the boundary layer is diluted by turbulent entrainment of less dense free tropospheric air into the boundary layer (*73*). Convective clouds, the collapse of the daytime boundary layer at night, and extreme variations in surface conditions across the study region can cause mixing of boundary layer air into the free troposphere, but these conditions are typically avoided for mass-balance aircraft flights.

Apportioning of CH_4 emissions between O/NG and other sources. Another source of uncertainty of aircraft-based, top-down surveys is determining the contribution of O/NG sources to total observed emissions. Aircraft surveys capture emissions from O/NG sources in addition to other thermogenic and biogenic sources in the contributing source region. Two approaches have been employed to separate the thermogenic from the biogenic. Source attribution is usually done

indirectly, using inventory data to account for biogenic sources and non-O/NG fossil fuel sources like coal mines (Table S2). Biogenic emissions may vary with temperature and meteorological conditions. The high variability in CH_4 and the low variability of ethane emissions observed by Karion (*71*) and Smith (*72*) in the Barnett would suggest that variability in biogenic emissions are a large source of uncertainty but it is unclear whether this is a consistent source of bias across all basins. A second approach to separating thermogenic emissions from biogenic emission has been the use of higher molecular weight hydrocarbons like ethane (e.g. (*49*, *72*)) and propane (*70*), which are co-emitted with CH_4 from O/NG sources. The challenge with the chemical fingerprinting approach is that the hydrocarbon ratios in a given basin can be highly spatially variable depending on the heterogeneity in reservoirs producing O/NG. The use of a median gas composition can lead to significant biases in our estimate of the fraction of the total measured by aircraft-based mass balance that is due to biogenic sources. Most studies use a combination of both approaches, where ethane measurements are used to validate inventory-based estimates for non-oil gas sources. Using isotopic instead of alkane tracers or the inventory method may reduce the uncertainty in the total amount of CH_4 attributed to O/NG sources in TD studies (*74*).

The magnitude of possible source attribution bias is constrained by the small-to-modest relative contribution of non-O/NG sources in the nine surveyed O/NG production areas, averaging 12% of the total TD CH_4 estimates with only one area >25%: Weld (27%, driven by biogenic sources).

Time of sampling – A potential source of uncertainty in comparing TD studies to annual emission inventories could result if emissions during aircraft sampling (typically centered on mid-day hours) are higher than at other times of day. This uncertainty is examined in more detail in Section S1.6.

S1.9. Alternative method to estimate emissions at production sites

We estimated national CH_4 emissions from O/NG production sites using the same dataset of production site emissions measurements used in the main Methods section (Section S1.2) but applying an alternative method in order to assess the robustness of our reported bottom up estimate. Instead of estimating emissions conditional on production, the alternative method assumes that, within discrete cohorts of gas production, a production site's emissions can be treated as independent from gas production. We identify three different gas production cohorts: >0-10 Mcf/d (43% of U.S. gas producing sites and 1% of U.S. gas production), 10-5,000 Mcf/d (57% and 53%) and >5,000 Mcf/d (0.45% and 40%)) (Figs. S2 and S9) and determine an emission factor (EF) for each cohort. These cut points were selected based on the observation that emissions for the lowest and highest producing cohorts exhibit lower and greater variability than the central cohort, based on rolling window averages (Fig. S9). To account for potential inter-basin variability in emission rates, we develop emission factors for each cohort-basin subset with sufficient measurements. We assume that the emission rate distributions within each cohort are lognormal, but require a minimum of 25 samples to produce a reasonably characterized emissions probability density function (pdf). This alternative method relies on 418 site-level measurements of natural gas production sites in six basins (*19–21*). SW PA was not included in the 10-5,000 Mcf/d cohort because there were only 15 measurements.

For the two lowest producing cohorts that have sufficiently large sample sizes, we determine characteristic emission factors by sampling from a lognormal distribution fitted to the corresponding measurements. Consequently, for the [10-5,000 Mcf/d] cohort we are able to produce an independent pdf for five of the six basins with site-level measurements (a SW PA cohort is not possible because N<25). For the [>0-10 Mcf/d] cohort we combine all samples from all basins and produce a single national lognormal-based EF. For the [> 5,000 Mcf/d] cohort, we do not produce a lognormal fit because the sample size is too small even after combining data from all basins, therefore we estimate the EF using a bootstrap resampling directly from the reported measurements instead.

Let x be the natural logarithm of CH_4 emissions (in kg/h) measured at a production site, the log likelihood function is:

$$l(\mu,\sigma) = S_0 ln\,\Phi\left(\frac{x^*-\mu}{\sigma}\right) - S_r ln\sigma - \sum_{i=1}^{S_r}\frac{(x_i-\mu)^2}{2\sigma^2} \quad (5)$$

where x^*, is the detection limit for the production site measurements (0.08 kg/h, 0.036 kg/h and 0.01 kg/h in Rella et al., Robertson et al., and Omara et al., respectively), S_0 is the number of measurements at or below the detection limit and S_r is the number of measurements above the detection limit, and $\Phi\left(\frac{x^*-\mu}{\sigma}\right)$ is the cumulative normal.

We estimate the μ, σ for each pdf by solving for the values that maximize Eq. 5 and use a direct search algorithm to calculate 95% confidence limits by inverting the Likelihood Ratio Test. Table S11 summarizes parameters, EFs for each cohort and Fig. S10 illustrates the estimation method.

National emissions from routine operations at natural gas producing sites

We estimate national emissions from production sites using a Monte Carlo sampling approach. For each of the 62 U.S. O/NG production basins:

- [>0-10 Mcf/d]: we draw a random mean from the emission distribution and apply it to all sites within this gas production cohort.
- [10-5,000 Mcf/d]: we randomly choose one of the five EF and draw a random mean from its emission distribution and apply it to all sites within this gas production cohort.
- [>5,000 Mcf/d]: we sample (with replacement) from the raw data to assign an EF to sites within this gas production cohort.

We repeat this process 10,000 times. Total emissions from routine operations at U.S. production sites with non-zero natural gas production are 8,400 Gg/y (95% CI: 6,700-11,000 Gg/y), compared to 6,700 Gg/y (95% CI: 5,100-8,600 Gg/y) from the method described in Section S1.2 (also see Table S6 and Table S12). If we use this alternative method of estimating production site emissions and determine a bottom-up emissions estimate of total O/NG emissions in the nine areas with reported top-down O/NG emission estimates, results are similar to those reported in main text. The 9-basin TD mean is 11% higher than the BU mean but the difference is not statistically significant (p=0.32 by a randomization test, Fig. S11). It is also worth noting that this alternative method and the one presented in the main text yield the same estimate for the 9-basin BU mean (320 Mg/h).

Even though results from the alternative method are similar to those in the main text (i.e., emissions conditional on production), we chose the latter as the primary method because it represents a single parsimonious solution that can be applied across the entire range of gas production represented in the sampled population of sites without having make somewhat arbitrary assumptions about the break points for gas production cohorts. The main method also allows for the use of the full dataset of measurements from SW PA and, consequently, the use of 6 basins instead of 5 when scaling the measurements to the national population (for all sites >0.68 Mcf/d, where emissions are concentrated).

Table S12 compares the results of the alternative and main method used in this paper. The alternative method assigns higher emissions across all but the highest natural gas production cohorts. Both methods assign the bulk of emissions (~90%) to the two central cohorts (10-5,000 Mcf/d) which correspond to the range for which the vast majority of production site emission measurements were made.

S2. Additional author disclosures

RAA serves on the Science Advisory Committee of the Center for Air, Climate, and Energy Solutions, for which ALR serves as co-director.

DRL is a member of the Technical Advisory Steering Committee for United States Department of Energy (USDOE) Projects DE-FE0029084/DE-FE0029085 on methane emissions quantification at natural gas gathering and storage facilities.

DTA has served as chair of the United States Environmental Protection Agency (USEPA) Science Advisory Board (2012−2015; in this role, he was a paid Special Governmental Employee); DTA has current research support from the National Science Foundation (NSF), the Texas Commission on Environmental Quality, and Exxon Mobil Upstream Research Company. Over the past five years, he has also worked on methane emission measurement projects that have been supported by multiple natural gas producers and Environmental Defense Fund (EDF). DTA has done work as a consultant for multiple organizations, including Colorado State University, Eastern Research Group, ExxonMobil, and Research Triangle Institute.

ZRB was partially supported by a Dennis and Joan Thomson Distinguished Graduate Fellowship in Meteorology from The Pennsylvania State University.

ARB has current and recent research support from a variety of governmental, non-profit, and industry sources. Governmental funding sources include funding from the California Environmental Protection Agency, California Air Resources Board, the National Renewable Energy Laboratory, Sandia National Laboratory, and Argonne National Laboratory. Non-profit funding sources include the Carnegie Endowment for International Peace, Novim (affiliated with University of California Santa Barbara), and EDF. Industry sources of funding include Aramco Services Company, Ford Motor Company Inc., and Seven Generations Energy Ltd. Internal Stanford University funding include Stanford Natural Gas Initiative and various graduate fellowships.

KJD's current research is supported by the National Aeronautics and Space Administration (NASA), the National Institute for Standards and Technology (NIST), and the NSF. He is a co-owner of CarbonNowCast, LLC, and in that capacity has received past research support from EDF.

SCH is employed by Aerodyne Research, Inc. (ARI), which has current and recent research support from government, non-profit, academic and industry sources. Government funding includes various agency Small Business Innovation Research programs and the State of Texas Air Quality Research Program. Non-profit funding sources include EDF. Academic sources include universities that purchase instrumentation associated with measurement of methane, ethane and propane. Industry funding sources include Eastern Research Group and Shell Oil Company.

DJJ has received methane-related research funding from NASA, USDOE Advanced Research Projects Agency-Energy, and Exxon-Mobil.

EAK has current research and past research support from NSF, NASA, NIST, and EDF. He has worked on methane emissions measurement projects supported by multiple natural gas producers and EDF.

BKL has served on science advisory panels for EDF related natural gas emission studies, for Bureau of Ocean Energy Management related to air quality modeling in the Arctic, and for USDOE supported studies of natural gas emissions. His current research is supported by United States Department of Agriculture (USDA), USEPA, Bureau of Land Management (BLM), and state clean air agencies in the Pacific Northwest. In recent years, he has worked on natural gas emission studies supported by EDF, USDOE, and NYSEARCH.

TL's current research is supported by NASA, NIST, and the Japanese Space Agency (JAXA). He is a co-owner of CarbonNowCast, LLC, and in that capacity has received past research support from the Bay Area Air Quality Monitoring Division and EDF.

AJM has current research support from the NSF and the USDOE. Over the past five years, he has worked on methane emissions studies that have been supported by EDF and multiple natural gas companies. AJM currently serves as a consultant with Abt Associates on methane emissions projects with USEPA and New York State Energy Research and Development Authority.

SWP has been one of two P.I.'s of a grant to Princeton University from BP to research the carbon and climate problem since 2000. He is also on the board of EDF. SWP has current and recent funding from National Oceanic and Atmospheric Administration, NSF, USDOE, NASA and USDA.

ALR has current research support from USEPA and NSF. Over the past five years, he has worked on methane emissions studies that have been supported by EDF and multiple natural gas companies. ALR currently serves as a member of the Health Effects Institute research

committee. In 2014-2015, he served on the Health Effects Special Scientific Committee on Unconventional Oil and Gas Development in the Appalachian Basin.

ATS has current research support for methane emissions research from the BLM, US Department of the Interior. Over the past five years, she has also worked on methane emissions projects that have been supported by EDF and other groups including NSF and Colorado Department of Public Health and Environment. She is currently an unpaid member of a scientific advisory panel for USDOE study investigating methane emissions quantification from natural gas storage and compressors. She previously served on a scientific advisory panel for an EDF study investigating methane emissions from natural gas distribution systems for which she was paid an honorarium ($2000). She has not done any paid consulting and does not have any financial interest in the oil and gas sector.

SPH serves as a member of the USEPA's Science Advisory Board (2013− present; in this role, he was a paid Special Governmental Employee for several years, currently is not); SPH served on and chaired the External Advisory Committee of the AirWaterGas NSF Sustainability Research Network at University of Colorado. He also serves as the Chief Scientific Officer of the International Methane Studies which are being carried out under the auspices of the Climate and Clean Air Coalition which is part of the United Nations Environment Program (SPH receives no compensation from the UN for these efforts).

Other co-authors (DZ-A, AK, JDM, MO, JP, PS, CS, SCW) report no additional disclosures.

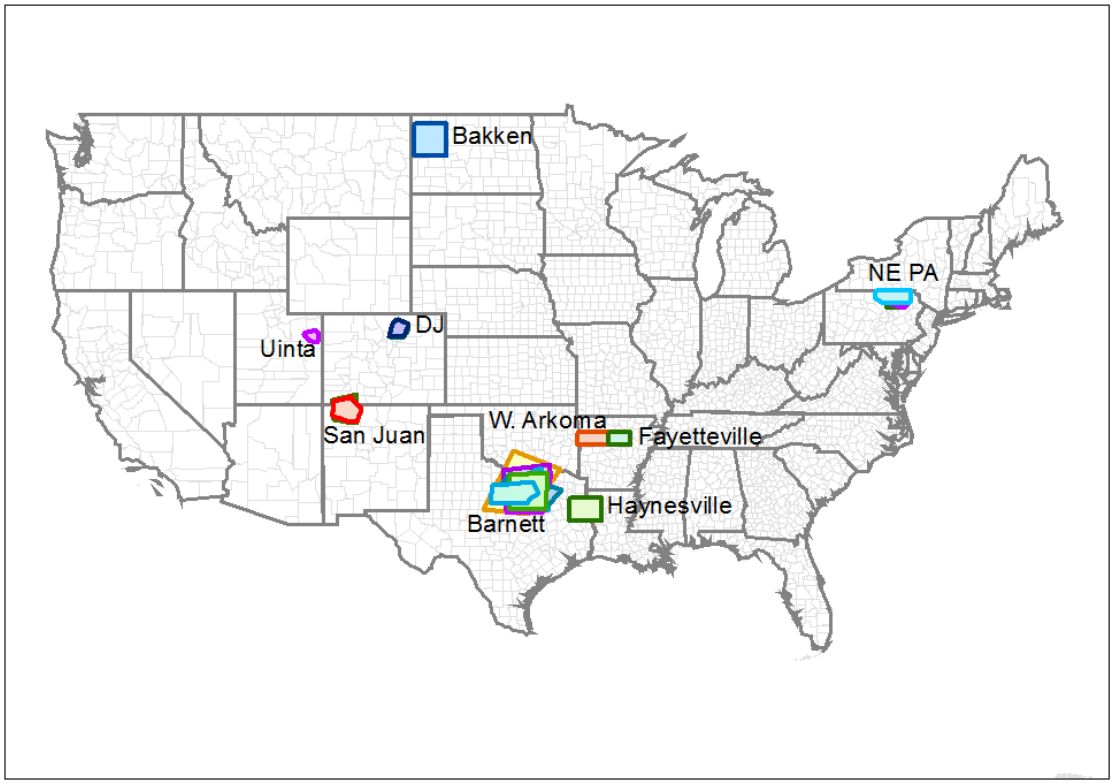


Fig. S1.
Map of O/NG production areas sampled in aircraft-based top-down studies (see Table S2). Colored polygons represent estimated boundaries of sampled source regions, which sometimes differ between multiple flights. Boundaries were determined from information in the original papers and/or in consultation with the corresponding authors (shapefiles provided as Supplementary Material).

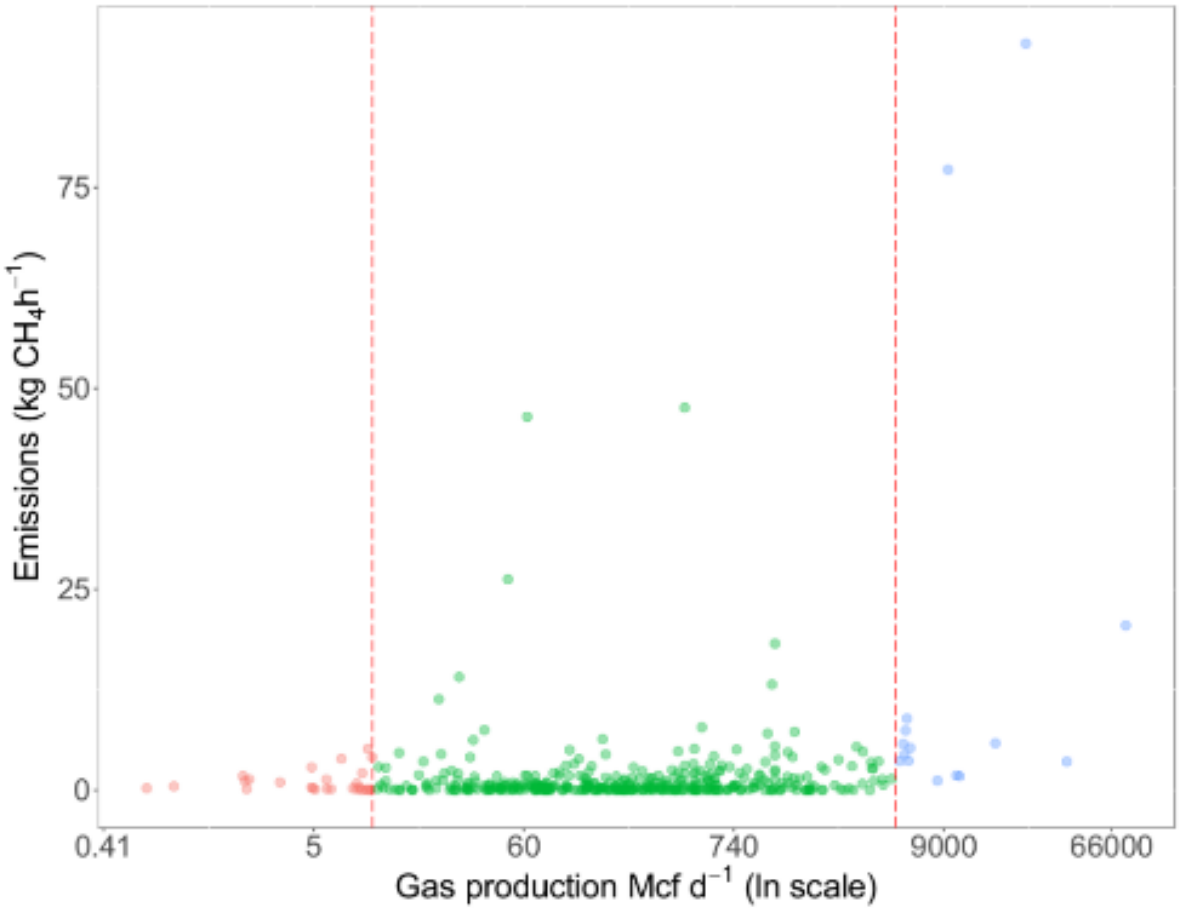


Fig. S2.

Measured CH_4 emissions (kg/h) from the sample of 433 natural gas producing sites that are used in this work to estimate emissions from the national population of natural gas producing sites, shown as a function of each site's natural gas production (Mcf/d). Red vertical lines and changes in symbol colors at 10 and 5,000 Mcf/d indicate the selected breakpoints between gas production cohorts used in an alternative methodology presented in Section S1.9.

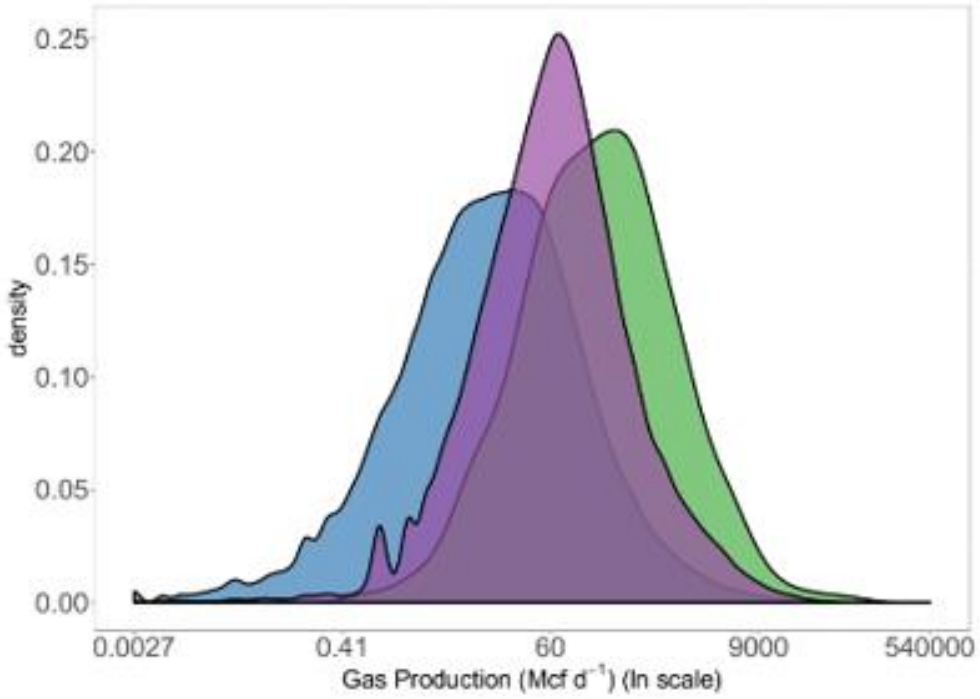


Fig. S3.
Density plot showing the distribution of daily natural gas production (in units of thousand standard cubic feet per day, Mcf d^{-1}) in the national population of production sites in 2015 (blue), production sites sampled in the datasets used in this work to estimate national production site emissions (green), and production sites in the 9 basins used for this work's comparison of TD and BU results (purple).

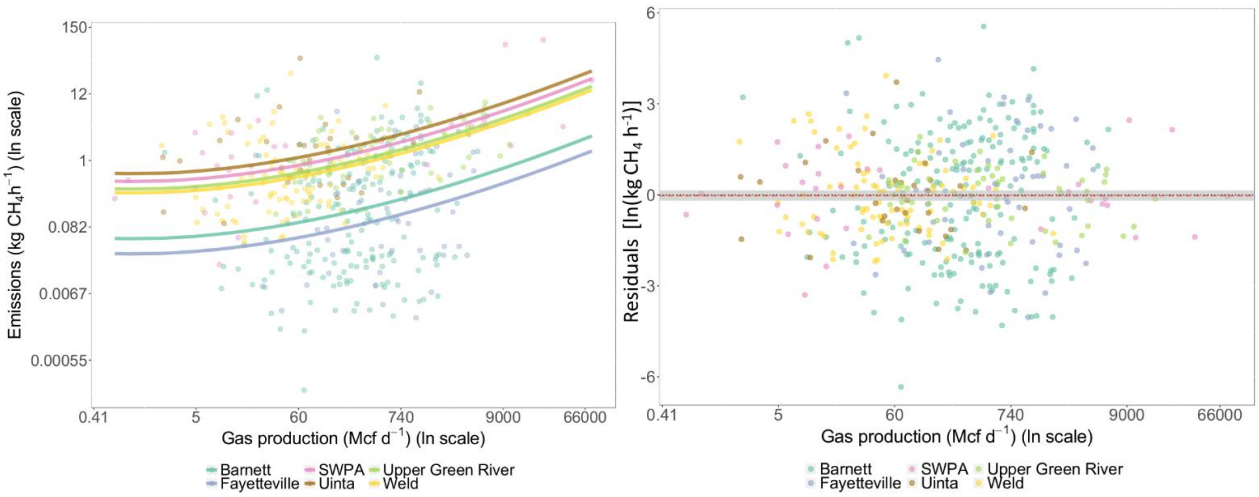


Fig. S4.
(Left panel) CH_4 emissions (in kg/h) from each measured natural gas producing site as a function of site-wide natural gas production (in Mcf/d). The solid lines show the non-linear model results using parameters in Table S5. (Right panel) Residuals (in ln(kg/h)) for the non-linear model that describes the relationship between emissions and production. The red dotted line and the gray shading in the right panel show the mean of all residuals and its 95% confidence interval.

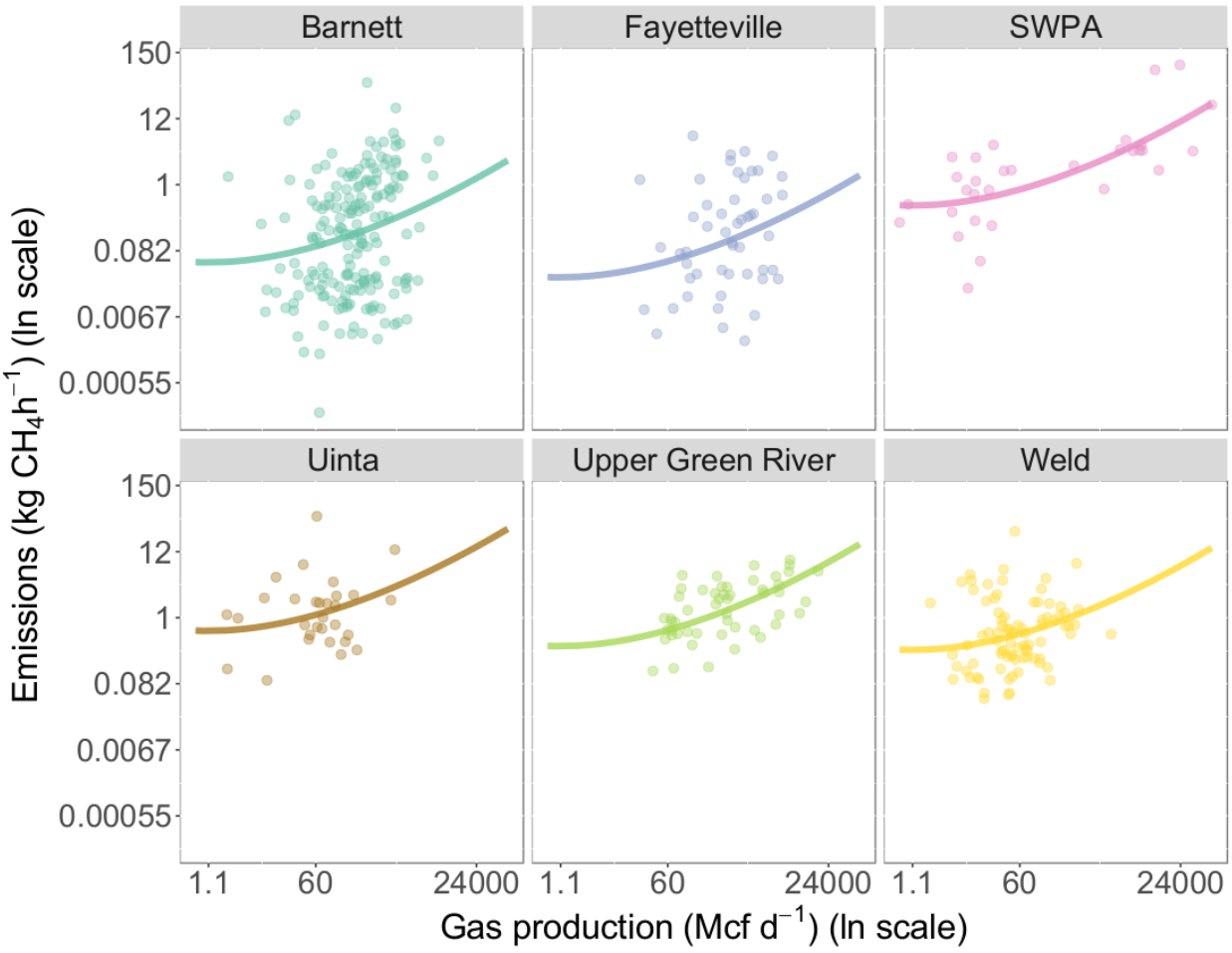


Fig. S5.
Basin-specific production site emission measurements with the fitted non-linear model results.

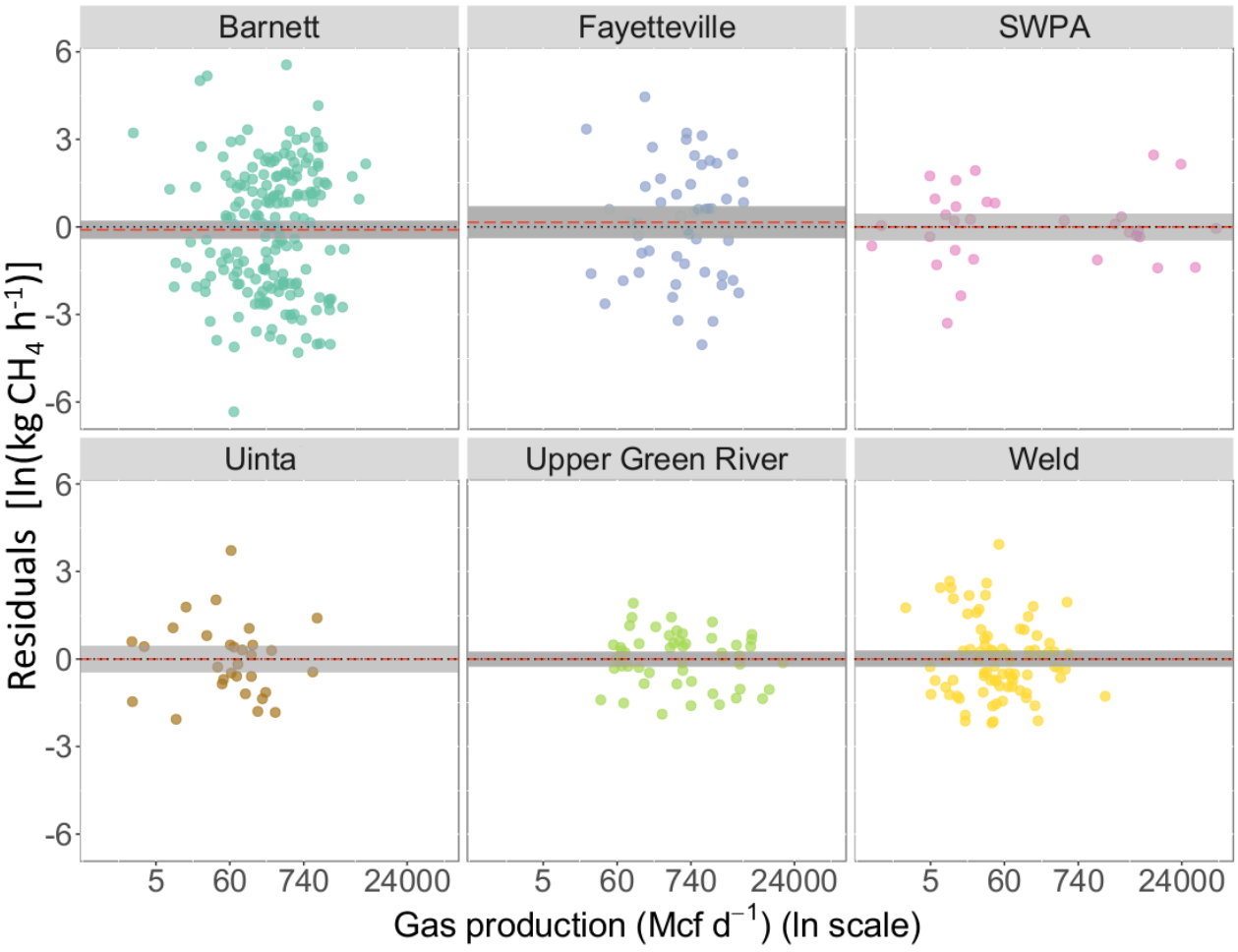


Fig. S6.
Residuals in Fig. S4 (in ln(kg CH4/h)) as a function of production for sampled basins. The red dotted line shows the basin-specific mean of residuals and grey shaded regions shows the 95% confidence interval.

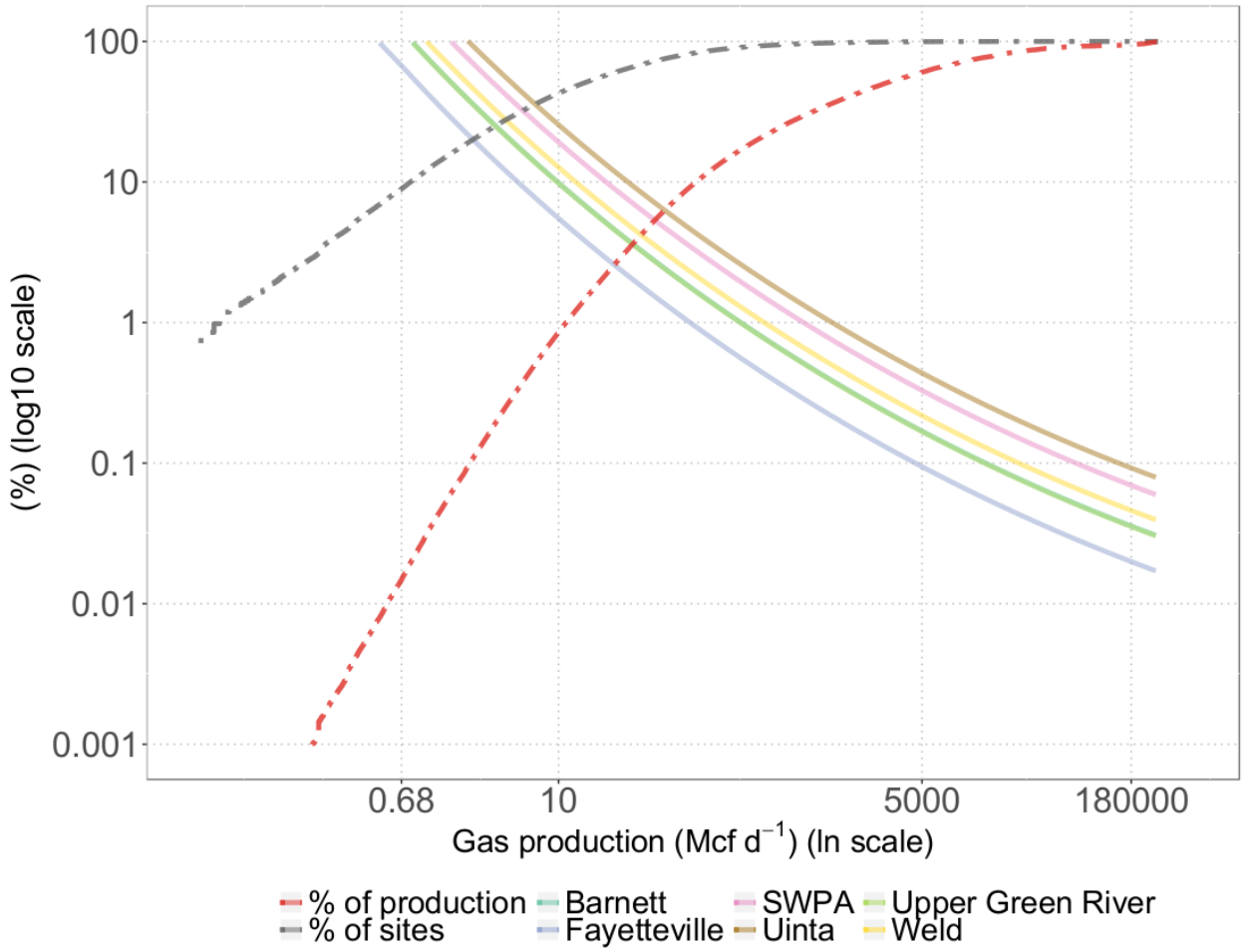


Fig. S7.
For each basin, the plot shows mean production normalized methane emission rates as a function of gas production classes, where methane emissions are normalized by methane produced, assuming 80% content (Barnett and Upper Green River curves overlap each other). The plot also shows the cumulative percent of sites and cumulative percent of production for the national population of gas producing sites.

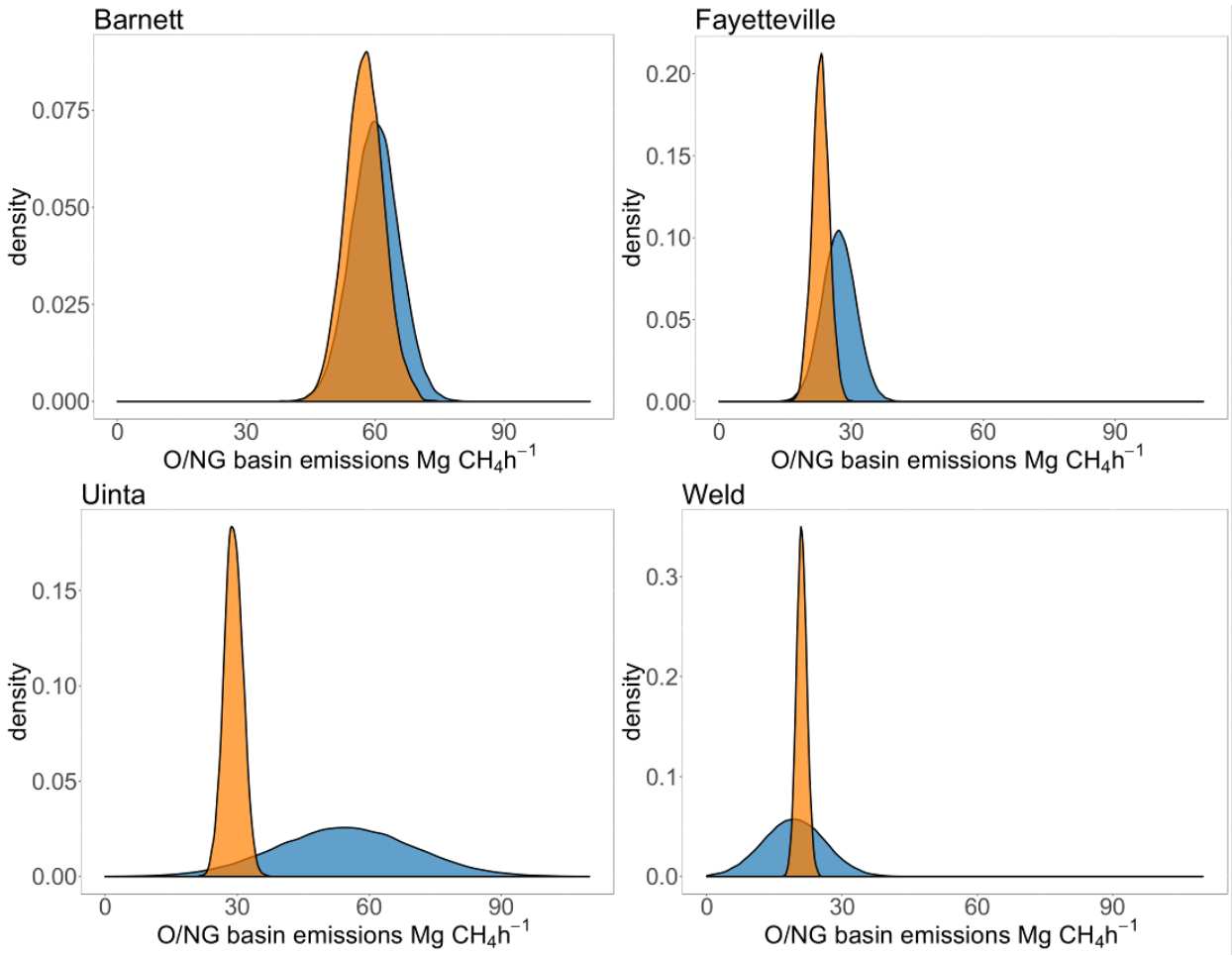
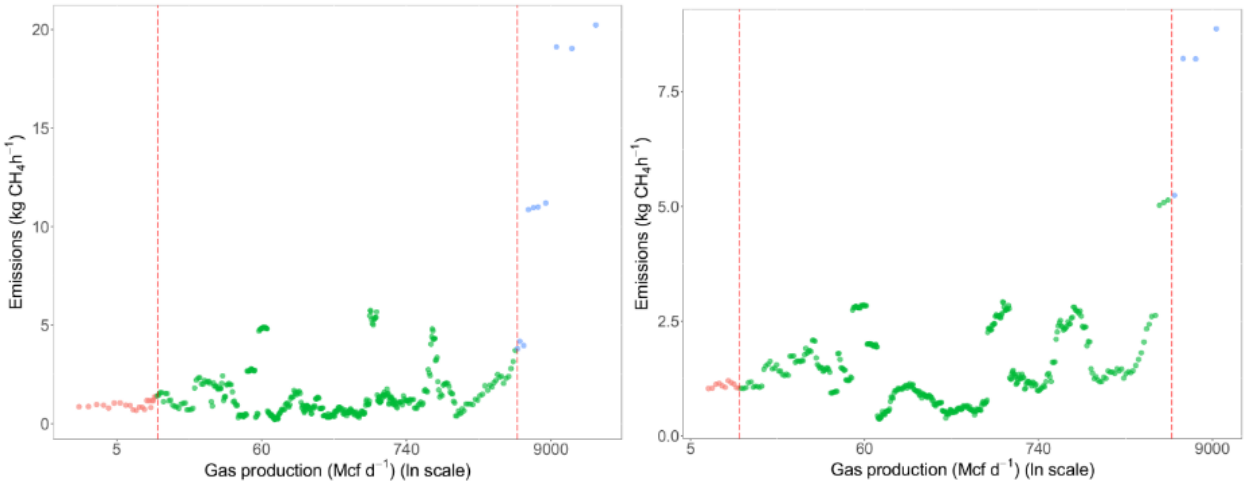


Fig. S8.
Individual density plots for the mean O/NG CH4 emissions in four basins with available top-down (TD) estimates (blue distributions) and local production site emission measurements (total bottom-up (BU) estimate of O/NG emissions shown as orange distribution). The favorable comparisons offer validation of the bottom-up methodology used in this work. Additional validation comes from the agreement between TD and BU estimates of total O/NG emissions in all nine sampled basins (Fig. 1b and Table S7.)



(a) **(b)**

Fig. S9.
Same data as Fig. 2 but using a moving window average of (a) 10 samples and (b) 30 samples.

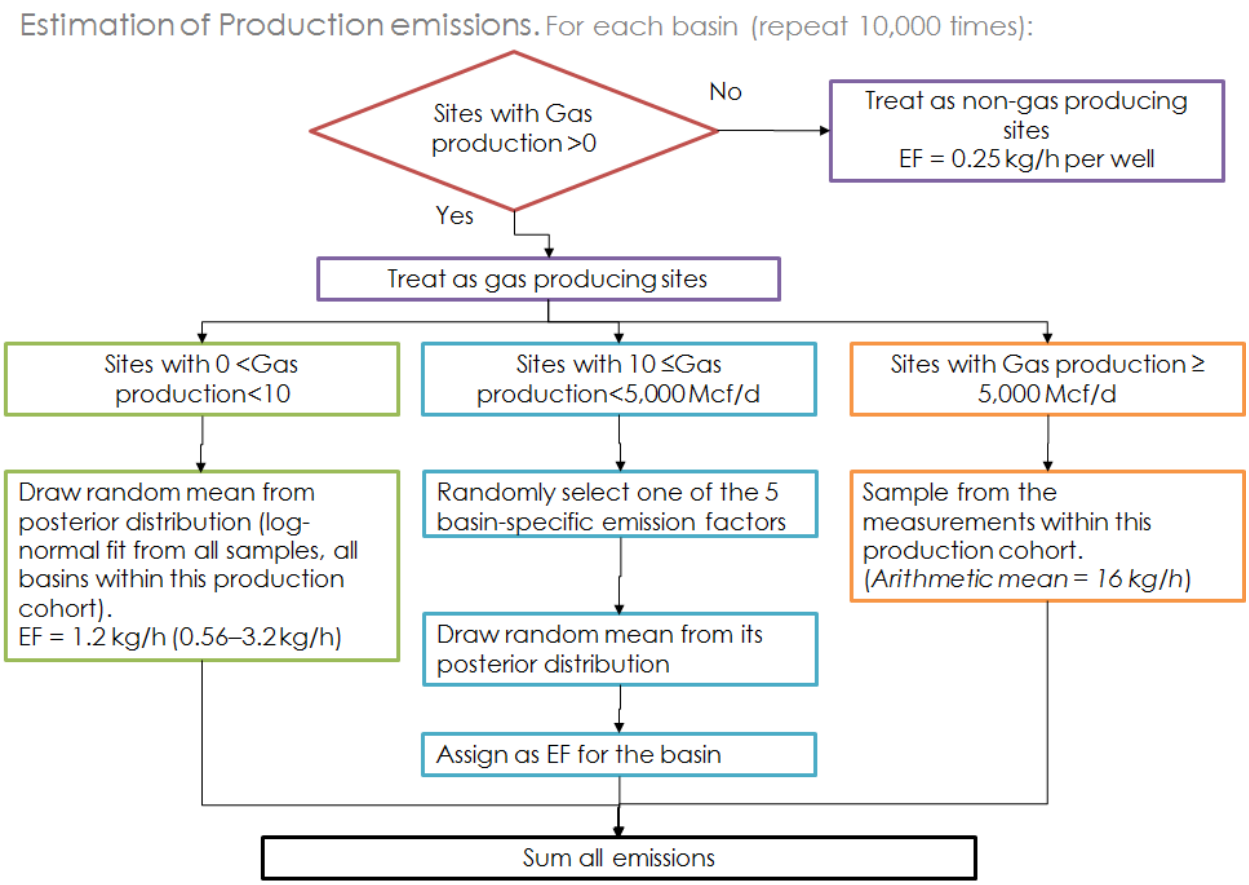


Fig. S10.
Flowchart summarizing the alternative methodology in Section S1.9 to estimate emissions from natural gas producing sites.

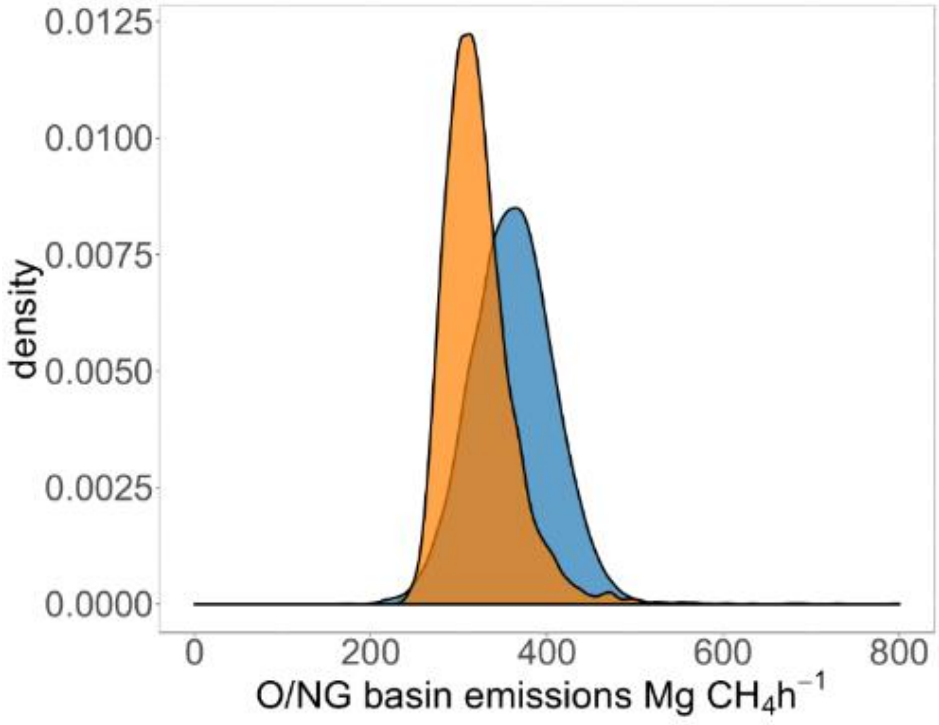


Fig. S11.
Distributions of the 9-basin sum of TD and BU mean estimates (blue and orange pdfs, respectively) when we estimate BU production emissions with the alternative method in Section S1.9.

Table S1. Datasets published since 2012 reporting source-specific emission measurements that comprised of 10 or more samples or used to characterize emissions from a population of sources. Italicized datasets were not used in this work because they could not be readily scaled to estimate national emissions.

Industry Segment	Source Category	Description	Reference for data source	Reference for additional analysis
O/NG Production	Production sites	186 sites in Barnett Shale	Rella (*19*)	Zavala-Araiza (*18*)
		31 sites in Marcellus Shale	Omara (*20*)	
		218 sites in 4 basins: Fayetteville (N=52), D-J (N=84), Upper Green River (N=51), Uinta (N=31)	Robertson (*21*), Brantley (*38*)	
		20 U.S. sites	*Allen* (*54*)	
	Pneumatic Controllers	377 controllers	Allen (*55*)	
	Equipment Leaks	278 leak measurements	Allen (*54*)	
	Pneumatic Pumps	62 chemical injection pumps	Allen (*54*)	
	Completions + Workovers	*27 completion flowback events*	*Allen* (*54*)	
	Abandoned and Orphaned Wells	42 wells (PA)	Kang (*45*)	
		138 Wells (OH, WY, UT, CO)	Townsend-Small (*44*)	
	Flares	37 flares in the Bakken	Gvakharia (*48*)	
	Liquids Unloading	*Unloading events at 107 gas wells*	*Allen* (*56*)	
Natural Gas Gathering	Gathering facilities	114 gathering facilities	Mitchell (*25*)	Marchese (*29*)
	Gathering Blowdowns	5 Events (10 plumes)	Mitchell (*25*)	Marchese (*29*), Zavala-Araiza (*18*)
Natural Gas Processing	Processing Plants	16 processing plants	Mitchell (*25*)	Marchese (*29*), Zavala-Araiza (*18*)
Transmission and Storage	T/S Stations	45 facility-level measurements; 1,398 on-site measurements (discrete sources)	Subramanian (*40*)	Zimmerle (*26*)
	Uncategorized/ Superemitters	2 facilities	Subramanian (40)	Zimmerle (*26*)
Multiple	Multiple	*140 production, compression and processing, facilities in the Barnett Shale*	*Lan* (39)	*Zavala-Araiza* (*18*)
		17 production, compression and processing, facilities in the Barnett Shale	*Yacovitch* (*65*)	*Zavala-Araiza* (*18*)
		13 production and gathering facilities in the Eagle Ford Shale	*Lavoie* (*75*)	
		14 compressor stations and production sites in the Marcellus Shale	*Goetz* (*76*)	
Local Distribution	Multiple distribution source types		Lamb (*50*)	
	Underground pipelines	*100 leaks from cast-iron distribution mains in Boston, MA*	*Hendrick* (*77*)	

Table S2. Reported estimates of O/NG CH_4 emissions from aircraft-based top-down (TD) studies, listed in decreasing order of natural gas production. Italicized values were calculated in this work; shaded rows indicate a second independent, statistically consistent set of reported measurements in two basins (not used directly in this work in favor of the more recent results based on more intensive sampling). Uncertainties are 2-sigma values calculated from reported uncertainties.

TD survey area	Reference	Date Sampled (Month/yr)	Days/flights/ downwind transects	NG production (bcf/d)	% CH_4 in NG	Upwind Background Method*	Total CH_4 Flux (Mg/h)	O/G apportionment method†	O/NG CH_4 flux (Mg/h)†	Production normalized emission rate$
Haynesville	Peischl (*51*)	6/2013	1/1/3	7.7	86%	UTA	80 ± 54	SE	*73 ± 54*	*1.3%*
Barnett	Karion (*71*)	3 & 10/2013	8/8/17	5.9	89%	DL	76 ± 13	E	60 ± 11	*1.4%*
NE PA	Barkley (*67*)	5/2015	4/4/7	5.8	95%	MUT	20 ± 17	SE	18 ± 14	*0.40%*
NE PA	Peischl (*51*)	7/2013	1/2	N/A	95%	UTA	15 ± 12	SE	13 ± 12	*0.30%*
San Juan	Smith (*52*)	4/2015	5/5/5	2.8	83%	DL	62 ± 46	N	*57 ± 54*	*3.0%*
Fayetteville	Schwietzke (*47*)	10/2015	2/2/4	2.5	97%	UTSV	31 ± 8	SE	27 ± 8	*1.4%*
Fayetteville	Peischl (*51*)	7/2013	1/1/2	N/A	97%	UT	39 ± 36	SE	35 ± 32	*1.9%*
Bakken	Peischl (*49*)	5/2014	3/3/5	1.9	47%	DL	28 ± 10	SE	27 ± 13	*3.7%*
Uinta	Karion (*69*)	2/2012	1/1/1	1.2	89%	UT	56 ± 30	S	55 ± 31	*6.6%*
Weld	Petron (*70*)	5/2012	2/2/3	1.0	79%	UT	26 ± 14	S	19 ± 14	*3.1%*
W Arkoma	Peischl (*51*)	7/2013	1/1/1	0.37	96%	UT	33 ± 30	S	26 ± 30	*9.1%*
9-basin total				29			410 ± 87		360 ± 92	1.8 ± 0.5%¶

* Upwind background methods: UT=upwind transect; UTSV = spatially variable upwind transect; UTA=upwind transect with adjustments to account for methane above background that flows into a region; DL = downwind lateral plume edges; MUT = model-assisted upwind transect

† Apportionment methods: S= subtraction of inventory-based estimates of non-O/NG sources; E = ethane; SE = subtraction with ethane as qualitative check; N = none

$ Methane emitted normalized by methane produced

¶ Production weighted

Table S3. CH_4 emissions from the U.S. O/NG supply chain in 2015 as estimated in the 2017 U.S. EPA Greenhouse Gas inventory (GHGI) (*17*) and in this work based on source-based (i.e., individual components inside sites or facilities, Section S1.4) and site-based (i.e., all sources at a site, Section S1.2) methodologies. Bold categories denote emission sources for which recent measurements have been reported.

Industry Segment	Source Category	2015 U.S. Emissions (Gg CH_4 y^{-1})		
		GHGI	This work (source–based)	This work (site–based)
O/NG Production	**Pneumatic Controllers**	1,800	1,100 (1,100 - 1,200)	7,200 (5,600 - 9,100)
	Equipment Leaks* $	360	620 (570 - 670)	
	Liquids Unloading	210	170 (170 - 200)	
	Pneumatic Pumps*	210	190 (180 - 200)	
	Oil & Condensate Tanks	100	100 (97 - 120)	
	Produced Water Tanks	40	360 (340 - 380)	
	Fuel combustion	240	98 (91 - 210)	
	Associated gas flaring and venting	150	71 (69 - 86)	
	Other production sources*	40	60 (58 - 68)	
	Routine Operations Subtotal	3,100	2,800 (2,700 - 2,900)	7,200 (5,600 - 9,100)
	Completions + Workovers	100	86 (80 - 120)	
	Abandoned and Orphaned Wells	NA	61 (59 - 360)	
	Onshore Production Subtotal	3,200	2,900 (2,900 - 3,300)	7,300 (5,700 - 9,300)
	Offshore Platforms	300	300 (240 - 380)	
	Production Total	3,500	3,200 (3,100 - 3,600)	7,600 (6,000 - 9,600)
Natural Gas Gathering	**Gathering Stations**	2,000	2,100 (2,100 - 2,200)	
	Gathering Episodic Events	200	170 (7 - 750)	
	Gathering Pipelines	160	310 (300 - 330)	
	Gathering Total	2,300	2,600 (2,400 - 3,200)	
Natural Gas Processing	**Processing Plants**	410	680 (610 - 880)	
	Routine Maintenance	36	36 (29 - 46)	
	Processing Total	450	720 (650 - 920)	
Transmission and Storage (T/S)	**T/S Stations**	1,100	1,100 (860 - 1,400)	
	T/S Uncategorized/Superemitters	NA	440 (350 - 570)	
	Transmission Pipelines	220	220 (180 - 290)	
	LNG Storage and Import Terminals	70	67 (54 - 87)	
	T/S Total	1,300	1,800 (1,600 - 2,100)	
Local Distribution	All sources through customer meters	440	440 (220 - 950)	
Petroleum Midstream	Oil Transportation + Refining	34	34 (26 - 84)	
Total U.S. Oil and Gas Supply Chain		8,100 (6,800 – 10,000)	8,800 (8,400 - 9,700)	13,000 (12,000 - 15,000)

* Denotes multiple GHGI source categories are combined into this source type.

$ GHGI combines compressor venting with compressor fugitives, thus we combine 48 Gg derived from GHGRP for compressor venting with 523 Gg for equipment leaks.

Table S4. Distribution of the activity data of U.S. oil and natural gas wells in 2015. The last row shows the percent of emissions from production sites calculated with the model described in this section. The production cohorts in this table were selected based on breakpoints evident in the dataset of production site emission measurements (Fig. S2 and Section S1.9), and 0.68 Mcf/d is the minimum production of the sampled population. The measurement dataset predominantly contains sites with gas production within the bolded gas production cohorts.

	% of US 2015 Activity Data by Gas Production Cohort				
Natural Gas Production Cohorts (Mcf d^{-1})	0	>0–0.68	**0.68–10**	**10–5,000**	>5,000
Sites*	15% (0%)	7.6% (8.9%)	29% (34%)	48% (57%)	0.38% (0.45%)
Wells	19%	5.1%	20%	53%	3.3%
Gas Production	0%	0.015%	0.84%	59%	40%
Oil Production	7.3%	0.49%	3.0%	74%	15%
Emissions*	6.4% (0%)	5.1% (5.5%)	20% (21%)	64% (68%)	4.8% (5.1%)

*The main value includes oil wells with zero reported gas production; the value in parentheses excludes them.

Table S5. Parameters that describe the emission distribution function conditional on production; 95% confidence intervals are shown between parentheses.

Basin	a	b	c	θ_1	θ_2	σ
Barnett Shale	0.83 (0.55, 1.1)	-2.2 (-2.6, -1.8)	0.20 (0.050, 0.42)	0.60 (0.44, 0.81)	1.4 (1.3, 1.8)	2.1 (2.0, 2.4)
Weld	2.6 (2.3, 2.8)					1.3 (1.1, 1.4)
Fayetteville	0.26 (-0.075, 0.54)					2.1 (1.9, 2.4)
SWPA	3.0 (2.6, 3.2)					1.3 (1.1, 1.6)
Uinta	3.3 (3.0, 3.5)					1.3 (1.1, 1.5)
Upper Green River	2.7 (2.5, 2.9)					0.90 (0.79, 1.0)

Table S6. Breakdown of this work's site-based, bottom-up (BU) estimate of CH_4 emissions from onshore U.S. oil and natural gas (O/NG) production.

Oil/Natural Gas Production Source	2015 U.S. Emissions (Gg CH_4/y)	Basis
Gas Producing Sites – routine operations	6,700 (5,100 - 8,600)	This work using measurements in (*19*–*21*)
Fayetteville – Manual Liquids Unloading	20	Schwietzke (*47*)
Bakken – Flaring	42	Gvakharia (*48*)
Oil-only sites	430 (210 - 640)	This work
Total O/NG sites – routine operations	7,200 (5,600- 9,100)	N/A
Completions, Workovers, & Well Testing	86 (80 - 120)	This work using GHGRP data (*46*)
Abandoned & Orphaned Wells	61 (59 - 360)	This work using measurements in (*44*, *45*)
Total	7,300 (5,700 - 9,300)	N/A

Table S7. Summary of central estimates (CE) of total oil and gas methane emissions based on top-down and this work's bottom-up estimates for each basin (Upper bound (UB) and Lower Bound (LB) represent the 95% confidence interval). Units are Mg CH_4/h.

TD Survey Area	Top-Down			Bottom-up		
	CE	LB	UB	CE	LB	UB
Barnett	60	49	71	57	49	66
Fayetteville	27	20	35	23	19	27
Uinta	55	24	85	29	25	33
Weld	19	5.8	33	21	19	23
Bakken	27	14	39	24	17	35
Haynesville	73	19	130	73	48	110
San Juan	57	4.3	110	57	31	93
NE PA	18	4.5	31	20	13	31
West Arkoma	26	-3.4	55	13	7.4	19
9-Basin Sum	360	270	450	320	270	380

Table S8. Sources of activity and emissions data used for the alternative, component-level estimates of source-specific methane emissions. Emission factor values are reported when applicable. Additional details are provided in the text description.

Sector	Source	Activity Data	Emissions Data	Emission Factors
Production	Associated Gas Venting	DI (*13*)	GHGRP (county-level CH4) (*46*)	NA
	Associated Gas Flaring			
	Liquids Unloading			
	Hydrocarbon Tanks			
	Centrifugal Compressors		GHGRP (basin-level CH4) (*46*)	
	Reciprocating Compressors			
	Dehydrators			
	Flares			
	Well Testing			
	Completions			
	Workovers			
	Combustion Exhaust		GHGRP (county-level CO2) (*46*); AP-42 (*57*)	100 (4 — 660) g CH4 MMBTU-1
	Equipment Leaks		Allen (*54*)	Zavala-Araiza (*18*)
	Pneumatic Controllers	DI production (*13*); GHGRP pump counts (*46*)	Allen (*55*)	High-bleed: 2.0 (1.5 – 2.5) MT CH4 yr-1
				Low-bleed: 0.42 (0.56 – 0.70) MT CH4 yr-1
				Intermittent-bleed: 0.15 (0.09 – 0.23) MT CH4 yr-1
				Malfunctioning: 7.3 (5.5 – 9.4) MT CH4 yr-1
				7% malfunctioning devices
	Pneumatic Pumps	DI production (*13*); GHGRP controller counts (*46*)	Allen (*54*)	1.9 (1.2 – 2.9) MT CH4 yr-1
	Produced Water Tanks	DI production (*13*); (*58*)	EPA O&G Tool (*59*)	Gas wells: 50 g CH4 bbl-1
				Oil wells: 14 g CH4 bbl-1
	Abandoned Wells	DI inactive wells (*13*)	Townsend-Small (*44*); Kang (*45*)	Plugged: 1.8 x 10-5 MT CH4 well-1 yr-1
				Unplugged: 8.8 x 10-2 MT CH4 well-1 yr-1
	Offshore	NA	GHGI (*17*)	NA
Gathering	Gathering Stations	DI gas production (*13*); Marchese (*29*)	Mitchell (*25*); Marchese (*29*); Zavala-Araiza (*18*)	NA
	Gathering Blowdowns	Gathering station emissions from Marchese (*29*)	Marchese (*29*); Zavala-Araiza (*18*)	10% (0.3 - 42%) station emissions
	Gathering Pipelines	DI gas producing wells (*13*)	GHGI (*17*)	0.95 miles pipeline well-1
				0.40 MT CH4 mile-1 yr-1
Processing		Maasakkers (*53*)	Marchese (*29*); Zavala-Araiza (*18*)	NA
Transmission & Storage	T&S Stations	Maasakkers (*53*)	GHGI (*17*); Zimmerle (*26*)	200 MT CH4 station-1 yr-1
	T&S Station Super-emitters		Zimmerle (*26*)	
	Transmission Pipelines		GHGI (*17*)	
	LNG Import and Export Terminals		GHGI (*17*)	
Local Distribution		Maasakkers (*53*)	GHGI (*17*); Lamb (*50*)	NA
Petroleum Transportation & Refining		Maasakkers (*53*)	GHGI (*17*); GHGRP (*46*)	NA
Other fossil and biogenic		Maasakkers (*53*)	GHGI (*17*)	NA

Table S9. Summary of liquids unloading activity in the six areas sampled to create the datasets of production site emissions used in this work's BU estimate. The expected percentage of gas wells in each area having emissions from liquids unloading at any one time were estimated using county-level 2015 data from GHGRP reporters, which account for 44-99% of gas wells in counties with site-level measurements (*60*). Unloading events were assumed to occur over 1.4 hours for non-plunger lift wells and 0.4 hours for plunger lift wells (*42*). Gas well counts from Drillinginfo were used extrapolate reported unloading events to non-reporters (*15*).

Production area	**Basin ID**	**% of wells reporting in the sampled area**	**% of wells unloading at any one time**
Fayetteville (6 counties)	345	99%	3.4%
SW PA + WV (5 counties)	160A	44%	0.14%
Barnett (9 counties)	350/415/420	82%	0.051%
Green River (Sublette County, WY)	535	98%	0.062%
Denver-Julesburg (Weld County, CO)	540	84%	0.078%
Uinta (Uintah County, UT)	575	92%	0.12%

Table S10. Comparison of the radiative forcing due to supply chain CH_4 losses across the natural gas fuel cycle (production through consumption) relative to the radiative forcing from the combustion CO_2 produced when natural gas is consumed.

Supply chain CH_4 loss rate (CH_4 emitted per CH_4 consumed)	Ratio of RF from CH4 to RF from CO2 (20-yr)	Ratio of RF from CH4 to RF from CO2 (100-yr)
0%	0%	0%
0.5%	16%	5%
1%	32%	11%
2%	63%	21%
3%	95%	32%
4%	130%	42%

Table S11. Parameters that describe the EFs for each gas production cohort, where $EF = e^{\mu+\frac{1}{2}\sigma^2}$. 95% confidence intervals are shown in parentheses.

Gas production cohort	Basin	Number of samples	EF (in kg/h)	μ	σ
>0-10 Mcf/d*	All 6 basins combined	27	1.2 (0.56, 3.2)	-0.91 (-1.5, -0.34)	1.5 (1.1, 1.9)
10-5,000 Mcf/d**	Barnett	181	1.7 (0.85,3.8)	-1.8 (-2.1, -1.4)	2.2 (1.8, 2.5)
	Fayetteville	51	1.9 (0.46, 10)	-2.2 (-2.9, -1.5)	2.4 (1.8, 2.9)
	Weld	74	1.2 (0.81,1.7)	-0.62 (-0.91, -0.34)	1.2 (1.0, 1.4)
	Uinta	25	3.0 (1.7, 5.8)	0.35 (-0.13, 0.83)	1.2 (0.88, 1.6)
	Upper Green River	45	1.9 (1.4, 2.8)	0.17 (-0.12, 0.46)	0.99 (0.78, 1.2)
>5,000 Mcf/d***	All basins combined	15	16 (4.6, 32)	-	-

*The minimum production of sites sampled in the measurement dataset is 0.68 Mcf/d. In the absence of additional data, we estimate emissions for sites with lower production (i.e., >0-0.68 Mcf/d) using the EF generated for sites with 0.68-10 Mcf/d (1.2 kg/h, which is 30% higher than the value used in the main method for sites with gas production <0.68 Mcf/d).
**We exclude SWPA from this cohort because a robust lognormal distribution could not be generated with only 15 measurements.
***N for this cohort is <25, therefore we do not generate a lognormal distribution and instead sample from the raw measurements (with replacement) to determine EFs.

Table S12. Comparison of emissions from U.S. natural gas production sites using the alternative method presented in Section S1.9 to the results of the main method (Section S1.2). Alternative/Main is the ratio of emissions estimated by the alternative method to emissions estimated by the main method.

Production cohort (Mcf/d)		>0—0.68	0.68—10	10—5,000	>5,000	All
Total emissions (Gg/y)	Main	360	1,400	4,600	340	6,700
	Alternative	550	2,000	5,500	330	8,400
Percent of emissions	Main	5.4%	21%	69%	5.1%	100%
	Alternative	6.5%	24%	66%	3.9%	100%
Ratio of Alternative/Main Emissions		1.5	1.4	1.2	0.97	1.3