https://doi.org/10.5194/essd-2024-72
Preprint. Discussion started: 2 April 2024

Open Access Earth System Science Data Discussions

Constructing a measurement-based spatially explicit inventory of US oil and gas methane emissions

Mark Omara[1,2], Anthony Himmelberger[2], Katlyn MacKay[1], James P. Williams[1], Joshua Benmergui[1,2,3], Maryann Sargent[3], Steven C. Wofsy[3], Ritesh Gautam[1,2]

[1]Environmental Defense Fund, New York, NY, USA 10010
[2]MethaneSAT, LLC, Austin, TX, USA 78701
[3]School of Engineering and Applied Science/Department of Earth and Planetary Science, Harvard University, Cambridge, MA, USA 02138

Correspondence to: Mark Omara (momara@edf.org) and Ritesh Gautam (rgautam@edf.org)

Abstract

Accurate and comprehensive quantification of oil and gas methane emissions is pivotal in informing effective methane mitigation policies, while also supporting the assessment and tracking of progress towards emissions reduction targets set by governments and industry. While national bottom-up source-level inventories are useful for understanding the sources of methane emissions, they are often unrepresentative across spatial scales, and their reliance on generic emission factors produces underestimations when compared with measurement-based inventories. Here, we compile and analyze previously reported ground-based facility-level methane emissions measurements in the major US oil and gas producing basins and develop representative methane emission profiles for key facility categories in the US oil and gas supply chain, including well sites, natural gas compressor stations, processing plants, crude oil refineries, and pipelines. We then integrate these emissions data with comprehensive spatial data on national oil and gas activity to estimate each facility's mean total methane emissions and uncertainties, from which we develop a mean estimate of national methane emissions, resolved at 0.1° × 0.1° spatial scales (~10 km × 10 km). From this measurement-based methane emissions inventory (EI-ME), we estimate total US national oil/gas methane emissions of 15.7 Tg (95% confidence interval of 14-18 Tg) in 2021 which is 2.5 times greater than the EPA Greenhouse Gas Inventory. Our estimate represents a mean gas production-normalized methane loss rate of 2.6%, consistent with recent satellite-based estimates. We find significant variability in both the magnitude and spatial distribution of basin-level methane emissions, ranging from <1% methane loss rates in the gas-dominant Appalachian and Haynesville regions to >3-6% in oil-dominant basins, including the Permian, Bakken, and the Uinta. Additionally, we present and compare novel comprehensive wide-area airborne remote sensing data and results of total area methane emissions and the relative contributions of diffuse and concentrated methane point sources as quantified using MethaneAIR in sub-regions of the Permian and Uinta basins that together indicate diffuse area sources accounting for the majority of total regional oil and gas emissions. Our assessment offers key insights into plausible underlying drivers of basin-to-basin variabilities in oil and gas methane emissions, emphasizing the importance of integrating measurement-based data in developing high-resolution, spatially explicit methane inventories in support of accurate methane assessment,

https://doi.org/10.5194/essd-2024-72
Preprint. Discussion started: 2 April 2024

Open Access Earth System Science Data Discussions

attribution, and mitigation. The high-resolution spatially explicit EI-ME inventory is publicly available at https://doi.org/10.5281/zenodo.10734300 (Omara et al. 2024)

1. Introduction

Accurate characterization of oil and gas methane emissions across spatial scales – from the facility-level to the basin- and national-level – is an essential component of methane reduction programs, integral to mitigating the near-term catastrophic impacts of human-induced global warming (IPCC, 2021). As governments, industry, and various stakeholders publicly commit to cut their methane emissions footprint (OGCI, 2021; GMP, 2021), accurate methane inventories will play a crucial role in the development and implementation of effective methane reduction approaches as well as tracking progress toward emission reduction targets.

At the national level, methane inventories are typically developed using "bottom-up" methods when reported as part of the UNFCCC Greenhouse Gas Inventory (UNFCCC, 2023). Bottom-up methane inventories are developed by applying generic, or in some cases, empirically determined, component- or source-level emission factors to national oil and gas activity data (EPA, 2022). While these inventories are useful as first-order estimates of the emission sources, they often lack the accuracy needed to characterize methane emissions, their sources, and their trends over time at the facility-scale to the basin-level.

In addition, scores of recent studies at specific oil and gas basins (Zhang et al., 2020), countries (Alvarez et al., 2018; Shen et al., 2021; Zavala-Araiza et al., 2021; Johnson et al., 2023), and globally (Shen et al., 2023) have consistently found an underestimation in bottom-up inventories when compared to measurement-based inventories, pointing to a need for improvements in the bottom-up methane inventory methodologies. Furthermore, satellite-based quantification of regional, national, and global methane emissions has emerged as crucial tools for assessing the accuracies of methane inventories (Jacob et al., 2022; Shen et al., 2023). However, when Bayesian inversion models are used for methane flux quantification, spatially explicit methane inventories are needed as a priori information (Shen et al., 2021, 2023). Past efforts have produced such a priori information by spatially disaggregating methane emissions inventories reported to the UNFCCC (Maasakkers et al., 2023; Scarpelli et al., 2022; EDGAR, 2023) which, as noted above, can have large underestimation and uncertainties in both the magnitude and spatial distributions of oil and gas methane emissions.

In this work, we utilize previous peer-reviewed facility-level measurement data for methane emissions at oil and gas sites in the major US oil and gas production basins to develop an improved assessment of national, basin-level, and facility-level methane emissions based on oil and gas activity in 2021. Our contributions are three-fold: First, we develop statistically robust facility-level methane emission models based on measurement data and use these models to estimate national methane emissions, on both an absolute basis (Tg/year) and production-normalized basis (% emitted relative to methane production). Second, we extend this approach to assess variability and underlying drivers of oil and gas methane emissions and methane loss rates across the major US oil and gas basins. As part of this assessment, we present and compare the quantification of total area methane emissions and the relative contributions of diffuse and large concentrated methane sources in the Permian and Uinta Basins, based on new remote sensing measurements by MethaneAIR (Staebell et al., 2021; Chulakadabba et al., 2023; Chan Miller et al., 2023), an

https://doi.org/10.5194/essd-2024-72
Preprint. Discussion started: 2 April 2024

Open Access Earth System Science Data Discussions

airborne precursor to MethaneSAT (www.methanesat.org). Finally, we construct a high-resolution spatially explicit oil and gas methane emissions inventory for 2021, aggregated at 0.1º × 0.1º (~10 km × 10 km) spatial scales, and use these results to characterize the spatial patterns in national emissions.

2. Methods

2.1. Oil and gas activity data

We follow the procedure developed by Omara et al. (2022) to assess the total number and site-level production characteristics of actively producing onshore oil and gas well sites in the US in 2021. Briefly, we use the monthly well-level oil and gas production data as reported by Enverus Prism (Enverus, 2023), which aggregates public and proprietary data on monthly well-level production. For each actively producing well, we derive average well-level oil (barrels per day, bpd), gas (1 thousand cubic feet per day, Mcfd), and combined oil and gas (barrels of oil equivalent per day; 1 boed = 6 Mcfd gas) production rates based on the reported number of production days, and assuming 365 calendar days in the year if production days were not reported. We filtered the data for only the actively producing wells (n = 824,003) and used geospatial clustering approaches, described in detail in Omara et al. (2022), to derive well site attributes (e.g., number of wells per site, site-level oil, gas, and boed production). Based on this analysis, we estimate a total of 660,149 actively producing onshore oil and gas well sites in the US in 2021 (Table 1), indicating an average of 1.2 wellheads per well site. Finally, we differentiate between low-production (≤ 15 boed) and non-low production (> 15 boed) oil and gas well sites based on their average site-level boed production rates in 2021. Our assessment indicates that low production well sites accounted for 82% of the total number of US onshore actively producing well sites in 2021 (Table 1).

We estimate the total number of operational gathering and transmission compressor stations, natural gas processing plants, and crude oil refineries, based on spatial data reported by Enverus Prism (Enverus, 2023), supplemented with additional spatial data from the Oil and Gas Infrastructure Mapping (OGIM) database (Omara et al., 2023), which consolidates public-domain data on oil and gas infrastructure locations and facility characteristics. For gathering and transmission pipelines, we estimate total pipeline miles based on the Enverus Prism (Enverus, 2023). In addition, we assess methane emissions associated with gas flaring activity, leveraging the natural gas flaring detections dataset based on VIIRS (visible infrared imaging radiometer suite) instruments onboard the Suomi National Polar-orbiting Partnership (NPP) and NOAA-20 satellites to identify the locations of gas flaring facilities or clusters of facilities and associated gas flared volumes (Elvidge et al., 2015). Table 1 shows the summary statistics for the oil and gas activity data used in this study.

2.2. Facility-level measurement-based methane emissions data

We begin by performing a comprehensive data review and assessment of previously published facility-level methane emissions measurements, focusing on ground-based facility-level measurement studies that report total facility-level methane emissions quantification for well sites, natural gas gathering and boosting compressor stations, natural gas transmission compressor stations, and natural gas processing plants. These ground-based measurement approaches include the dual tracer downwind mobile measurements (Mitchell et al., 2015; Omara et al., 2016, 2018),

https://doi.org/10.5194/essd-2024-72
Preprint. Discussion started: 2 April 2024

Open Access Earth System Science Data Discussions

Table 1. Oil and gas activity data and estimates of total methane emissions

Facility category	**Facility sub-category**	**Units**	**Activity data**	**Measurement-based data sources (sample size)[a]**	**Estimated total methane emissions, 2021 (Tg, 95% CI)**	**EPA GHGI (2022) (Tg)[b]**
Well sites	Low production	# of well sites	541,987	*n* = 1,153, see footnote for study references	4.3 (2.9-6.0)	3.4
	Non-low production	# of well sites	118,162		5.1 (3.6-7.4)	
Natural gas compressor stations	Gathering and boosting stations	# of stations	4,651	*n* = 116 (Mitchell et al. (2015)), *n* = 180 (Zimmerle et al. (2020))	1.6 (0.9-3.0)	1.4
	Transmission stations	# of stations	2,107	*n* = 47 (Subramanian et al. (2015))	1.7 (0.7-4.5)	1.6
Natural gas processing plants	--	# of plants	908	*n* = 16 (Mitchell et al. (2015))	1.6 (0.7-3.7)	0.51
Crude oil refineries[b]	--	# of refineries	143	*n* = 28 (see footnote)	0.14(0.1-0.18)	0.03
Pipelines	Gathering pipelines	Pipeline miles	367,717	EPA Greenhouse Gas Inventory (EPA, 2022)	0.13 (0.12-0.14)	0.13
	Transmission pipelines	Pipeline miles	552,150		0.47 (0.46-0.48)	0.17
Natural gas flaring detections	# of flaring detections	# of detections	3,153	*n* = 3,153; Elvidge et al. (2015)	0.56 (0.55-0.57)	-
	Estimated gas flared volumes	MMcf/year	344,217	Elvidge et al. (2015)		
				Total estimated methane emissions	15.7 (14.1-18.0)	8.3 (7.0-9.6)

a Measurements at well sites include 1,153 facility-level measurements from nine studies in eight basins or production regions in the US. Studies include Brantley et al. (2014), Robertson et al. (2017), Robertson et al. (2020), Omara et al. (2016), Omara et al. (2018), Caulton et al. (2019), Rella et al. (2015), Lan et al. (2015), and Yacovitch et al. (2015). For crude oil refineries, available facility-level measurements are based on aerial remote sensing quantification (Duren et al., 2019; Lavoie et al., 2017).
b EPA GHGI total includes 0.5 Tg methane from natural gas distribution, LNG storage, and other sources not shown in this table.

https://doi.org/10.5194/essd-2024-72
Preprint. Discussion started: 2 April 2024

the EPA Other Test Method (OTM-33A) downwind stationary measurements (Brantley et al., 2014; Robertson et al., 2017, 2020), and downwind mobile measurements with Gaussian plume transport modelling (Caulton et al., 2019; Omara et al., 2018). Omara et al. (2018) provides a detailed overview of these ground-based measurement methods. Other recent published studies that used chamber flux quantification approaches and reported only wellhead methane emissions quantification (e.g., wellhead methane emissions in Deighton et al. (2020) and Riddick et al. (2019)) are not included, as unquantified methane sources (e.g., from separators, tanks, pneumatic devices, etc) likely lead to a low bias in facility-level total methane emissions. However, we use the total facility-level methane emissions data reported by Zimmerle et al. (2021) for natural gas gathering and boosting stations, based on aggregation of each facility's onsite component-level measurements performed using high flow sampler following leak detection with an infrared camera. We acknowledge possible low bias in this dataset given the limitations of facility-level measurements using high-flow samplers, including inability to access all methane emitting sources and/or to quantify large emission sources beyond the high-flow sampler capacity (Zimmerle et al., 2021). Finally, given their large size and difficulty of quantifying facility-wide emissions with ground-based measurement approaches, we use available measurement-based methane emissions data for crude oil refineries based on aerial remote sensing methods (Lavoie et al., 2017; Duren et al., 2019).

For non-low production well sites, we use previously published facility-level measurement data collected in eight US basins, including the Barnett (*n* = 254; Brantley et al. 2014; Lan et al., 2015; Rella et al., 2015; Yacovitch et al., 2015), Denver-Julesburg (*n* = 46; Robertson et al., 2017; Brantley et al., 2014; Omara et al., 2018), Eagle Ford (*n* = 3; Brantley et al., 2014); Fayetteville (*n* = 47; Robertson et al., 2017), Marcellus Shale (*n* = 572; Omara et al., 2016; Omara et al., 2018; Caulton et al., 2019), Permian (*n* = 72; Robertson et al., 2020), Uinta (*n* = 31; Robertson et al., 2017; Omara et al., 2018), and Upper Green River (*n* = 129; Brantley et al., 2014; Robertson et al., 2017). The consolidated site-level measurement data (*n* = 1,153) included only data from studies that reported total facility-level emissions quantification in addition to the production characteristics (i.e., gas and/or oil production rates). We use each study's reported facility-level methane loss rate, computed as methane emissions relative to methane production at each facility, in our modeling of methane emissions. Where methane loss rates were not reported, we compute the percent methane loss rates as follows, based on the reported average gas production rate at the time of measurement:

$$methane\ loss\ rate = CH_4[\frac{kg}{h}] \times \frac{1}{Gas\ [Mcfd]} \times \frac{1\ Mcf}{19.2\ [kg\ CH_4]} \times \frac{1}{\sigma_{CH4}} \times \frac{24h}{1d}$$

where σ_{CH4} is the assumed methane fraction in the produced natural gas (we assume an average of 80% methane content in the produced natural gas).

For low-production well sites (≤15 boed), we use the same facility-level methane emissions data and emissions assessment methods as described in detail in Omara et al. (2022). For non-low production well sites (>15 boed), we use the reported site-level measurement data described above and shown in Fig. 1a, which indicates an inverse relationship between methane loss rates and facility-level gas production rates (Omara et al., 2018). The measurement-based data includes measurements that were reported as zeros or below the method detection limits of

https://doi.org/10.5194/essd-2024-72
Preprint. Discussion started: 2 April 2024

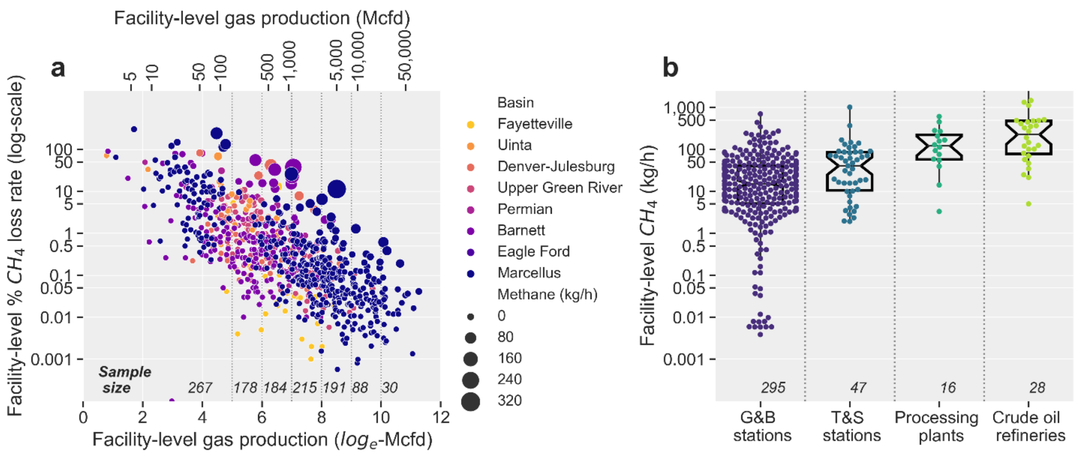


Figure 1. Previously reported facility-level measurement-based methane emissions data. **a.** Facility-level methane emissions data (percent methane loss rate) as functions of gas production rates (n = 1,153). The bottom x-axis shows the log-normalized gas production rates, with dashed vertical lines delineating the seven production cohorts used to model total methane emissions. Sample sizes for each production cohort are shown at the bottom of the Fig. panel above the x-axis tick labels. The top x-axis shows the same production data in Mcfd. Each point is color-coded by basin and sized by the quantified methane emission rate in kg/h. Not shown are measurements that were reported as below the method detection limits. **b.** Absolute methane emission rate data (kg/h) for gathering and boosting (G&B) compressor stations (n = 295), transmission and storage (T&S) compressor stations (n = 47), natural gas processing plants (n = 16), and crude oil refineries (n = 28). The swarm plots show individual facility-level measurements, while the notched box plots show the distribution, where the boxes represent the 25th and 75th percentiles and the whiskers extend to 1.5× the interquartile range.

0.036 kg/h (Robertson et al. 2017; Brantley et al., 2014) for the OTM-33A methods and 0.01 kg/h (Omara et al., 2016) for the dual tracer flux quantification.

To assess facility representativeness, we compare and find reasonable overlap in the distribution of facility-level gas production rates for the measured non-low production sites with the distribution for the national population of non-low production sites (Supplementary Fig. 3). We account for potential bias in oversampling of the higher producing well sites by using the gas production-normalized methane loss rate models in our estimates of total methane emissions for the non-low production well sites.

Figure 1b shows the previously reported facility-level measurements at midstream/downstream facilities, including natural gas gathering and boosting compressor stations, transmission compressor stations, processing plants, and crude oil refineries. In all cases, we use the average facility-level methane emissions data as reported, acknowledging inherent limitations in these measurement approaches (e.g., pseudo-random facility-level measurements with small sample sizes in ground-based approaches, or difficulty quantifying large emitters using high flow samplers in component-level measurements, etc.) likely increases the uncertainties in our estimates of total, regional, and national methane emissions.

https://doi.org/10.5194/essd-2024-72
Preprint. Discussion started: 2 April 2024

2.3. Facility-level methane emissions model development and estimation of total national methane emissions

For non-low production well sites (average facility-level production rates > 15 boed), we group the facility-level methane loss rates into seven log-normalized gas production (Mcfd) cohorts, shown in Fig. 1a and delineated by dashed vertical lines (log-Mcfd ≤ 5, 5-6, 6-7, 7-8, 8-9, 9-10 and log-Mcfd > 10). We use one log-e space (between log-Mcfd ≤ 5 to log-Mcfd > 10) to develop these production cohorts, given the inverse relationship between facility-level methane loss rate and production rates, and selected to provide sufficient sample sizes for emissions distribution modeling for each production cohort (Fig. 1a). For each cohort, we simulate the frequency of finding a site emitting below the method detection limits (reported as zeros or below the method detection limit) through a random bootstrapping procedure, repeated 10^4 times, with replacement. From this simulation, we develop a frequency distribution for the sites below the detection limits, which averaged roughly 20% to 30% for all of the cohorts, with the exception of the last production cohort (>10 Mcfd), where the frequency drops to roughly 10 to 20% (Supplementary Fig. 1).

For the measured oil and gas production well sites with emissions above the method detection limits, we begin by applying a log-transformation to the reported facility-level methane loss rates in each cohort and assessing the goodness of fit for the empirical distributions to a lognormal distribution, using the Kolmogorov-Smirnov test with significance established at $p < 0.05$. For all seven cohorts, we find that the lognormal distribution assumption is valid, with $p > 0.05$ (Supplementary Fig. 2). For each cohort's empirical distribution, we assume a univariate normal likelihood with mean μ and standard deviation σ and use Bayesian models with weakly informative priors to estimate μ and σ, for example, as μ~Normal(-10,5) and σ~HalfNormal(3) for the first cohort of non-low production sites. For Bayesian inference, we draw 5,000 posterior samples from the posterior distribution using the PyMC3 (Salvatier et al., 2016) implementation of the No-U-Turn Sampler (NUTS) algorithm (Hoffman and Gelman, 2014) from which we estimate μ and σ, as well as the 94% highest posterior density intervals (HPD). Note that the mean facility-level methane loss rate for each cohort can be computed as $exp(\mu+0.5\sigma^2)$. From the posterior results, we generate 5,000 predictions of the facility-level methane loss rate for each measured well site within each production cohort. Fig. 2 shows the cumulative probability distribution function for the observed data and 500 random samples from the model predictions.

We follow the above Bayesian modeling procedure to develop predictions of emission distributions (kg/h/facility), conditional on empirical data, for the gathering and boosting compressor stations, transmission compressor stations, natural gas processing plants, and crude oil refineries. We then proceed as follows to estimate methane emissions for the total population of facilities: for every facility in each facility category and/or production cohort, we randomly draw an emission rate from the modeled posterior predictions (Fig. 2). For non-low production oil and gas production facilities, we randomly draw a methane loss rate estimate which is then multiplied by the facility's average methane production rate to estimate methane emissions in kg/h. As some facilities can have emissions below the method detection limits, we decrement the total estimated emission rate based on a randomly sampled frequency of BDL sites (f_{BDL}). We repeat this procedure 500 times and develop a methane emission distribution for the total methane emissions for each facility category or production cohort.

https://doi.org/10.5194/essd-2024-72
Preprint. Discussion started: 2 April 2024

Open Access Earth System Science Data Discussions

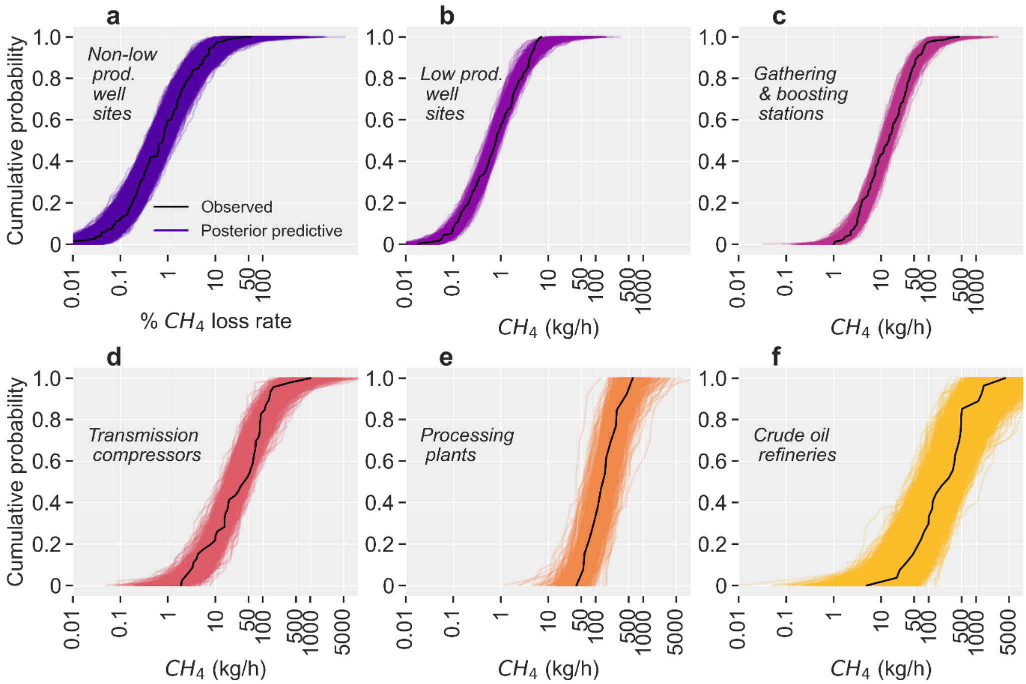


Figure 2. Empirical cumulative distribution functions of observed data and model predictions. Empirical CDFs are shown in solid black lines while thin colored lines show 500 random samples drawn from the model predictions. Sample sizes and data sources for empirical data are shown in Table 1. **a.** Non-low production well sites (modeled as facility-level methane loss rates), the Fig. shows the CDF for the 5 < log-Mcfd < 6 production cohort. Supplementary Fig. 2 shows the CDFs for all seven non-low production cohorts in Fig. 1. **b.** Low-production well sites (kg/h/site), **c.** Gathering and boosting compressor stations, **d.** Transmission compressor stations, **e.** Natural gas processing plants, **f.** Crude oil refineries.

Given the scarcity of facility-level measurements for gathering and transmission pipelines, we use the emission factors estimated by the US EPA Greenhouse Gas Emission Inventory (EPA, 2022; 285 kg methane/mile/year and 582 kg methane/mile/year, respectively) and assume normal distributions of emission factors with 50% uncertainty.

We also estimate the methane emissions associated with gas flaring activities using location-specific gas flaring data from the VIIRS instrument (Elvidge et al., 2015) and apply average effective methane destruction removal efficiency of 91% (Plant et al., 2022; 95% confidence interval of ~90—92%).

Finally, we combine the emission distributions for all facility categories and sources using Monte Carlo methods to estimate the mean total national methane emissions and the 95% confidence interval based on the 2.5th and the 97.5th percentiles of the modeled distributions. Fig. 3 shows a general schematic of the emissions model development and estimation of total methane emissions.

https://doi.org/10.5194/essd-2024-72
Preprint. Discussion started: 2 April 2024

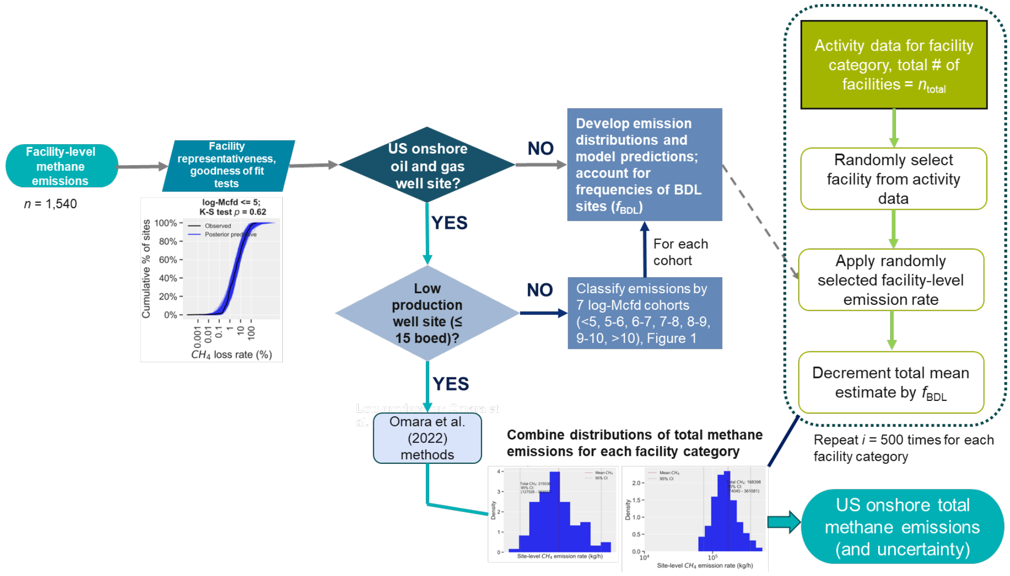


Figure 3. General schematic for model development and estimation of total methane emissions, given activity data for each facility category.

2.4. Spatial allocation of estimated methane emissions

For each facility with known location (latitude, longitude), our assessment includes 500 different estimates of likely facility-level methane emission rates (in kg/h), from which we derive 500 different estimates of total national methane emissions. We use a search algorithm to identify a random sample of the facility-level emission rate distribution that most closely matches the computed mean estimate for the population of facilities. We use a similar approach to select a random sample of the facility-level emissions distributions representing uncertainties in the total emission estimates (i.e., the distribution that most closely matches the lower bound and upper bound of the 95th percent confidence intervals on the total estimated methane emissions). We then aggregate total mean methane emissions (and associated upper and lower bound estimates) on regular grids of 0.1º × 0.1º decimal degrees (~10 km × 10 km), to produce spatially explicit oil and gas methane emission inventories and related uncertainties.

2.5. Model uncertainties and limitations

In our modeling, we use the average facility-level emissions data as is, while assuming facility emissions arise from an underlying methane emissions distribution that is statistically described by lognormal distributions. The implementation of these probabilistic models produces emission distribution models (Fig. 2) that account for uncertainties in each facility's measured average methane emission rate and facility-to-facility variability in methane emissions within and across multiple oil and gas production regions. The 95% confidence intervals obtained through the Monte Carlo methods above reflect these uncertainties, as well as the model uncertainties in predictions of

https://doi.org/10.5194/essd-2024-72
Preprint. Discussion started: 2 April 2024

emissions distributions, given the limited sample sizes used herein. Additional uncertainties that are difficult to quantify include uncertainties in the oil and gas activity data and uncertainties in the potential impacts of recently promulgated federal/state-specific regulations or operator-specific practices regarding regular facility-level methane emissions monitoring and repair. In addition, due to data limitations, our national estimates do not include methane emissions from downstream natural gas distribution, LNG storage, post-meter emissions, and abandoned oil and gas wells. The EPA GHGI (EPA, 2022) estimates these sources account for ~0.5 to 1 Tg a year of total methane emissions, the vast majority of these would be distributed in urban locations outside of major oil and gas production regions.

3. Results and discussion

3.1 Total national oil and gas supply chain methane emissions

We estimate a measurement-based methane emission inventory (EI-ME) of total national oil and gas methane emissions for the onshore US as 15.7 Tg (95% confidence interval of 14 – 18 Tg or -10%/+15% uncertainty; Table 1; Fig. 4) for the year 2021. Our central estimate and confidence bounds are in reasonable agreement with recent measurement-based facility-level emission estimates (Alvarez et al., 2018; Rutherford et al., 2021 (production sector only)) and satellite-derived oil and gas methane emissions, including GOSAT (Lu et al., 2022, 2023) and TROPOMI (Shen et al., 2022) quantification (Fig. 3b). However, consistent with previous findings (Alvarez et al., 2018; Rutherford et al., 2021; Shen et al., 2022), our central estimate is significantly greater than inventories developed using the traditional bottom-up source-level emission factor approaches: we find a factor of 1.9× and 1.8× greater total methane emissions than is estimated by the EPA Greenhouse Gas Inventory (EPA, 2022) and EDGAR v8 (EDGAR, 2023) inventories for the year 2021. (Fig. 5a).

We attribute the largest discrepancy between our measurement-based estimates and the EPA GHGI to the estimated emissions for the oil and gas production sector, which we estimate accounts for approximately 60% of the total onshore methane emissions for a total of ~9 Tg in 2021, or roughly 2.6× greater than the EPA GHGI (Fig. 4; Table 1). These results are in reasonable agreement with previous measurement-based inventories (Alvarez et al., 2018; Rutherford et al., 2021; Omara et al., 2022) and, as has been noted elsewhere (Alvarez et al., 2018; Rutherford et al., 2021; Omara et al., 2018), likely reflect the use of emission factors in the EPA GHGI that do not adequately characterize the contributions of high-emitting methane sources that have been consistently observed in measurement-based studies. Furthermore, within the oil and gas production sector, we find that the low-production well site cohort (<15 boed) accounts for roughly one-half of total production site methane emissions in 2021, consistent with recent findings based on 2019 oil and gas activity (Omara et al. 2022). As Table 1 shows, the estimated total methane emissions from the low-production well site cohort alone are ~26% more than the total methane emissions from all low-production and non-low production well sites based on the EPA GHGI.

In 2021, we estimate a national methane loss rate of 2.6% (95% CI: 2.3 – 2.9%) relative to gross natural gas production, assuming 80% methane content in natural gas. Similar results have been reported based on recent satellite-derived estimates (Shen et al., 2022 using TROPOMI and Lu et al., 2023 using GOSAT). Lu et al. (2023) reports a steadily declining national methane loss rate between 2010 (~3.7%) and 2019 (~2.5%) and attributes these trends to

https://doi.org/10.5194/essd-2024-72
Preprint. Discussion started: 2 April 2024

Open Access Earth System Science Data Discussions

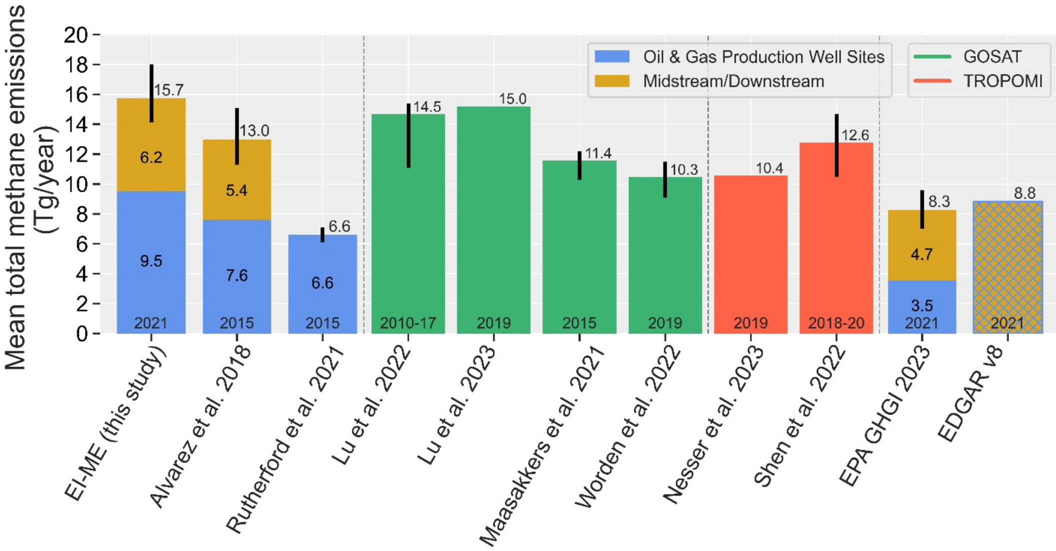


Figure 4. Comparison of this study's national estimate of total methane emissions from the oil and gas supply chain with previous measurement-based estimates. The first three bars show the oil and gas methane emissions estimated from facility-level measurements (this study, Alvarez et al. 2018) and production-sector-only methane emissions estimate by Rutherford et al. (2021) using component-level measurement data. Blue bars show the estimated emissions for the production sector, gold bars show the estimated emissions for the midstream and downstream facilities (compressor stations, processing plants, refineries, gathering and transmission pipelines). Error bars show the estimated 95% confidence bounds on the mean total methane emissions estimates.

two likely factors: (i) a slower increase/decrease in absolute methane emissions compared to the increase in methane production during this period and (ii) the impact of national regulations, such as the EPA's New Source Performance Standards, promulgated in 2012, which focused on reducing emissions from newly constructed well sites, among other requirements. As we discuss further below, we find significant variability in the total methane emissions as well as spatial distributions of the estimated emissions at regional/basin-level for the of oil and gas activity in 2021.

3.2. Variability in estimated basin-level methane emissions

Among the major oil and gas production basins, we identify the Permian, Appalachian, Anadarko, Eagle Ford, Haynesville, and the Barnett basins as the top six methane emitting basins, with estimated mean total basin-level methane emissions ranging from approximately 70 t/h to 340 t/h (Table 2, Fig. 5). These six basins account for 72% of onshore total combined oil and gas production (boe), and 52% of estimated total oil and gas methane emissions. Among these basins, we estimate considerable variability in gas production-normalized methane loss rates, with the lowest mean methane loss rates of <1% in the Appalachian and the Haynesville basins and the highest mean methane loss rates of 3-4% in the Permian, Anadarko, and the Barnett (Table 2). The basin-level differences in methane loss rates among basins are consistent with the GOSAT-derived estimates for 2019 (Lu et al. 2023, Table 2), except for

https://doi.org/10.5194/essd-2024-72
Preprint. Discussion started: 2 April 2024

Open Access Earth System Science Data Discussions

Table 2. Top six methane emitting basins' production, loss rate, and comparison with the EPA GHGI (traditional bottom-up inventory) and Lu et al. 2023[a] (satellite-derived estimates).

Basin	**Basin area (km^2)**	**Well site count (% from low-prod.)**	**Total annual (2021) gas production (Tcf[b])**	**EI-ME methane emissions, 2021 (t/hr, 95% CI) | *% from well sites***	**EPA GHGI methane emissions, 2018 (t/hr)**	**EI-ME methane loss rate, 2021 (%, 95% CI)[c]**	**EPA GHGI methane loss rate, 2018 (%)**	**Lu et al. GOSAT methane loss rate, 2019 (%)[e]**
Permian	165,325	129,364 *(78%)*	6.5	335 (274 - 428) | *69%*	98	2.6 (2.1-3.3)	1.2	2.7 (1.6-3.0)
Appalachian	415,446	167,132 *(97%)*	12.7	231 (165 - 324) | *75%*	140	0.92 (0.66-1.30)	0.67	0.45 (0.40-0.47)
Anadarko	42,479	24,180 *(64%)*	1.9	119 (93 - 166) | *55%*	32	3.2 (2.5-4.4)	0.74	3.4 (2.1-3.6)
Eagle Ford	50,179	24,377 *(54%)*	2.3	90 (73 - 119) | *75%*	27	2.0 (1.7-2.7)	0.56	1.1 (0.78-1.3)
Haynesville	28,922	23,895 *(78%)*	4.8	75 (59 - 95) | *69%*	29	0.80 (0.63-1.0)	0.46	1.2 (0.89-1.2)
Barnett	68,146	25,760 *(79%)*	0.92	74 (57 - 96) | *68%*	33	4.1 (3.1-5.3)	1.4	4.0 (3.3-4.1)

Note the differences in the temporal resolution of the studies used for the comparison, specifically the EPA GHGI basin-level estimates are based on Maasakkers et al. (2023) for the year 2018 and Lu et al. (2023) GOSAT estimates are for the year 2019. **a** Loss rates calculated assuming 90% methane content in each basin, for ease of comparison with Lu et al. **b** Tcf = Trillion cubic feet **c.** Methane loss rate calculated using 2021 production data from Enverus Prism **d.** Loss rate calculated using 2019 production data from Enverus Prism.

the Appalachian and the Eagle Ford Basins, where this study's estimates are roughly 2× greater (Table 2). As with our findings on the comparative assessments with the EPA GHGI at the national level, our basin-level methane emissions estimates are consistently greater than the EPA GHGI estimates (Maasakkers et al., 2023) by a factor of 1.7× (Appalachian) to ~4× (Anadarko).

The confluence of various modeled factors, including the spatial density and characteristics of methane emitting oil and gas infrastructure and basin-level operational characteristics (gas-dominant versus oil-dominant, intensive flaring versus basins with negligible flaring, etc) contribute to the differences in the modeled basin-level methane emissions. In each basin, we estimate a predominant contribution of total methane emissions from well site

https://doi.org/10.5194/essd-2024-72
Preprint. Discussion started: 2 April 2024

Open Access Earth System Science Data Discussions

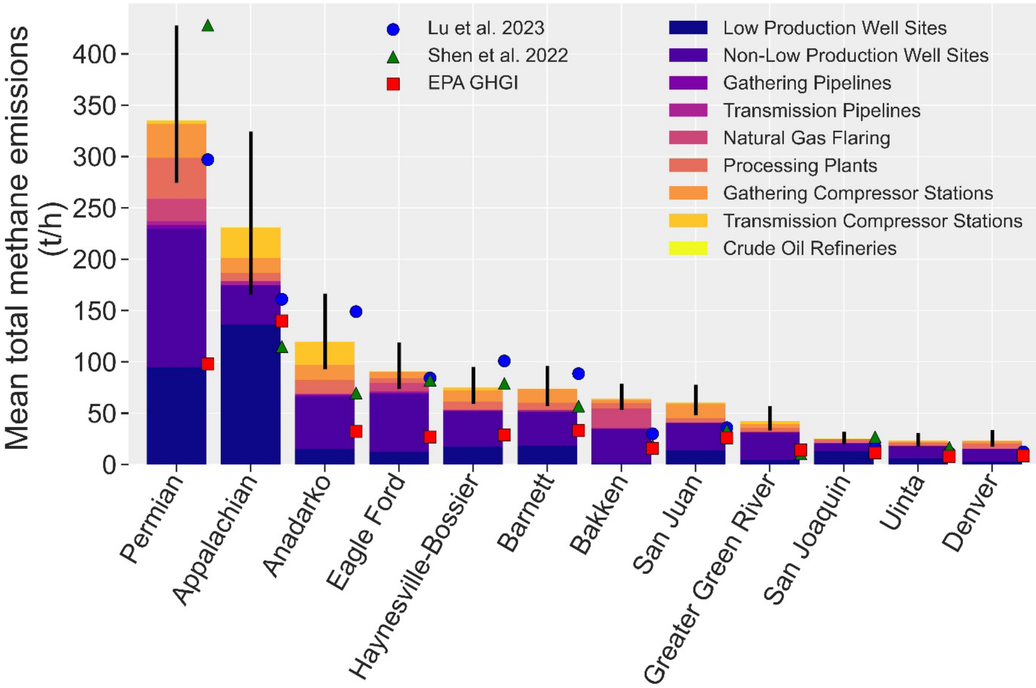


Figure 5. Basin-level differences in modeled mean total methane emissions and comparison with the EPA GHGI (Maasakkers et al., 2023), TROPOMI-derived estimates (Shen et al., 2022) and GOSAT-derived estimates (Lu et al., 2023).

infrastructure, ranging from 55% to 75% of the total basin-level methane emissions (Table 2; Fig. 5). Well site infrastructure characteristics vary significantly among basins, for example, the Appalachian Basin is characterized by a large population of old, leak-prone low-producing gas well sites (Omara et al., 2016; Deighton et al., 2020; Riddick et al., 2019) even as more than 95% of the gas produced comes from ~3% of well sites that are unconventional non-low production well sites (Enverus, 2023). This contrasts with the San Joaquin Basin, where well site infrastructure is dominated by low-producing oil pump jacks with limited onsite processing equipment, which in turn contrasts with the oil-dominant Bakken, dominated by high-producing horizontally-drilled well site facilities, typically with multiple wellheads and auxiliary processing equipment including separators, storage tanks, and flare stacks. Such basin-level oil and gas infrastructure characteristics likely contribute to the modeled emission differences, given that the empirical data synthesized herein reveal the weakly correlated methane emission profiles with well site production characteristics (loss rates, Fig. 1a) and infrastructure category (absolute emissions, Fig. 1b, Fig. 2).

Furthermore, the magnitude of modeled methane emissions varies by basin-level operational characteristics. For example, the Permian Basin with significant new oil and gas development stands in contrast to the relatively

https://doi.org/10.5194/essd-2024-72
Preprint. Discussion started: 2 April 2024

mature basins such as the Barnett or Uinta with steadily declining gas production and aging well site infrastructure. As Lu et al. (2023) observed, high methane loss rates tend to be associated with oil-dominant basins where production activities are focused on oil production even as substantial associated gas is co-produced along with oil (e.g., Permian, Eagle Ford, Bakken). In these basins, potentially higher methane emissions may occur due to venting and/or inefficient flaring of the co-produced gas, especially when there is insufficient infrastructure to gather, process, and transport to market the associated gas production, as has been postulated for the Permian Basin (Lyon et al., 2021; Varon et al., 2023; Lu et al., 2023).

3.3. Sub-basin methane assessment and comparison with emissions quantification using MethaneAIR

This study's EI-ME inventory provides methane emission estimates at geolocated oil and gas facilities, making it possible to develop aggregate methane emissions estimates across sub-basin to basin and national levels. We compare our sub-basin estimates for the Delaware portion of the Permian Basin and Uinta Basin with new remote sensing-based quantification by MethaneAIR (Staebell et al., 2021; Chulakadabba et al., 2023; Chan Miller et al., 2023), an airborne precursor to MethaneSAT satellite, which is scheduled to launch in 2024. MethaneAIR and MethaneSAT missions are managed by MethaneSAT LLC (www.methanesat.org), which is a wholly owned subsidiary of Environmental Defense Fund. Both MethaneAIR and MethaneSAT are designed to produce quantitative data on total regional methane emissions while spatially disaggregating diffuse area emissions and detecting high-emitting point sources. Detailed description of the MethaneAIR instrument technical specifications, instrument calibration, retrieval methods and point source detections and validation can be found in recent works by Staebell et al., 2021, Conway et al. (2023), Chulakadabba et al., 2023, El Abbadi et al. (2023), Miller et al. (2023), and Omara et al. (2023).

In August 2021, MethaneAIR flew a ~10,000 km^2 area in the Delaware sub-basin of the Permian Basin (research flight RF-06) and Uinta Basin (research flight RF-08) and produced quantification of total area methane emissions using a geostatistical inverse modeling (GIM) framework (based on Miller et al., 2023). The GIM framework was applied to inversion of the column mean methane dry air mole fraction retrieved using MethaneAIR measurements flying at 40,000 ft above ground aboard the NCAR GV aircraft (https://www.eol.ucar.edu/field_projects/methaneair). For MethaneAIR and MethaneSAT, the GIM framework is specialized to exploit the instrument's high spatial resolution, wide spatial coverage, and high precision, and ingests high-emitting point source detections which are quantified using the modified integrated mass enhancement method (Chulakadabba et al., 2023). As such, remote sensing measurements by MethaneAIR at 40,000 ft above ground produce a high resolution, spatially explicit quantification of the total area methane emissions as well as high-emitting methane point sources emitting above ~200 kg/h.

We compare these new MethaneAIR total area methane quantification with the EI-ME modeled results, as well as the results from previous peer-reviewed studies in overlapping domains with the Permian RF-06 and the Uinta RF-08 regions. For both regions, we find good agreement, within uncertainty bounds, of the MethaneAIR quantification with other studies (Fig. 6), with emission rates quantification that within a representative range of

https://doi.org/10.5194/essd-2024-72
Preprint. Discussion started: 2 April 2024

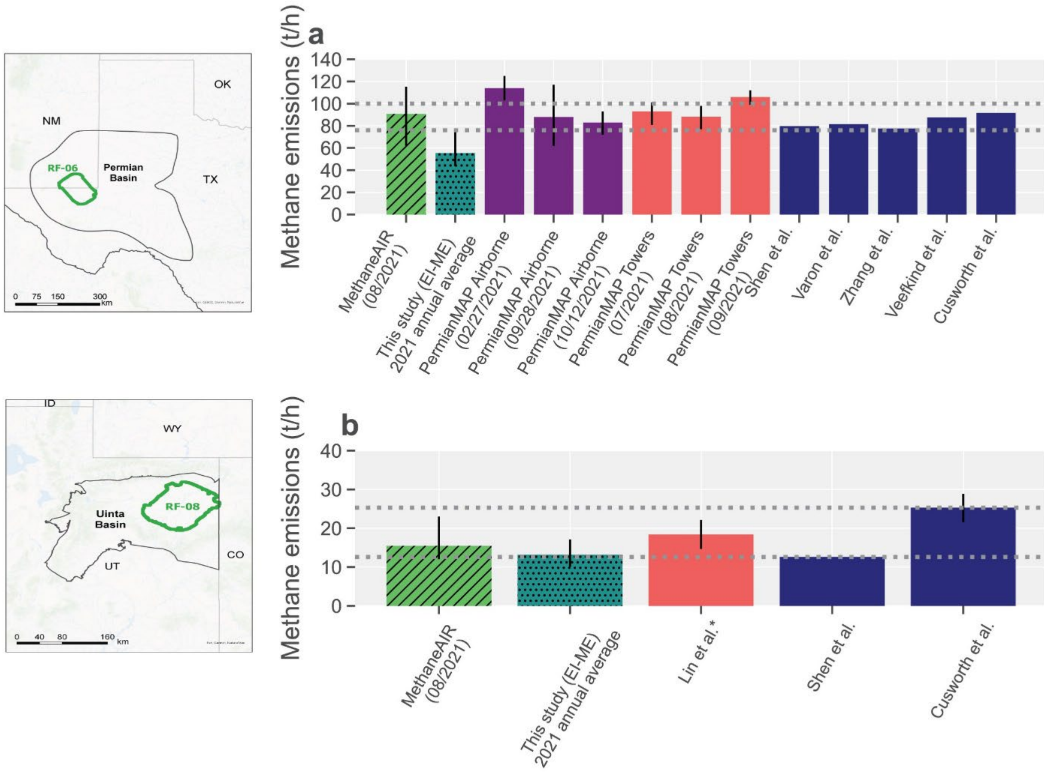


Fig. 6. Comparison of the EI-ME inventory with MethaneAIR and other peer-reviewed studies for two sub-regions of the Permian and Uinta Basins. Bars are color-coded by emission quantification method (MethaneAIR – hatched green bar; EI-ME – hatched dark green; PermianMAP airborne studies – purple; PermianMAP or tower-based study – red color; TROPOMI studies – dark blue bars). Lin et al. (2021) reports total Uinta Basin methane emissions estimates; we adjusted their estimate by the ratio of gas production in RF-08 region to total gas produced in Uintah and Duchesne counties in 2021 (RF-08 accounts for 74% of the total production in the two counties). For all other studies, we use only the reported emission estimates that overlap the MethaneAIR target boundaries. The dashed horizontal lines show a representative range of sub-basin methane emissions, computed via a bootstrapping procedure of all previously reported methane emissions (including the EI-ME results) to derive a lower bound and upper bound on the mean total methane emissions, based on the 2.5th and 97.5th percentiles of the resulting bootstrap distribution. Map credit: ESRI, 2023.

mean total sub-basin methane emissions of 80 – 100 t/h and 17 – 24 t/h for RF-06 Permian and RF-08 Uinta, respectively (horizontal dashed lines in Fig. 6).

Based on MethaneAIR quantification for these two regions, we estimate that diffuse area emissions (which are assessed using the GIM modeling framework) account for the majority of methane emissions in both sub-basins, representing 63% and 88% of the total area methane emissions in the RF-06 and RF-08 regions, respectively. The remainder (37% and 12% of the total in RF-06 and RF-08 respectively), is attributable to the quantified high-emitting methane point sources with facility-specific methane emission rates in excess of ~200 kg/h/facility. These results are in reasonable agreement with the EI-ME results –averaged over the year—for the same spatial domains,

https://doi.org/10.5194/essd-2024-72
Preprint. Discussion started: 2 April 2024

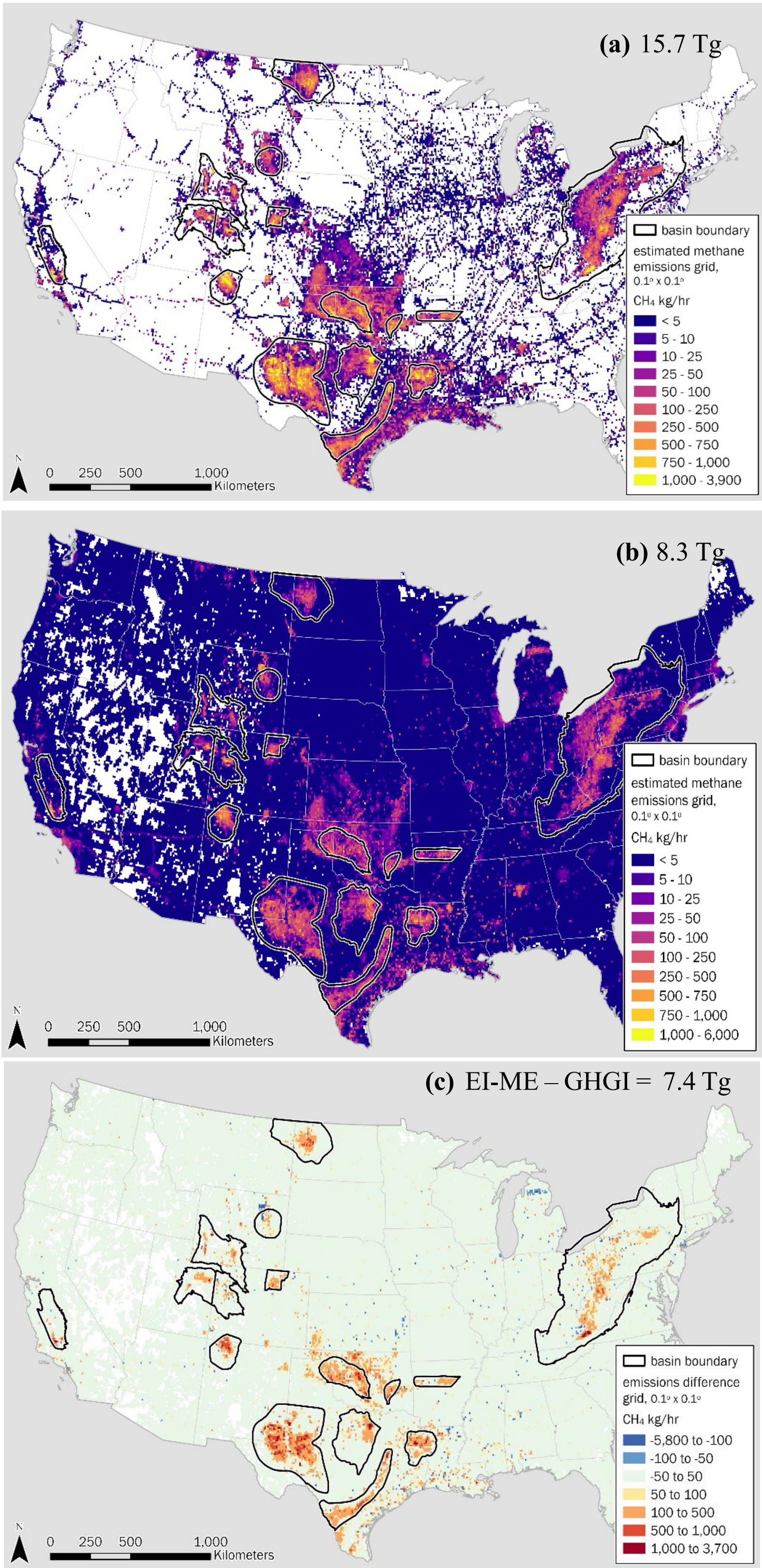
in which oil and gas methane sources with mean methane emission rates <200 kg/h account for 85% and 90% of the total estimated methane emissions for RF-06 and RF-08, respectively. Furthermore, Cusworth et al. (2021) reports similar results for the same regions overlapping these domains in the Permian and Uinta, finding that methane sources below 200 kg/h account for 70% and 88% of total area emissions, which were quantified based on area-inversion of TROPOMI satellite observations and point source detections by AVIRIS-NG in 2021 and 2020 for RF-06 and RF-08, respectively. Furthermore, for a different sub-region of the Permian Basin, Kunkel et al. (2023) observed that facility sized emission sources with rates below 280 kg/h contribute 67% of the total emission rate from all sources with rates above 10 kg/h. At the national level, Omara et al. (2023) previously showed that the large population of low producing well sites (also known as marginal wells), with population average methane emissions rates of ~1 kg/h/site, account for roughly one-half of all production site methane emissions. Taken together, these results underscore the importance of small methane emitting sources dispersed across areas that, while individually emitting at low rates, nevertheless contribute, in aggregate, a disproportionate fraction of regional total methane emissions. Williams et al. (2024) expands on these assessments, providing a detailed look at facility-level methane emissions distributions at the basin- and national-level.

3.4. Variability in estimated spatial distribution of methane emissions

Our spatially explicit EI-ME inventory suggests that basin-level differences also manifest as differences in the spatial distribution of total methane emissions. On average, we find oil and gas methane emission hotspots in every major US oil and gas production basin, including the Permian (the largest oil producing basin in the US located in western Texas and southern New Mexico), the Appalachian (Pennsylvania, Ohio, West Virginia, New York), Anadarko (Oklahoma, Texas), Eagle Ford (Texas), Bakken (North Dakota), and the Haynesville basins (Texas, Louisiana; Fig. 6)., Supplementary Fig. 4). Our analysis suggests methane emission hotspots tend to concentrate in high oil and gas production areas, for example, as evidenced by the two large hotspots in the rapidly developing, high-producing Delaware and Midland sub-basins of the Permian Basin in (Fig. 7a), consistent with spatial distributions for the satellite-observed methane emissions quantification in this region (Zhang et al., 2020; Varon et al., 2023). In addition, as with the total basin-level emissions, methane emissions spatial distributions are functions of oil and gas activity and their related facility-level emission characteristics. For example, substantial low-production oil and gas well site activity yields modeled methane emission hotspots in the southwestern tip of the Appalachian Basin (Fig. 7a), even as this region is not an oil and gas production hotspot (Supplementary Fig. 5). Furthermore, our analysis suggests spatial correlation of methane emission hotspots with intensive gas flaring activity, particularly for the oil producing basins with substantial associated gas production, including the Permian, the Eagle Ford, and the Bakken regions (Fig. 4, Supplementary Fig. 5).

We further assess variability in the spatial distribution of modeled methane loss rates, which reveals areas (25×25 km^2 grids) in each major basin where methane loss rates are <0.25-1% of methane production. These areas, in general, are characterized by significant unconventional oil and gas production, for example, in the Appalachian Basin (northeastern Pennsylvania and the tri-state corner of southern Pennsylvania, eastern Ohio, and northern West Virginia) as well as in the Permian Delaware and Midland sub-basins, and parts of the Haynesville, Eagle Ford, and

https://doi.org/10.5194/essd-2024-72
Preprint. Discussion started: 2 April 2024



https://doi.org/10.5194/essd-2024-72
Preprint. Discussion started: 2 April 2024

Open Access Earth System Science Data Discussions

Figure 7. Estimated spatial distribution of total methane emissions and comparison with the EPA GHGI estimates. **a.** This study's assessment of spatial distribution of US total oil and gas supply chain methane emissions. For visualization and comparison with the EPA GHGI inventory, the total methane emissions are gridded to 0.1 × 0.1 degrees. Major basin boundaries are outlined using black polygons. **b.** Estimated spatial distribution of total oil and gas methane emissions based on the EPA GHGI (2022; Maasakkers et al., 2023). **c.** Difference in spatially explicit methane emissions between this study's estimates and the EPA GHGI. Warmer colors indicate comparatively higher estimates from this study relative to the EPA GHGI. We acknowledge the comparison is limited by the different time periods in the two studies – 2021 in this study versus 2018 for the EPA GHGI. Nevertheless, as both studies report annual averages, it is unlikely that significant differences in aggregate spatial distribution would have occurred over the three intervening years. Map credit: ESRI, 2023. Basin boundaries based on US EIA basin boundaries data (https://www.eia.gov/maps/maps.php)

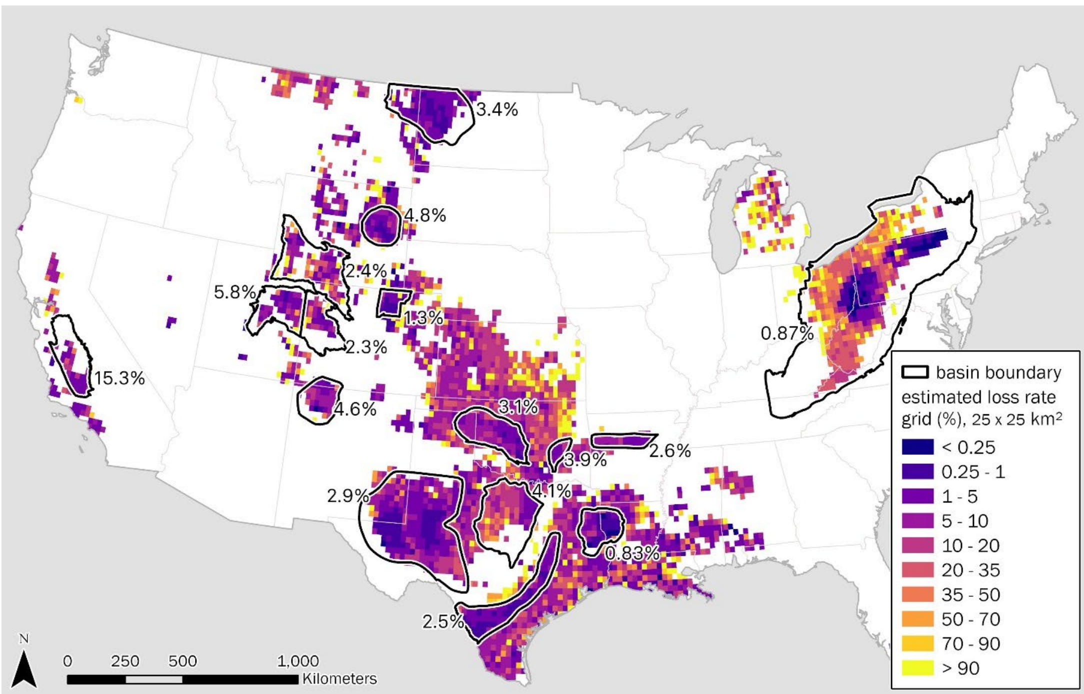


Figure 8. Estimated mean spatial distribution of production-normalized methane loss rates. For ease of visualization, we aggregate our facility-level methane inventories to a coarser spatial grid (25 × 25 km^2) and normalize each grid's total estimated methane emissions relative to total methane production to derive spatially explicit methane loss rates, assuming 80% methane content in natural gas. Major basin boundaries are outlined in black and mean basin-level methane loss rates are shown as %. Map credit: ESRI, 2023. Basin boundaries based on US EIA basin boundaries data (https://www.eia.gov/maps/maps.php).

https://doi.org/10.5194/essd-2024-72
Preprint. Discussion started: 2 April 2024

the Bakken (Fig. 8). We also estimate areas with excessive methane loss rates >10% of methane production (Fig. 8) in each major producing basin, particularly in the Appalachian Basin, in the Michigan Basin, and in the greater Anadarko area of Missouri (Fig. 8). High methane loss rates are likely linked to the predominance of old, leak-prone low-producing well sites (e.g., in parts of the Appalachian and San Joaquin basins) or may be associated with modeled midstream infrastructure emissions from sources not collocated with significant oil and gas production.

The updated 2018 gridded EPA GHGI inventory for oil and gas systems (Maasakkers et al., 2023) uses the same source of oil and gas activity data as this study (Enverus, 2023), and allocates GHGI emissions to specific emission source categories using infrastructure locations and methane emission scaling factors (e.g., scaled using well count, oil, and/or gas production for well sites depending on source category). The estimated methane emission hotspots (Fig. 7b) are in reasonable agreement with this study's estimated spatial distributions (Fig. 7a, $r = 0.64$), with notable exceptions in parts of the Michigan Basin (Michigan), the Appalachian Basin (Pennsylvania, Ohio, West Virginia) Basin, the Powder River Basin (Wyoming), the Barnett (east Texas), the Permian (west Texas), and the San Joaquin (southern California) Basins (Fig. 6b). In parts of these basins, strong methane hotspots appear in regions that likely reflect a dependence of emissions spatial allocation on spatial density of infrastructure (Supplementary Fig. 5). This differs with this study's spatial allocation which leverages not just infrastructure locations, but simultaneously integrates the empirically observed facility-level methane emissions characteristics (Fig. 1, Fig. 2), which can vary among populations of the same facility category (e.g., the distinction between emission profiles for low/non-low production well sites or among different production cohorts of the non-low production well site category).

In general, we find large differences in the magnitude of methane emissions in all of major basins shown in Fig. 7c when comparing the spatially explicit methane emissions in the GHGI and this study's estimates (Fig. 7c, Supplementary Fig. 6). We also find large differences in the spatial distributions of the methane emissions when comparing this study's spatially explicit emissions inventory with the EDGAR v8 inventory (EDGAR, 2023; Supplementary Fig. 7). Note that the EDGAR v8 total methane emissions are similar in magnitude to the EPA GHGI inventory estimates (Fig. 1), although emissions spatial allocation methods are primarily dependent on scaling by oil production characteristics, such that large methane hotspots are estimated in the oil-dominant basins of the Permian, the Bakken, and the Eagle Ford Basins (Supplementary Fig. 57).

Our results suggest both an underestimation in the magnitude of spatially explicit emissions in key US oil and gas basins as well as potentially unrepresentative spatial distributions of these emissions in the EDGAR v8 and the EPA GHGI gridded inventories. These results carry important implications for the use of traditional bottom-up inventories as a priori information in Bayesian inversions of satellite observations for methane quantification, since both the magnitude and spatial allocation of emissions could influence the posterior results from these modeling systems under certain observational data constraints, such as insufficient observational data density (Shen et al., 2022).

4. Data availability

EI-ME_v1.0 can be accessed at https://doi.org/10.5281/zenodo.10734300 (Omara et al. 2024) in an open-access GeoPackage file format.

https://doi.org/10.5194/essd-2024-72
Preprint. Discussion started: 2 April 2024

5. Code availability

Python 3.7 code used for emissions modeling, extrapolation to population of facilities, and data visualization is available from the corresponding author upon reasonable request.

7. Conclusions

Accurate and comprehensive assessment of oil and gas methane emissions is pivotal in informing effective methane mitigation policies. In this study, we develop robust statistical models based on measured facility-level methane emissions and integrate these models with comprehensive oil and gas activity data for onshore US oil and gas facilities to estimate total national oil and gas methane emissions for the year 2021. We estimate a total of 15.7 (14 – 18) Tg of oil and gas methane emissions in 2021, representing a mean methane loss rate of 2.6% of gross gas production. Our national methane emission estimate, while in reasonable agreement with previous measurement-based estimates using facility-level measurements and satellite observations, are nevertheless roughly a factor of 2× greater than official inventories from the EPA Greenhouse Gas Inventory (GHGI). This improved assessment of national methane emissions underscores the importance of integrating measurement-based data to develop robust methane emission inventories which, as we show in this work, exhibit substantial variability in both the magnitude and spatial distribution of total methane emissions across major oil and gas basins.

Further improvements to methane emission inventories are possible through greater integration of measurement-based data including remote sensing approaches that can provide comprehensive area-wide total methane emissions, quantification of high-emitting methane point sources, as well as high-resolution spatial disaggregation of total methane emissions. In this study, we present the first set of such remote sensing quantification, based on MethaneAIR measurements in sub-basins of the Permian and Uinta and demonstrate reasonable agreement with several previous peer-reviewed assessments of total area methane emissions over similar spatial domains and time periods. These comprehensive area wide assessments also enable a detailed characterization of the importance of diffuse area emissions viz-a-viz high-emitting methane point sources, revealing the relative importance of diffuse area emissions and their variability across unique US oil and gas producing basins.

The EI-ME inventory provides a detailed characterization of total methane emissions by key facility categories at the national level as well as at the regional/basin-level, thus helping provide policy-relevant information that is important in developing and tracking effective methane mitigation strategies. The quantified uncertainties in our methane emission estimates could be improved upon in future studies through additional peer-reviewed data collection efforts, which are needed to develop further insights in response to ongoing methane mitigation efforts. There is a research need to develop robust statistical methods for effective integration of lower-detection-limit ground-based facility-level methane emissions data with the growing number of airborne facility-level measurement studies, which generally have higher method detection limits. These improved integrated assessments of facility-level, regional, and national methane emission inventories, as demonstrated herein, support ongoing efforts to accurately quantify methane emissions, identify key methane sources and regions for targeted methane reductions, and track progress toward methane reduction goals.

https://doi.org/10.5194/essd-2024-72
Preprint. Discussion started: 2 April 2024

Author contributions

MO and RG conceptualized the study. MO developed the facility-level methane emission models. Formal data analysis, interpretation, and visualization were performed by MO, AH and RG, with assistance from KM and JPW. MethaneAIR GIM modeling and high emitting methane point source assessments were performed by JB, MS and SCW. MO wrote the manuscript with contributions from all co-authors.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgments

This work was made possible by support from the Bezos Earth Fund. We thank Madeleine A. O'Brien for providing the OGIM database and related geospatial data gridding script. We are grateful to Daniel Zavala-Araiza for providing comments.

References

Alvarez, R. A., Zavala-Araiza, D., Lyon, D. R., Allen, D. T., Barkley, Z. R., Brandt, A. R., Davis, K. J., Herndon, S. C., Jacob, D. J., Karion, A., Kort, E. A., Lamb, B. K., Lauvaux, T., Maasakkers, J. D., Marchese, A. J., Omara, M., Pacala, S. W., Peischl, J., Robinson, A. L., Shepson, P. B., Sweeney, C., Townsend-Small, A., Wofsy, S. C., Hamburg, S. P. Assessment of Methane Emissions from the U.S. Oil and Gas Supply Chain. Science, 361, 186–188, https://doi.org/10.1126/science.aar7204, 2018.

Brantley, H. L., Thoma, E. D., Squier, W. C., Guven, B. B., Lyon, D. Assessment of Methane Emissions from Oil and Gas Production Pads Using Mobile Measurements. Environ. Sci. Technol., 48 (24), 14508–14515, https://doi.org/10.1021/es503070q, 2014.

Caulton, D. R., Lu, J. M., Lane, H. M., Buchholz, B., Fitts, J. P., Golston, L. M., Guo, X., Li, Q., McSpiritt, J., Pan, D., Wendt, L., Bou-Zeid, E., Zondlo, M. A. Importance of Superemitter Natural Gas Well Pads in the Marcellus Shale. Environ. Sci. Technol., 53(9), 4747–4754, https://doi.org/10.1021/acs.est.8b06965, 2019.

Chan Miller, C., Roche, S., Wilzewski, J. S., Liu, X., Chance, K., Souri, A. H., Conway, E., Luo, B., Samra, J., Hawthorne, J., Sun, K., Staebell, C., Chulakadabba, A., Sargent, M., Benmergui, J. S., Franklin, J. E., Daube, B. C., Li, Y., Laughner, J. L., … Wofsy, S. C. Methane retrieval from MethaneAIR using the CO 2 Proxy Approach: A demonstration for the upcoming MethaneSAT mission [Preprint]. Gases/Remote Sensing/Data Processing and Information Retrieval, https://doi.org/10.5194/egusphere-2023-1962, 2023.

Chulakadabba, A., Sargent, M., Lauvaux, T., Benmergui, J. S., Franklin, J. E., Chan Miller, C., Wilzewski, J. S., Roche, S., Conway, E., Souri, A. H., Sun, K., Luo, B., Hawthrone, J., Samra, J., Daube, B. C., Liu, X., Chance, K. V., Li, Y., Gautam, R., … Wofsy, S. C. Methane Point Source Quantification Using MethaneAIR: A New Airborne Imaging Spectrometer [Preprint]. Gases/Remote Sensing/Validation and Intercomparisons, https://doi.org/10.5194/egusphere-2023-822, 2023.

Conway, E. K., Souri, A. H., Benmergui, J., Sun, K., Liu, X., Staebell, C., Chan Miller, C., Franklin, J., Samra, J., Wilzewski, J., Roche, S., Luo, B., Chulakadabba, A., Sargent, M., Hohl, J., Daube, B., Gordon, I., Chance, K., Wofsy, S. Level0-to-Level1B processor for MethaneAIR [Preprint]. Gases/Remote Sensing/Data Processing and Information Retrieval, https://doi.org/10.5194/amt-2023-111, 2023.

Cusworth, D. H., Duren, R. M., Thorpe, A. K., Olson-Duvall, W., Heckler, J., Chapman, J. W., Eastwood, M. L., Helmlinger, M. C., Green, R. O., Asner, G. P., Dennison, P. E., Miller, C. E. Intermittency of Large Methane Emitters in the Permian Basin. Environ. Sci. Technol. Lett., 8 (7), 567–573, https://doi.org/10.1021/acs.estlett.1c00173, 2021.

https://doi.org/10.5194/essd-2024-72
Preprint. Discussion started: 2 April 2024

Deighton, J. A., Townsend-Small, A., Sturmer, S. J., Hoschouer, J., Heldman, L. Measurements Show That Marginal Wells Are a Disproportionate Source of Methane Relative to Production. J. Air Waste Manag. Assoc., 70 (10), 1030–1042, https://doi.org/10.1080/10962247.2020.1808115, 2020

Duren, R. M., Thorpe, A. K., Foster, K. T., Rafiq, T., Hopkins, F. M., Yadav, V., Bue, B. D., Thompson, D. R., Conley, S., Colombi, N. K., Frankenberg, C., McCubbin, I. B., Eastwood, M. L., Falk, M., Herner, J. D., Croes, B. E., Green, R. O., Miller, C. E. California's methane super-emitters. Nature, 575(7781), 180–184, https://doi.org/10.1038/s41586-019-1720-3, 2019.

EDGAR (Emissions Database for Global Atmospheric Research) Community GHG Database, a collaboration between the European Commission, Joint Research Centre (JRC), the International Energy Agency (IEA), and comprising IEA-EDGAR CO2, EDGAR CH4, EDGAR N2O, EDGAR F-GASES version 8.0, European Commission, JRC (Datasets), https://edgar.jrc.ec.europa.eu/report_2023, 2023.

El Abbadi, S., Chen, Z., Burdeau, P., Rutherford, J., Chen, Y., Zhang, Z., Sherwin, E., Brandt, A. Comprehensive evaluation of aircraft-based methane sensing for greenhouse gas mitigation [Preprint]. Engineering, https://doi.org/10.31223/X51D4C, 2023.

Elvidge, C., Zhizhin, M., Baugh, K., Hsu, F.-C., Ghosh, T. Methods for Global Survey of Natural Gas Flaring from Visible Infrared Imaging Radiometer Suite Data. Energies, 9(1), 14, https://doi.org/10.3390/en9010014, 2015.

Enverus Prism, www.enverus.com, last accessed February 06, 2024

EPA: United States Environmental Protection Agency, Inventory of US Greenhouse Gas Emissions and Sinks, https://www.epa.gov/ghgemissions/inventory-us-greenhouse-gas-emissions-and-sinks (last access: 20 December 2023), 2022.

GMP: Global Methane Pledge, https://www.ccacoalition.org/en/resources/global-methane-pledge (last access: 12 December 2023), 2021.

Hoffman, M. D., Gelman, A. The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. J. Mach. Learn. Res., 15 (1), 1593-1623, https://api.semanticscholar.org/CorpusID:12948548, 2014.

IPCC: The Intergovernmental Panel on Climate Change, Sixth Assessment Report. Climate Change 2021: The Physical Science Basis, Summary for Policy Makers, https://www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_WGI_SPM.pdf (last access: 12 December 2023), 2021.

Jacob, D. J., Varon, D. J., Cusworth, D. H., Dennison, P. E., Frankenberg, C., Gautam, R., Guanter, L., Kelley, J., McKeever, J., Ott, L. E., Poulter, B., Qu, Z., Thorpe, A. K., Worden, J. R., Duren, R. M. Quantifying methane emissions from the global scale down to point sources using satellite observations of atmospheric methane. Atmos. Chem. Phys., 22(14), 9617–9646, https://doi.org/10.5194/acp-22-9617-2022, 2022.

Johnson, M. R., Conrad, B. M., Tyner, D. R. Creating Measurement-Based Oil and Gas Sector Methane Inventories Using Source-Resolved Aerial Surveys. Commun. Earth Environ., 4 (1), 139, https://doi.org/10.1038/s43247-023-00769-7, 2023.

Kunkel, W. M., Carre-Burritt, A. E., Aivazian, G. S., Snow, N. C., Harris, J. T., Mueller, T. S., Roos, P. A., Thorpe, M. J. Extension of Methane Emission Rate Distribution for Permian Basin Oil and Gas Production Infrastructure by Aerial LiDAR. Environ. Sci. Technol., 57 (33), 12234–12241, https://doi.org/10.1021/acs.est.3c00229, 2023.

Lan, X., Talbot, R., Laine, P., Torres, A. Characterizing Fugitive Methane Emissions in the Barnett Shale Area Using a Mobile Laboratory. Environ. Sci. Technol., 49(13), 8139–8146, https://doi.org/10.1021/es5063055, 2015.

Lavoie, T. N., Shepson, P. B., Gore, C. A., Stirm, B. H., Kaeser, R., Wulle, B., Lyon, D., Rudek, J. Assessing the Methane Emissions from Natural Gas-Fired Power Plants and Oil Refineries. Environ. Sci. Technol., 51(6), 3373–3381, https://doi.org/10.1021/acs.est.6b05531, 2017.

Lin, J. C.; Bares, R.; Fasoli, B.; Garcia, M.; Crosman, E.; Lyman, S. Declining Methane Emissions and Steady, High Leakage Rates Observed over Multiple Years in a Western US Oil/Gas Production Basin. Sci. Rep. 2021, 11 (1), 22291. https://doi.org/10.1038/s41598-021-01721-5.

https://doi.org/10.5194/essd-2024-72
Preprint. Discussion started: 2 April 2024

Open Access Earth System Science Data Discussions

Lu, X., Jacob, D. J., Wang, H., Maasakkers, J. D., Zhang, Y., Scarpelli, T. R., Shen, L., Qu, Z., Sulprizio, M. P., Nesser, H., Bloom, A. A., Ma, S., Worden, J. R., Fan, S., Parker, R. J., Boesch, H., Gautam, R., Gordon, D., Moran, M. D., Reuland, F., Villasana, C. A. O., Andrews, A. Methane Emissions in the United States, Canada, and Mexico: Evaluation of National Methane Emission Inventories and 2010–2017 Sectoral Trends by Inverse Analysis of in Situ (GLOBALVIEWplus CH₄ ObsPack) and Satellite (GOSAT) Atmospheric Observations. Atmospheric Chem. Phys., 22 (1), 395–418, https://doi.org/10.5194/acp-22-395-2022, 2022.

Lu, X., Jacob, D. J., Zhang, Y., Shen, L., Sulprizio, M. P., Maasakkers, J. D., Varon, D. J., Qu, Z., Chen, Z., Hmiel, B., Parker, R. J., Boesch, H., Wang, H., He, C., Fan, S. Observation-Derived 2010-2019 Trends in Methane Emissions and Intensities from US Oil and Gas Fields Tied to Activity Metrics. Proc. Natl. Acad. Sci., 120 (17), e2217900120, https://doi.org/10.1073/pnas.2217900120, 2023.

Lyon, D. R., Hmiel, B., Gautam, R., Omara, M., Roberts, K. A., Barkley, Z. R., Davis, K. J., Miles, N. L., Monteiro, V. C., Richardson, S. J., Conley, S., Smith, M. L., Jacob, D. J., Shen, L., Varon, D. J., Deng, A., Rudelis, X., Sharma, N., Story, K. T., Brandt, A. R., Kang, M., Kort, E. A., Marchese, A. J., Hamburg, S. P. Concurrent Variation in Oil and Gas Methane Emissions and Oil Price during the COVID-19 Pandemic. Atmospheric Chem. Phys., 21 (9), 6605–6626, https://doi.org/10.5194/acp-21-6605-2021, 2021.

Maasakkers, J. D., McDuffie, E. E., Sulprizio, M. P., Chen, C., Schultz, M., Brunelle, L., Thrush, R., Steller, J., Sherry, C., Jacob, D. J., Jeong, S., Irving, B., Weitz, M. A Gridded Inventory of Annual 2012–2018 U.S. Anthropogenic Methane Emissions. Environ. Sci. Technol., 57 (43), 16276–16288, https://doi.org/10.1021/acs.est.3c05138, 2023.

Mitchell, A. L., Tkacik, D. S., Roscioli, J. R., Herndon, S. C., Yacovitch, T. I., Martinez, D. M., Vaughn, T. L., Williams, L. L., Sullivan, M. R., Floerchinger, C., Omara, M., Subramanian, R., Zimmerle, D., Marchese, A. J., Robinson, A. L. Measurements of Methane Emissions from Natural Gas Gathering Facilities and Processing Plants: Measurement Results. Environ. Sci. Technol., 49 (5), 3219–3227, https://doi.org/10.1021/es5052809, 2015.

OGCI: The Oil and Gas Climate Initiative Reporting Framework. Version 3.5, https://www.ogci.com/wp-content/uploads/2021/11/OGCI_Reporting_Framework_ver_3.5_Nov_2021.pdf (last access: 12 December 2023), 2021.

Omara, M., Gautam, R., O'Brien, M. A., Himmelberger, A., Franco, A., Meisenhelder, K., Hauser, G., Lyon, D. R., Chulakadabba, A., Miller, C. C., Franklin, J., Wofsy, S. C., Hamburg, S. P. Developing a spatially explicit global oil and gas infrastructure database for characterizing methane emission sources at high resolution. Earth Syst. Sci. Data, 15(8), 3761–3790, https://doi.org/10.5194/essd-15-3761-2023, 2023.

Omara, M., Gautam, R., Himmelberger, A. et al. Measurement-Based Spatially Explicit Methane Emission Inventory (EI-ME) https://doi.org/10.5281/zenodo.10734300, 2024.

Omara, M., Zavala-Araiza, D., Lyon, D. R., Hmiel, B., Roberts, K. A., Hamburg, S. P. Methane emissions from US low production oil and natural gas well sites. Nat. Commun., 13(1), 2085, https://doi.org/10.1038/s41467-022-29709-3, 2022.

Omara, M., Sullivan, M. R., Li, X., Subramanian, R., Robinson, A. L., Presto, A. A. Methane Emissions from Conventional and Unconventional Natural Gas Production Sites in the Marcellus Shale Basin. Environ. Sci. Technol., 50 (4), 2099–2107, https://doi.org/10.1021/acs.est.5b05503, 2016.

Omara, M., Zimmerman, N., Sullivan, M. R., Li, X., Ellis, A., Cesa, R., Subramanian, R., Presto, A. A., Robinson, A. L. Methane Emissions from Natural Gas Production Sites in the United States: Data Synthesis and National Estimate. Environ. Sci. Technol., 52 (21), 12915–12925, https://doi.org/10.1021/acs.est.8b03535, 2018.

Plant, G., Kort, E. A., Brandt, A. R., Chen, Y., Fordice, G., Gorchov Negron, A. M., Schwietzke, S., Smith, M., Zavala-Araiza, D. Inefficient and unlit natural gas flares both emit large quantities of methane. Science, 377(6614), 1566–1571, https://doi.org/10.1126/science.abq0385, 2022.

Rella, C. W., Tsai, T. R., Botkin, C. G., Crosson, E. R., Steele, D. Measuring Emissions from Oil and Natural Gas Well Pads Using the Mobile Flux Plane Technique. Environ. Sci. Technol., 49 (7), 4742–4748, https://doi.org/10.1021/acs.est.5b00099, 2015.

Riddick, S. N., Mauzerall, D. L., Celia, M. A., Kang, M., Bressler, K., Chu, C., Gum, C. D. Measuring Methane Emissions from Abandoned and Active Oil and Gas Wells in West Virginia. Sci. Total Environ., 651, 1849–1856, https://doi.org/10.1016/j.scitotenv.2018.10.082, 2019.

Robertson, A. M., Edie, R., Field, R. A., Lyon, D., McVay, R., Omara, M., Zavala-Araiza, D., Murphy, S. M. New Mexico Permian Basin Measured Well Pad Methane Emissions Are a Factor of 5–9 Times Higher Than U.S. EPA Estimates. Environ. Sci. Technol., 54 (21), 13926–13934, https://doi.org/10.1021/acs.est.0c02927, 2020.

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

Rutherford, J. S., Sherwin, E. D., Ravikumar, A. P., Heath, G. A., Englander, J., Cooley, D., Lyon, D., Omara, M., Langfitt, Q., Brandt, A. R. Closing the Methane Gap in US Oil and Natural Gas Production Emissions Inventories. Nat. Commun., 12 (1), 4715, https://doi.org/10.1038/s41467-021-25017-4, 2021.

Salvatier, J., Wiecki, T. V., Fonnesbeck, C. Probabilistic Programming in Python Using PyMC3. PeerJ Comput. Sci., 2, e55, https://doi.org/10.7717/peerj-cs.55, 2016.

Scarpelli, T. R., Jacob, D. J., Grossman, S., Lu, X., Qu, Z., Sulprizio, M. P., Zhang, Y., Reuland, F., Gordon, D., Worden, J. R. Updated Global Fuel Exploitation Inventory (GFEI) for methane emissions from the oil, gas, and coal sectors: Evaluation with inversions of atmospheric methane observations. Atmos. Chem. Phys., 22(5), 3235–3249, https://doi.org/10.5194/acp-22-3235-2022, 2022.

Shen, L., Jacob, D. J., Gautam, R., Omara, M., Scarpelli, T. R., Lorente, A., Zavala-Araiza, D., Lu, X., Chen, Z., Lin, J. National quantifications of methane emissions from fuel exploitation using high resolution inversions of satellite observations. Nat. Commun., 14(1), 4948, https://doi.org/10.1038/s41467-023-40671-6, 2023.

Shen, L., Zavala-Araiza, D., Gautam, R., Omara, M., Scarpelli, T., Sheng, J., Sulprizio, M. P., Zhuang, J., Zhang, Y., Qu, Z., Lu, X., Hamburg, S. P., and Jacob, D. J.: Unravelling a large methane emission discrepancy in Mexico using satellite observations, Remote Sens. Environ., 260, 112461, https://doi.org/10.1016/j.rse.2021.112461, 2021.

Shen, L., Gautam, R., Omara, M., Zavala-Araiza, D., Maasakkers, J. D., Scarpelli, T. R., Lorente, A., Lyon, D., Sheng, J., Varon, D. J., Nesser, H., Qu, Z., Lu, X., Sulprizio, M. P., Hamburg, S. P., Jacob, D. J. Satellite Quantification of Oil and Natural Gas Methane Emissions in the US and Canada Including Contributions from Individual Basins. Atmospheric Chem. Phys., 22 (17), 11203–11215, https://doi.org/10.5194/acp-22-11203-2022, 2022.

Staebell, C., Sun, K., Samra, J., Franklin, J., Chan Miller, C., Liu, X., Conway, E., Chance, K., Milligan, S., Wofsy, S. Spectral calibration of the MethaneAIR instrument. Atmos. Meas. Tech., 14(5), 3737–3753, https://doi.org/10.5194/amt-14-3737-2021, 2021.

Subramanian, R., Williams, L. L., Vaughn, T. L., Zimmerle, D., Roscioli, J. R., Herndon, S. C., Yacovitch, T. I., Floerchinger, C., Tkacik, D. S., Mitchell, A. L., Sullivan, M. R., Dallmann, T. R., Robinson, A. L. Methane Emissions from Natural Gas Compressor Stations in the Transmission and Storage Sector: Measurements and Comparisons with the EPA Greenhouse Gas Reporting Program Protocol. Environ. Sci. Technol., 49 (5), 3252–3261, https://doi.org/10.1021/es5060258, 2015.

UNFCCC. The United Nations Framework Convention on Climate Change: GHG data from UNFCC. https://unfccc.int/topics/mitigation/resources/registry-and-data/ghg-data-from-unfccc (last access: 12 December 2023)

Varon, D. J., Jacob, D. J., Hmiel, B., Gautam, R., Lyon, D. R., Omara, M., Sulprizio, M., Shen, L., Pendergrass, D., Nesser, H., Qu, Z., Barkley, Z. R., Miles, N. L., Richardson, S. J., Davis, K. J., Pandey, S., Lu, X., Lorente, A., Borsdorff, T., … Aben, I. Continuous weekly monitoring of methane emissions from the Permian Basin by inversion of TROPOMI satellite observations. Atmos. Chem. Phys., 23(13), 7503–7520, https://doi.org/10.5194/acp-23-7503-2023, 2023.

Williams, J. P., Regehr, A., Kang, M. Methane Emissions from Abandoned Oil and Gas Wells in Canada and the United States. Environ. Sci. Technol., 55 (1), 563–570, https://doi.org/10.1021/acs.est.0c04265, 2024.

https://doi.org/10.5194/essd-2024-72
Preprint. Discussion started: 2 April 2024

Williams, J.P., Gautam, R., Omara, M., Himmelberger, H., MacKay, K. et al. Contribution of Small Methane Emission Sources to Total Regional Emissions from the US Oil and Gas Sector. In preparation.

Yacovitch, T. I., Herndon, S. C., Pétron, G., Kofler, J., Lyon, D., Zahniser, M. S., Kolb, C. E. Mobile Laboratory Observations of Methane Emissions in the Barnett Shale Region. Environ. Sci. Technol., 49 (13), 7889–7895, https://doi.org/10.1021/es506352j, 2015.

Zavala-Araiza, D., Omara, M., Gautam, R., Smith, M. L., Pandey, S., Aben, I., Almanza-Veloz, V., Conley, S., Houweling, S., Kort, E. A., Maasakkers, J. D., Molina, L. T., Pusuluri, A., Scarpelli, T., Schwietzke, S., Shen, L., Zavala, M., and Hamburg, S. P.: A tale of two regions: methane emissions from oil and gas production in offshore/onshore Mexico, Environ. Res. Lett., 16, 024019, https://doi.org/10.1088/1748-9326/abceeb, 2021.

Zhang, Y., Gautam, R., Pandey, S., Omara, M., Maasakkers, J. D., Sadavarte, P., Lyon, D., Nesser, H., Sulprizio, M. P., Varon, D. J., Zhang, R., Houweling, S., Zavala-Araiza, D., Alvarez, R. A., Lorente, A., Hamburg, S. P., Aben, I., Jacob, D. J. Quantifying Methane Emissions from the Largest Oil-Producing Basin in the United States from Space. Sci. Adv., 6 (17), eaaz5120, https://doi.org/10.1126/sciadv.aaz5120, 2020.

Zimmerle, D., Vaughn, T., Luck, B., Lauderdale, T., Keen, K., Harrison, M., Marchese, A., Williams, L., Allen, D. Methane Emissions from Gathering Compressor Stations in the U.S. Environ. Sci. Technol., 54 (12), 7552–7561, https://doi.org/10.1021/acs.est.0c00516, 2020.