Author response to reviewer comments

Anonymous Referee # 3

The authors developed a detailed compilation of activity data and emissions measurements to estimate 2021 US oil and gas methane emissions. The measurements were drawn from published studies and used to develop emission rate distributions. These distributions were used to estimate methane emissions at the facility level and then, estimate methane emissions at the national level. The authors also compare their results to airborne measurements using MethaneAIR. Overall, their spatially explicitly inventory is a valuable contribution, especially for remote measurements using satellites. In addition, they find important spatial trends that can be used to improve emission estimates and inform mitigation. Below are some high-level and detailed comments that can improve clarity of the paper.

We thank Reviewer #3 for these detailed comments and review of our manuscript. We provide below point-by-point responses.

High-level comments:

The authors use measurement data from various years (not 2021) to estimate methane emissions for 2021. However, due to regulations, technology advancements, and other factors, methane emissions distributions may be changing over time, as the authors acknowledge. It doesn't appear that the authors try to correct for this temporal variability or address this in their discussion.

As we acknowledged in the Main Text, specific uncertainties related to the impact of changing operator practices and/or promulgated regulations of oil and gas methane emissions are difficult to quantify due to lack of data. However, for well sites (the largest contributor to the estimated methane emissions in this study), the large body of evidence from the measurement data that we synthesized, which span the years post-2011 EPA NSPS to 2020, do suggest that (i) for low production well sites, absolute methane emissions are weakly correlated with production rates and (ii) for non-low production well sites, newly developed high producing well sites exhibit lower methane loss rates (absolute emissions normalized by production) compared with aging and lower producing sites in this category. We leverage these insights based on empirical observations, which we assume generally hold true across basins, to model national methane emission rates, with specific application to activity data (both well site count and production rates) in 2021. As such, temporal variability that are predictable as part of the insights derived from these empirical observations (i.e., how do facility-level emissions sizes vary with changes in production rates—e.g., Figure 1a) are implicitly accounted for by constraining estimates to activity data specific to 2021.

We have included the following sentences in Section 3.2 to clarify:

> *"As noted previously, basin-level differences in total methane emissions could also be impacted by federal/state-level regulations of oil and gas methane emissions and/or operator-specific practices, affecting both the magnitude and temporal variability in emissions. While our methods are based on insights derived from empirical observations and robust modelling to estimate methane*

emissions specific to oil and gas activity in 2021, we lack sufficient data to characterize the impacts of specific regulations or operator practices. Further studies are needed to assess oil and gas methane emission trends and corresponding underlying drivers."

We provide further details below in response to similar comments from Reviewer #3.

The definition of the methane loss rate is unclear. Although the authors mention in the Results (though I would have expected this in the Methods) that an 80% methane content was assumed, it remains unclear how the conversion was done for oil facilities.

We include the following clarification text in Section 2.4 of Methods:

"We compute basin-level and national methane loss rates as the ratio of estimated basin-level methane emissions to gross methane production in 2021, based on gross natural gas production data from Enverus Prism (Enverus, 2024) and an assumed average methane content of 80% in natural gas. Our assumption of an average 80% methane content in natural gas is informed by regional estimates of methane composition in natural gas based on the EPA GHGI (EPA, 2022). We acknowledge that uncertainties in methane composition across basins likely increases uncertainties in our overall methane loss rate calculations. Further studies on basin-level methane composition are needed to constrain these uncertainties. This methane intensity metric allows for a direct comparison of estimated methane losses relative to gross methane production across different basins. While our use of gross methane production accounts for emissions from associated gas produced during oil operations, the results are not intended to represent lifecycle emission intensities, which are outside the scope of this work."

The authors need to provide a clear definition for "measurement-based inventories". The authors do not measure all sources but use methane emission rate distributions based on available measurements, which does not cover all sites but some subset. Therefore, the question is how many and which measurements are needed to have a representative sample that can be used to create "measurement-based inventories".

We appreciate Reviewer # 3's comment on clarifying our definition for "measurement-based inventories." We include the following revisions in the Introduction Section 1:

"In this work, we utilize previous peer-reviewed facility-level measurement data for methane emissions at oil and gas facilities 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. <u>Our measurement-based inventory differs from other "bottom-up" inventories that use generic emission factors (e.g., EPA GHGI) in that we leverage empirical observations to derive insights on facility-level methane emission distributions useful for estimating population mean total methane emissions."</u>

Previous research have used a limited sample of measurement-based data on facility-level oil and gas methane emissions to develop an improved inventory of basin-level or regional methane emissions. For example, Zavala-Araiza et al. (2016) used facility-level measurement datasets of the order of hundreds of measurements in combination with a robust statistical model to characterize facility-level methane emission distributions and estimate the total Barnett oil and gas methane emissions, which were then validated (and showed good agreement) with independent airborne measurements. Alvarez et al. (2018) used facility-level methane measurements from 433 production sites and measurements at other facility types (e.g., compressor stations, processing plants) to develop emission models to estimate basin-level methane emissions that were validated with independent airborne measurements, and the insights from these models were used to estimate national methane emissions conditional on oil and gas activity data in 2015. These works (Zavala-Araiza et al., 2015 and Alvarez et al., 2018) demonstrated that measurement-based inventories developed using these methods produce results that are in good agreement, within statistical uncertainty, of independent airborne measurements of total area methane emissions. Our study follows similar approaches and yields results that are comparable, within statistical uncertainty, of independent airborne and satellite-based estimates, as we discuss in the Results and Discussion section.

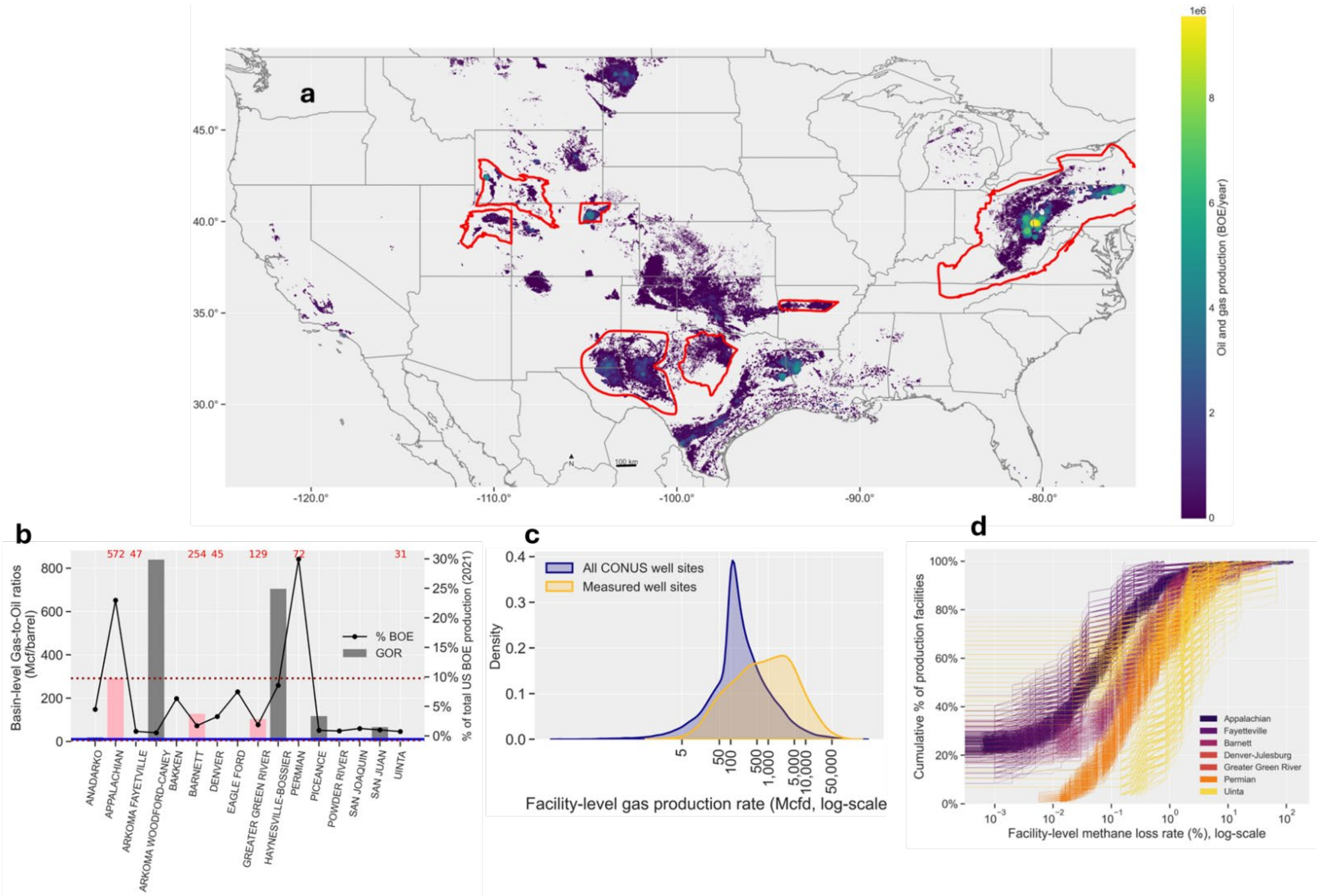
To characterize how many samples are required to minimize uncertainties in the development of regional measurement-based inventories, well-designed coordinated measurement campaigns, employing multiscale measurements from ground-based facility-level measurements to top-down airborne/satellite measurements would be needed. Such assessment and related analyses are beyond the scope of this work.

We have revised Section 2.3 of our manuscript to include the following clarification texts and discussion of sample representativeness:

> *"Our approach for estimating regional and national oil and gas methane emissions builds upon previous works that used data from hundreds to thousands of ground-based facility-level measurements (Zavala-Araiza et al., 2015; Alvarez et al., 2018; Omara et al. 2018; Omara et al., 2022) in combination with robust probabilistic models integrated with oil and gas activity data. Zavala-Araiza et al. (2015) and Alvarez et al. (2018) demonstrated that measurement-based inventories developed using these methods produce total methane emission results that are in good agreement, within statistical uncertainty, of independent airborne measurements of total area methane emissions.*
>
> *For non-low production well sites (average facility-level production rates > 15 boed), we begin by evaluating facility representativeness on the basis of (i) geographical diversity of measurements, (ii) distribution of facility-level production rates of measurements compared with the national population of well site facilities, and (iii) the distribution of facility-level methane emission rates across basins (Supplemental Fig. 3). Our measurement data, while limited in sample size, covers eight major US oil and gas basins with diverse oil and gas production characteristics, including the Appalachian, the Permian, Uinta, Barnett, Fayetteville, Greater Green River, and Denver-Julesburg. The wide range in basin-level gas-to-oil ratios (~1 to 800 Mcf/barrel) is well represented in the data for the sampled basins (Supplementary Fig. 3b).*
>
> *In addition, the distribution of facility-level natural gas production rates shows reasonable overlap with that for the national population of non-low production facilities, and the broad range in distribution of facility-level production rates across the national population of sites (~90 Mcfd to >50,000 Mcfd) is well represented in the sampled sites (Supplementary Fig. 3c). However, the distribution of production rates for the sampled sites suggests potential bias toward higher-producing sites relative to the national distribution (Supplementary Fig. 3c). We account for any such potential biases by developing emission models based on production-normalized methane loss*

rate distributions (methane emitted relative to methane produced) across seven cohorts of specific gas production rates (further details below).

We develop and use probabilistic emission rate distributions based on production-normalized methane loss rates, which shows a wide range <0.01% to >90% (Figure 1a) across all basins (Supplementary Fig. 3d), reflecting, in part, the diversity in production characteristics within and across basins. We use production-normalized methane loss rate distributions because (i) the empirical data across a wide diversity of oil and gas production facilities suggests an inverse relationship in which high-producing facilities exhibit comparatively lower methane loss rates, and vice versa (Figure 1a) and (ii) the consolidated dataset includes measurements collected in earlier years before 2021. By using the production-normalized methane loss rate distribution models for specific cohorts of facility-level production rates, we do not model any particular site that is active in 2021 as exhibiting the same emission rate size as observed when measurements were taken in the past, as the empirical data and the model constrains facility-level methane loss rates to production levels, which will be time-variant. As such, we provide a necessary constraint on our estimates, effectively adjusting modelled facility-level methane emission rates if production rates have substantially changed over time."



Supplementary Fig. 3: Geographical coverage, distribution of facility-level production rates, and emission rates for well site measurements used in this study.

a. The map shows the well site oil and gas production data for 2021, color-coded by combined oil and gas production rates (boe/year). Major basins for which substantial measurement-based data on oil and gas methane emissions are available are highlighted in red. **b.** Assessment of the basin-level production characteristics, based on average gas-to-oil ratios in Mcf/barrel in 2021. The bar plots show the basin-level GOR ratios, light pink bars correspond to basins for which measurement-based data (see Main Text on criteria) are available. The number of samples are shown in top x axis in red. The solid blue line shows the average GOR ratio for all sites in the US in 2021 (average of 11 Mcf/bbl) and the dotted dark red lines show the minimum (5 Mcf/barrel) and maximum (230 Mcf/barrel) for all the basins for which we have measurement-based data. The right y-axis shows the % of total US onshore BOE production that is accounted for by each major US basin. **c.** Histogram of gas production rates comparing the distribution for the sampled non-low production sites with that for the population of non-low production sites in the US. **d.** Comparison of empirical distribution of facility-level production-normalized methane loss rates for the major basins for which measurements are available. We performed non-parametric bootstrap resampling, with replacement, of the data for each basin, repeated 10^4 times to generate the likely extent or uncertainties of the distributions conditional on empirical observations. For each basin, we plot the distributions based on the 10^4 bootstrap results.

Detailed Comments:

L16: Provide a clear definition of "measurement-based inventories". It will be helpful to clarify the extent to which measurements are conducted for the inventory to be considered "measurement-based".

Please see author response above.

L18: How is "representativeness" assessed?

Please see author response above.

L20: How is the comprehensiveness of the spatial data assessed?

Comprehensiveness of spatial data is based on the comparison of our spatially explicit data (sourced primarily from Enverus Prism, a proprietary data source, www.enverus.com) with reported metrics in official energy statistics. For example, the US EIA reports a total gross national gas production of ~42 Tcf in 2021 (https://www.eia.gov/dnav/ng/ng_prod_sum_dc_NUS_mmcf_a.htm) consistent with the totals from the spatially explicit database (wellheads with allocated production data) available from Enverus Prism (~42 Tcf in 2021). Similarly, our spatially-explicit data for midstream facilities are generally consistent with official estimates from the EPA GHGI (for example, the GHGI reports a total of ~2,000 transmission compressor stations, similar to the spatially explicit activity data used in this work). In Section 2.5 of the Main Text, we acknowledge that there could be uncertainties in oil and gas activity data that are difficult to quantify because much of this information is based on operator-reported data, with unknown uncertainties.

We include the following clarification sentences in the revised manuscript in Section 2.1:

> *"We consider these spatial data as comprehensive for the US oil and gas production well sites as it is consistent with the official gross oil and gas production reported by the US Energy Information Administration for 2021 (e.g., the sum of gross gas production from spatially explicit well-level production data from Enverus Prism is consistent with the total of ~42 Tcf of US natural gas gross withdrawals reported by the US Energy Information Administration, https://www.eia.gov/dnav/ng/ng_prod_sum_dc_NUS_mmcf_a.htm)."*

L25: Is production data at the annual level?

Yes. We have revised this sentence for clarity:

> *"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 for the year 2021, from which we develop a mean estimate of annual national methane emissions, resolved at $0.1^o \times 0.1^o$ spatial scales (~10 km × 10 km)."*

L28-32: Very long sentence. Break sentence in two.

We have revised this sentence as follows:

"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 2021. The MethaneAIR assessment showed reasonable agreement with independent regional methane quantification results in sub-regions of the Permian and Uinta basins and indicated that diffuse area sources accounted for the majority of total oil and gas emissions in these two regions."

L46: Countries submit national inventory reports to the UNFCCC but there is no UNFCCC Greenhouse Gas Inventory. They are required to follow IPCC guidelines.

We have revised this sentence as follows:

"At the national level, methane inventories are typically developed using "bottom-up" methods, for example, these methods are used by most countries that report annual inventories to the UNFCCC (UNFCCC, 2023)."

L53: Define what is meant by "measurement-based inventories". See previous and high-level comments.

Please see response above. We have included the following sentence for clarification:

"In this work, we utilize previous peer-reviewed facility-level measurement data for methane emissions at oil and gas facilities 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. <u>Our measurement-based inventory differs from other "bottom-up" inventories that use generic emission factors (e.g., EPA GHGI) in that we leverage empirical observations to derive insights on facility-level methane emission distributions useful for estimating population mean total methane emissions</u>."

L64: Reading on, it appears that the paper uses measurements not from 2021. How is the data corrected for temporal variability?

Temporal variability in facility-level methane emission rates could be influenced by several factors, including changes in facility operations (e.g., increased or decreased production activities over time, installation of emission control devices, or removal of auxiliary processing equipment, etc), frequency of scheduled or unscheduled maintenance activities leading to intentionally vented emissions, regulatory requirements on emission controls, voluntary efforts on emission controls, etc. We unfortunately lack the data needed to assess facility-level methane emission trends. However, for well sites—which accounts for the majority of our estimated emissions—the one important attribute that is consistently reported across the consolidated measurements in our study is the facility-level production rate. It is important to note that production declines substantially over time at a well site over time, and our consolidation of facility-level emission rate data for a wide range of production rates allows us to develop key insights that make it possible to use data collected in previous years to estimate emissions for the year 2021. Specifically, for the non-low production well site category, by using stratified methane loss rate distributions for specific cohorts of production rates (see Methods), we do not model any particular site that is active in 2021 as exhibiting the exact emission rate size exactly as observed when measurements were taken in the past, as the empirical data and the model constrains facility-level methane loss rates to production levels, which will be time-variant. This provides a necessary constraint on our model regarding temporal variability, but as we have acknowledged, our results will have uncertainties that are difficult to quantify and related to, among other factors, potential changes in emissions as regulations are enacted and implemented. Further comprehensive studies on total area methane emissions are needed to better understand temporal variability in emissions.

> *"We develop and use probabilistic emission rate distributions based on production-normalized methane loss rates, which shows a wide range <0.01% to >90% (Figure 1a) across all basins (Supplementary Fig. 3d), reflecting, in part, the diversity in production characteristics within and across basins. We use production-normalized methane loss rate distributions because (i) the empirical data across a wide diversity of oil and gas production facilities suggests an inverse relationship in which high-producing facilities exhibit comparatively lower methane loss rates, and vice versa (Figure 1a) and (ii) the consolidated dataset includes measurements collected in earlier years before 2021. By using the production-normalized methane loss rate distribution mode for specific cohorts of facility-level production rates, we do not model any particular site that is active in 2021 as exhibiting the same emission rate size as observed when measurements were taken in the past, as the empirical data and the model constrains facility-level methane loss rates to production levels, which will be time-variant. As such, we provide a necessary constraint on our estimates, effectively adjusting modelled facility-level methane emission rates if production rates have substantially changed over time."*

L65: Measurement data from which year(s)?

We have revised this sentence as follows:

> *"First, we develop statistically robust facility-level methane emission models based on measurement data collected in the years post-2011 (when EPA's NSPS were first proposed) through 2020. We use these models to estimate national methane emissions, on both an absolute basis (Tg/year) and production-normalized basis (% emitted relative to methane production)."*

L67: relative to methane or natural gas production?

Production-normalized methane loss rates are relative to methane production.

L81-83: Above, the authors mention that the loss rates are normalized by methane production. How is oil/gas production converted to methane production? Is the production data for 2021 used or is the production data corresponding to the month/year of measurement used?

For the measurement-based data, methane loss rates, as reported in the different studies, are based on methane production specific to the time in which measurements occurred (averaged to hourly production rates based on the production data for the month in which measurements occurred because production data are generally reported on a monthly basis).

L78-91 describes the assessment of the production characteristics (annual production for 2021, which is then expressed as a daily average in Mcfd, bbl per day, etc based on the number of production days in the year). This assessment is specific to the national population of well sites, based on monthly data that is reported at the well-level.

L141 describes the method for calculating methane loss rate, which is defined methane emitted relative to methane produced. In this study, methane produced at a well site facility is a factor of gross gas production and the methane content. Further details can be found in the revised Section 2.4, which includes the following paragraph on computation of basin-level methane loss rates:

> *"We compute basin-level and national methane loss rates as the ratio of estimated basin-level methane emissions to gross methane production in 2021, based on gross natural gas production data from Enverus Prism (Enverus, 2024) and an assumed average methane content of 80% in natural gas. Our assumption of an average 80% methane content in natural gas is informed by regional estimates of methane composition in natural gas based on the EPA GHGI (EPA, 2022). We acknowledge that uncertainties in methane composition across basins likely increases uncertainties in our overall methane loss rate calculations. Further studies on basin-level methane composition are needed to constrain these uncertainties. This methane intensity metric allows for a direct comparison of estimated methane losses relative to gross methane production across different basins. While our use of gross methane production accounts for emissions from associated gas produced during oil operations, the results are not intended to represent lifecycle emission intensities, which are outside the scope of this work."*

Table 1: Are the estimated total methane emissions reported in column 6 done by the authors here in this paper or are these previous results? If they are estimated in this paper, they are better placed in the results.

The EI-ME estimated total methane emissions in Table 1 are new results estimated as part of this study. They are included in this Table at this point in the manuscript to contextualize both the activity data for various infrastructure categories, the methane measurement data sources, as well as the EPA's GHGI estimates. The discussion of these results is presented in the Results and Discussion section, supplemented with relevant figures.

L130-135: Many of these measurements were conducted before 2021. There needs to be a description as to how these measurements can be used to estimate emissions in 2021, and if some adjustments are needed.

Please see detailed responses above.

L141: what are the units for the methane loss rate? If it's unitless, it should say so. Is "CH_4" methane lost or measured? What is "Gas"? All the variables here need to defined and their units clearly provided after the equation.

The production-normalized methane loss rate is unitless. We have revised the equation as follows:

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

where CH_4 is the measured facility-level methane emission rate, Gas [Mcfd] is the reported gas production rate in thousand cubic feet [Mcf] per day, 19.2 kg/Mcf is the methane density at 60 ^{o}F (15.5 ^{o}C) and 1 atm, and 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).

Figure 2. These distributions are better placed in the Results section.

We view these distributions as an important component of the methods and models for estimating total methane emissions. We focus the Results and Discussion section on the outputs of these models.

Figure 3. The K-S test provides a measure of goodness of fit. How can it be used to assess the representativeness of the underly methane emissions measurements?

The inset figure showing distribution and K-S test results was not intended to suggest a method for assessing facility representativeness. The goodness of fit tests occurs as part of the emission distribution modeling. We have revised this general schematic as follows:

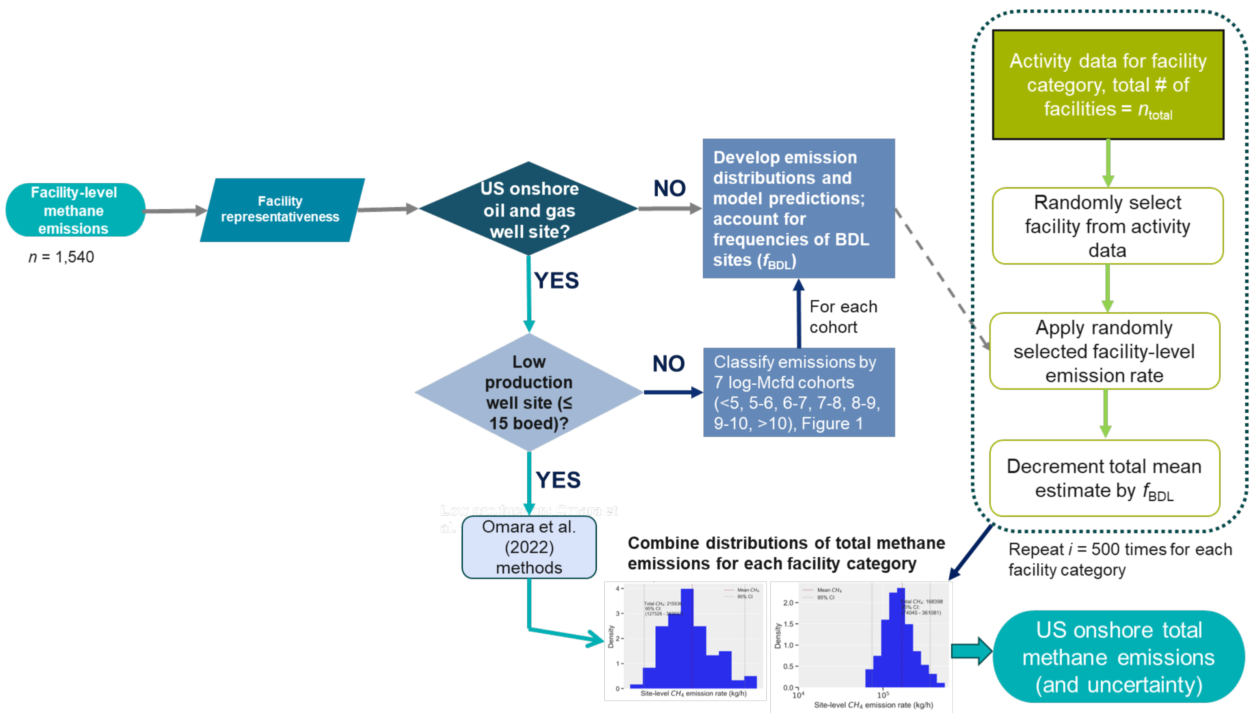


Figure 3. There are some arrows missing. I suggest an arrow be added to the black line from the dashed rounded box to the site-level CH4 emission rate histogram.

We thank Reviewer 3 for this suggestion. We have revised Figure 3 accordingly.

L240: How was 500 selected?

500 was selected as a reasonably simulation size that is not too computationally intensive to implement but that also gives sufficient statistical power to develop robust model uncertainty assessment.

We have included the following clarification sentence:

> *"We use 500 simulation results for each facility as a reasonable simulation size that is not too computationally intensive to implement but that also gives sufficient statistical power to develop robust model uncertainty assessment."*

L260-261: How was data limitation determined? There are published studies on downstream natural gas, post-meter, and abandoned well emissions. Are the authors looking for some specific number of measurements?

Given the focus of our study on developing spatially-explicit measurement-based methane emissions inventory, we did not include these sources due to a general lack of comprehensive spatially explicit activity data.

We have revised the sentence to clarify the lack of comprehensive spatially explicit data for these sources:

> *"In addition, due to lack of comprehensive spatially explicit data, our measurement-based inventory does not include methane emissions from downstream natural gas distribution, LNG storage, post-meter emissions, and abandoned oil and gas wells."*

L271: Remove the word "However"

We have revised L271 as follows:

> *"In addition, 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)."*

L288-289: methane content in natural gas can be variable. Does the 95% CI include methane content variability or is it assume to be fixed at 80%?

Nationally and across regions we assume an average 80% methane content in natural gas.

We include the following paragraph in Section 2.4 to clarify:

> *"We compute basin-level and national methane loss rates as the ratio of estimated basin-level methane emissions to gross methane production in 2021, based on gross natural gas production data from Enverus Prism (Enverus, 2024) and an assumed average methane content of 80% in natural gas. Our assumption of an average 80% methane content in natural gas is informed by regional estimates of methane composition in natural gas based on the EPA GHGI (EPA, 2022). We acknowledge that uncertainties in methane composition across basins likely increases uncertainties in our overall methane loss rate calculations. Further studies on basin-level methane composition are needed to constrain these uncertainties. This methane intensity metric allows for a direct comparison of estimated methane losses relative to gross methane production across different basins. While our use of gross methane production accounts for emissions from associated gas produced during oil operations, the results are not intended to represent lifecycle emission intensities, which are outside the scope of this work."*

Fig. 4: The legends should be moved outside of the plot. It would be helpful if the groupings of bars separated by dashed lines were annotated – e.g., green bars should just be labeled GOSAT.

We have updated Fig. 4 as follows:

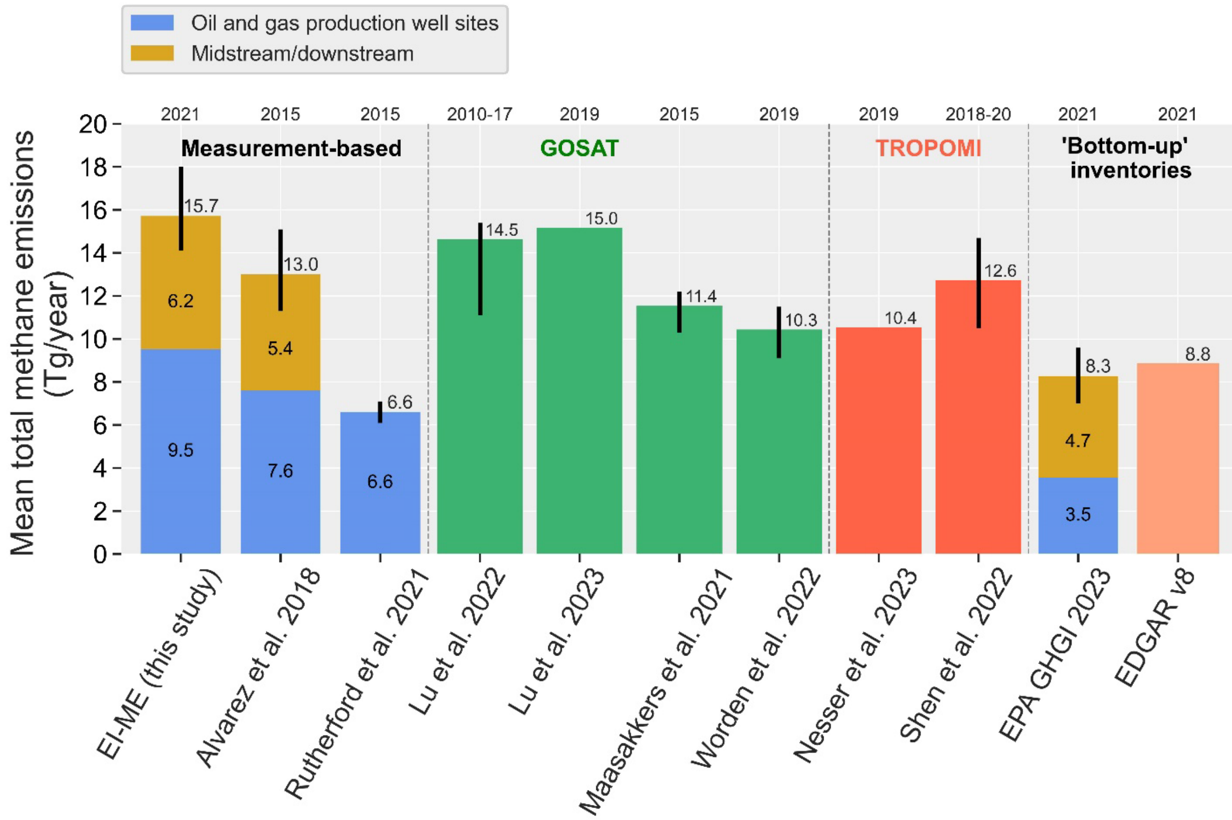


Figure 4. Comparison of this study's national estimate of total methane emissions from the oil and gas supply chain with previous measurement-based estimates. The first three bars show the oil and gas methane emissions estimated from facility-level measurements (this study, Alvarez et al. 2018) and production-sector-only methane emissions estimate by Rutherford et al. (2021) using models developed from 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. This study's estimate of total national methane emissions include ~0.1 Tg/year of estimated methane emissions for Alaska. The green bars and the red bars show the satellite-derived estimates for contiguous US based on GOSAT and TROPOMI observations, respectively. The last two bars show the "bottom-up" inventories from EPA GHGI and EDGAR v8 for the contiguous US. In all cases, the years for which methane emissions are estimated are shown on the top x-axis.

L294-299: The caption describes the first three bars only but should describe the rest as well.

We have revised Fig. 4 caption as follows:

> ***"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. This study's estimate of total national methane emissions include ~0.1 Tg/year of estimated methane emissions for Alaska. The green bars and the red bars show the satellite-derived estimates for contiguous US based on GOSAT and TROPOMI observations, respectively. The last two bars show the "bottom-up" inventories from EPA GHGI and EDGAR v8 for the contiguous US."

L345-346: If weakly correlated, should factors other than infrastructure be considered?

Other possible predictors of facility-level methane emissions could indeed be assessed; unfortunately, such data and related attributes were unavailable in the reported facility-level measurement data synthesized herein.

L501: How are production rates determined for midstream infrastructure?

Our estimate of facility-level emissions for compressor stations and processing plants are independent of throughput rates. In our spatial aggregation of methane loss rates, there will be cases where emissions in certain locations are dominated by midstream infrastructure in those locations that are handling oil and gas produced from well sites that are located in a different grid. This is possible given the grid resolution of ~25 km x 25 km. In these cases, the methane loss rates could be much higher than would be if expected if the production from the well site infrastructure were collocated with the gathering/transportation infrastructure within the chosen grid resolution of ~25 km x 25 km.

L502-503: If this study uses the data for 2021, would it not be different from the 2018 gridded EPA GHGI inventory data?

We expect minor differences in GHGI total methane emissions year-to-year. We have updated our assessment to use the spatially explicit GHGI data for the latest year for which these data are now available, i.e., 2020. We acknowledge we are not comparing the estimates for the same year although we do not expect the overall conclusions to change. Not that this only affects our comparison of the spatially explicit inventory from the EPA GHGI (based on Maasakkers et al., 2023) which extends only up to the year 2020. The official report from the EPA GHGI does include the total inventory estimates for the year 2021, which we compare in Table 1 and in Figure 4.

L509: Figure 6b shows methane emissions for a sub-region of the Uintah Basin, for which the agreement was good. Therefore, I don't think it's the correct figure to be pointing to here.

We have revised this to reference Figure 7.

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.

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.

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.

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.

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.

Zavala-Araiza, D.; Lyon, D. R.; Alvarez, R. A.; Davis, K. J.; Harriss, R.; Herndon, S. C.; Karion, A.; Kort, E. A.; Lamb, B. K.; Lan, X.; Marchese, A. J.; Pacala, S. W.; Robinson, A. L.; Shepson, P. B.; Sweeney, C.; Talbot, R.; Townsend-Small, A.; Yacovitch, T. I.; Zimmerle, D. J.; Hamburg, S. P. Reconciling Divergent Estimates of Oil and Gas Methane Emissions. Proc. Natl. Acad. Sci. 2015, 112 (51), 15597–15602. https://doi.org/10.1073/pnas.1522126112.