Author response to reviewer comments

Anonymous Referee # 2

Omara et al. constructed a high-resolution inventory for methane emissions from the U.S. oil and gas industry based on reported site-level measurements. The work provides a baseline that incorporates the best information for future evaluation of oil & gas methane emissions in the U.S., thus an important contribution to the field. I appreciate that the statistical method applied in the study is carefully designed with adequate sophistication. I'd recommend publication of the manuscript in ESSD, after the following comments are addressed.

We thank Reviewer #2 for these detailed comments and review of our manuscript. We provide below point-by-point responses.

1. The title indicates the inventory is for "US oil and gas methane emissions". However, the work is actually for "contiguous US onshore up- and mid-stream oil and gas emissions". The language can be more precise in places like abstract, conclusion, and Section 3.1 (when national totals are compared). While the focus on "onshore up- and mid-stream" is explained in the main text, I do not find any explicit language about the spatial extent (can only be inferred based on Fig. 7 and 8). As Alaska is an important oil & gas production region, I am concerned about if the comparisons are "apple to apple" in e.g. Section 3.1 when varied "national" totals are compared and discussed.

   Our estimates of oil and gas methane emissions are indeed for the continental United States, including Alaska. The full inventory, including estimated emissions for Alaska, is included in the GeoPackage data file (EI_ME_v1.0.gpkg, https://zenodo.org/records/10909191). For Alaska, we estimate total onshore oil and gas methane emissions of ~0.1 Tg in 2021, representing ~0.6% of the estimated national total. As such, including or excluding the estimated onshore-only methane emissions for Alaska does not alter the overall conclusions from our study in comparison with previous studies on national oil and gas methane emissions. We note that Alaska is generally excluded in recent works on satellite-based inversion studies of methane emissions, where primary focus has been on assessing the emissions from the contiguous (lower 48 states) United States (e.g., Shen et al., 2022; Lu et al., 2023; Nesser et al., 2023). We do provide a netcdf for only the contiguous US so as to allow for more of an "apple-to-apple" comparison with these studies.

   We have made revisions throughout the manuscript to clarify, where needed, that our estimates include estimates of ~0.1 Tg/year of for Alaska. For example, in the figure caption for Figure 4, we include the following sentence for clarity:

   "This study's estimate of total national methane emissions include ~0.1 Tg/year of estimated methane emissions for Alaska."

2. The method for low-production wells is not described in the manuscript. Reference to Omara et al. (2022) is provided. However, given the importance of low-production sites found in this work, a brief description of the main idea (e.g., method and data source) of Omara et al. (2022) seems necessary.

We have included the following sentences to briefly describe the methods for the estimation of methane emissions from low production sites:

"Briefly, we use the reported empirical observations (n = 240; Omara et al., 2022) in a hybrid Monte-Carlo and non-parametric probabilistic model that simultaneously estimates the frequency of below-detection-limit sites, the frequency of high-emitting sites representing the top 5% of emitting facilities based on absolute methane emissions, and the distribution of high-emitter methane emissions, while accounting for the weakly observed positive relationship between emission rates and production rates for the bottom 95% of emitting well sites. We integrate this model with spatially explicit activity data on low-production oil and gas well sites in 2021 (Enverus, 2024) to estimate their total methane emissions."

In addition, there is inconsistency in the current description of the well-site measurements (Table 1, Line 145-146, and Fig. 1a). Table 1 shows that there are n=1153 samples for low-production and non-low production sites combined. But line 145-146 and the caption of Fig.1a indicate that the figure is for non-low production sites only and includes n=1153 samples.

We have revised the Figure 1 Caption to fix the typo in the number of non-low production well sites:

"Facility-level methane emissions data (percent methane loss rate) as functions of gas production rates (n = 961 non-low production well sites)."

Line 142: The fraction of methane in produced natural gas should vary greatly from basin to basin. Is there better information for this parameter? What's the impact of this assumption on the uncertainty?

We do expect variability in the fraction of methane in produced natural gas across various basins, given the differences in geologic characteristics. However, the lack of comprehensive spatial data on methane composition across basins limits our ability to assess the impact of this parameter on our estimates of basin-level and national methane loss rates. Our assumption of an average 80% average methane content across basins is informed by estimates from the EPA Greenhouse Inventory, which reports regional variability of 77.1% in the Rocky Mountains region to 91.9% in the West Coast region, with an overall national average of 82.5%. Using an assumed higher methane content in natural gas leads to lower methane loss rate calculation, and vice versa. If we assume the full range of ~77% to 90% of regional variability in methane content, our computed average methane loss rate ranges from ~2.3% to 2.7%, which falls within our overall 95% confidence bounds of 2.3 to 2.9%.

We have included to following sentence in Section 3.1 for additional information:

"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 an average of 80% methane content in natural gas."

In Section 2.4, we provide additional clarification on the computation of 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."

3. Line 182-187: (1) Based on the description, it is unclear whether the distribution of fBDL or only the mean of fBDL is used in the "decrement total mean estimate by fBDL" step. (2) fBDL is defined below in L210 for mid-stream facilities, but the concept first appears here but fBDL is not defined.

We now define f_{BDL} in the first paragraph of Section 2.3:

"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* (f_{BDL})*, 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)."*

We also clarify that f_{BDL} is used to decrement the mean based on random draws from the modelled distribution:

> *"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}), randomly drawn from the modelled distributions."*

4. Section 2.1 Non-SI units are used throughout the text. It'd better to provide a conversion for SI units.

   The units used in the manuscript for oil and gas production are standard units used by oil and gas industry in the US. We have provided the following conversion: 1 ft3 = 0.0283 m3 and 1 bbl crude oil ~ 0.136 tonnes.

   > *"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; 1 barrel crude oil ~ 0.136 tonnes), gas (1 thousand cubic feet per day, Mcfd; 1 $ft^3 = 0.0283\ m^3$), 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, which occurred at <5% of producing wells (Supplementary Fig. 10)."*

5. Line 151: "as a function of"?

   We have revised this sentence to read:

   > *"Facility-level methane emissions data (percent methane loss rate) as a function of gas production rate."*

6. Table 2: Shen et al. (2022) results presented in Fig. 5. can also be shown here.

   We have now included the percent methane loss rate results from Shen et al. (2022) in Table 2.

7. A recent publication by Sherwin et al. (2024) in Nature reported a large dataset of aerial site measurements over US oil & gas basins. A discussion, if possible, can provide interested readers with useful information. For instance, (1) How does this study compare with Sherwin et al. (2024) at the basin level? (2) What's the implication of this large measurement data to the national inventory compilation?

   Sherwin, E.D., Rutherford, J.S., Zhang, Z. et al. US oil and gas system emissions from nearly one million aerial site measurements. Nature 627, 328–334 (2024).

   Sherwin et al. uses data from different snapshot facility-level aerial surveys conducted over multiple years (2017 to 2021) in combination with component-level simulations of missed emissions (i.e., below detection limits of aerial methods) to estimate the emission size distribution, i.e., the proportion of sites responsible for the majority of emissions in select regions across the US. The focus of the Sherwin et al. study and the methods used are different from the present study's methods and scope. Specifically, the study focuses on characterizing emission size distributions using aerial remote sensing data as opposed to the present study which is focused on the development of high-resolution spatially-explicit total methane emissions at the basin and national scale. Direct comparison with our study's results is limited by these and other caveats noted in the study. We acknowledge that more measurements and analyses are needed, specifically, direct quantification of total area methane emissions in combination with the assessment of the emissions from high emitting facilities will help constrain uncertainties in the assessment of emission size distributions. Interested readers are referred to Williams et al. (2024) for a detailed discussion of emission size distributions and uncertainties for the US upstream and midstream facilities and comparison with different studies that have explored this subject in recent years.

In Section 7 (Conclusions) we emphasize that further improvements to measurement-based methane emission inventories are possible:

> *"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."*
>
> *"There is a research need to develop robust statistical methods for effective integration of lower-detection-limit ground-based facility-level methane emissions data (such as data synthesized herein) with the growing number of airborne facility-level measurement studies, which generally have higher method detection limits (e.g., airborne methane remote sensing data in Duren et al., 2019; Cusworth et al., 2021; Sherwin et al., 2024). As demonstrated herein, improved integrated assessments of facility-level, regional, and national methane emission inventories, based on measurement data, 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."*

References

Cusworth, D. H.; Thorpe, A. K.; Ayasse, A. K.; Stepp, D.; Heckler, J.; Asner, G. P.; Miller, C. E.; Yadav, V.; Chapman, J. W.; Eastwood, M. L.; Green, R. O.; Hmiel, B.; Lyon, D. R.; Duren, R. M. Strong Methane Point Sources Contribute a Disproportionate Fraction of Total Emissions across Multiple Basins in the United States. Proc. Natl. Acad. Sci. 119 (38), e2202338119. https://doi.org/10.1073/pnas.2202338119, 2022.

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.

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.

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.

Nesser, H., Jacob, D. J., Maasakkers, J. D., Lorente, A., Chen, Z., Lu, X., Shen, L., Qu, Z., Sulprizio, M. P., Winter, M., Ma, S., Bloom, A. A., Worden, J. R., Stavins, R. N., Randles, C. A. High-resolution US methane emissions inferred from an inversion of 2019 TROPOMI satellite data: contributions from individual states, urban areas, and landfills, Atmos. Chem. Phys., 24, 5069–5091, https://doi.org/10.5194/acp-24-5069-2024, 2024.

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.

Sherwin, E.D., Rutherford, J.S., Zhang, Z. et al. US oil and gas system emissions from nearly one million aerial site measurements. Nature 627, 328–334 (2024), https://doi.org/10.1038/s41586-024-07117-5

Williams, J.P., Omara, M., Himmelberger, A., Zavala-Araiza, D., MacKay, K., Benmergui J., Sargent, M., Wofsy, S., Hamburg, S.P., Gautam, R. Small emission sources disproportionately account for a large majority of total methane emissions from the US oil and gas sector. [Preprint] https://doi.org/10.5194/egusphere-2024-1402