Author response to reviewer comments

Anonymous Reviewer #1

This work uses emission factors from ~20 published studies across ~9 regions to estimate national methane emissions from active mid- and up-stream oil/gas production facilities for 2021. Using infrastructure inventories (Enverus, OGIM database), regional emission rates were modelled and validated with airborne surveys.

The manuscript is well written and the subject is of suitable content for EGUsphere. The subject is timely as there is active discussion regarding how mitigation funding can most effectively be used to reduce fugitive emissions from O&G. The figures are well designed and informative.

We thank the reviewer for the valuable comments and edits, and we hope the following responses address their concerns.

All of the page references in the responses below reference the attached manuscript with tracked changes included. Text in ***"bold blue italics"*** references prior text from the manuscript, and text in ***"bold red italics"*** references new added text.

I have two main hesitations that together question the novelty of this work and the contributions that it provides. First, the chosen methodology, which is complicated and I am not convinced contributes to the authors results, discussion, or conclusions (see general comment 1). Second, the close similarity of this work prior work from this group (see Omara et al. 2022; 2024) questions the novelty of this manuscript. Specifically, the aggregation of emission factors is already published in Omara et al. 2018, 2022, & 2024. Scaling from emission factors to national budgets using Enverus is repeated from Omara et al. 2022 & 2024. Lastly, comparison of national/regional/basin-level emissions to airborne studies was previously done in Omara et al. 2022 & 2024.

Our understanding is that the reviewer is suggesting that Conclusions 1 and 2 [Section 5 – page 25] in the paper can be reached solely from the empirical measurements and emission factors without any extrapolation/modeling of distributions, which we respectfully disagree with.

Multiplying an average methane emission factor by the number of facilities can produce a rough estimate of total methane emissions, but is not a suitable approach for characterizing facility-level methane emission distributions, which must account for the stochasticity in facility-level methane emissions profiles and related uncertainties (for references that discuss this stochasticity: https://www.nature.com/articles/ncomms14012, https://pubs.acs.org/doi/full/10.1021/acs.est.6b04303, https://pubs.acs.org/doi/full/10.1021/acs.est.8b03535) Developing robust methods for characterizing such distributions at the basin- and national-scale is the focus of this work. While we do present estimates of total emissions estimates at the national/basin/aerial spatial scale, these are a by-product of our methodology and not the main findings, which are the detailed distributions of individual facility-level emission rate and the large majority contribution of total emissions linked to an aggregate of smaller emitting sources (i.e., the distributions presented in Figures 3, 5, and 6).

If, for example, an EPA GHGI emission factor (e.g., average methane emission rate per facility) and the associated confidence bounds (e.g., standard deviation of the mean) are applied to each individual facility to provide an independent emission rate, and this is repeated for all facilities in the CONUS, this simplified approach would not produce an accurate distribution of emission rates because a representative methane emission factor would still need to account for (i) facilities that may be non-emitting at any one time, (ii) the fact that different facility categories (including different production ranges of well sites) can emit at different rates at any one time, and (iii) the representativeness of facility-level empirical data (and inherent uncertainties in emissions quantification) when compared with the national population of facilities.

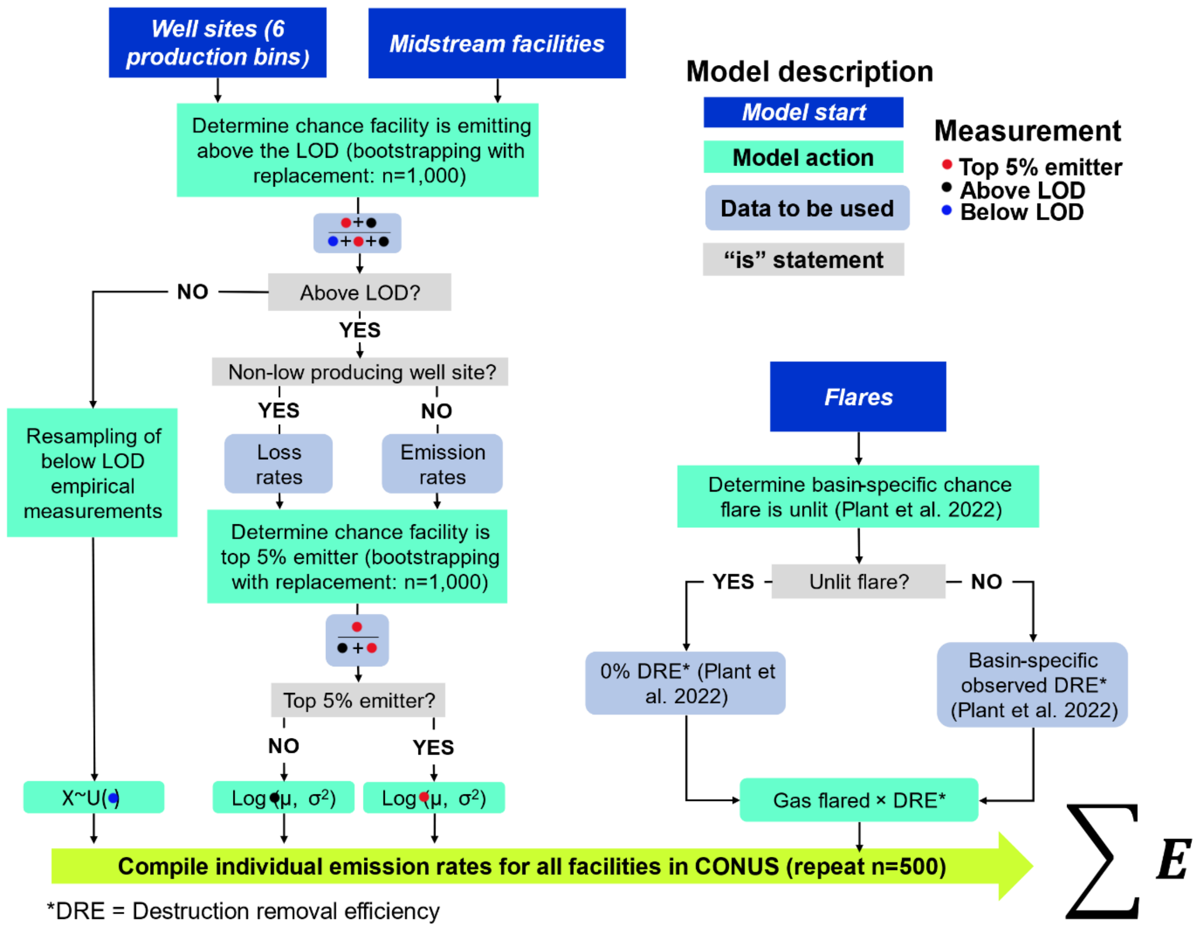
For these reasons, we believe a probabilistic modeling approach that accounts for these factors (and others) is essential to assessing emissions distributions and underpins the novel findings we present in this work. Moreover, the conclusions in terms of the specific emission rate thresholds and the aggregate emissions below those and their relative fractions to the total emissions across the US oil and gas upstream and midstream sectors as well as over each individual oil/gas basin has not been produced before based on empirically derived measurement-based analysis, which this study presents as a major step forward in our understanding the dynamics of oil/gas emissions and their source contributions which have important policy implications for measurement and mitigation, as we have discussed throughout the manuscript.

For the second comment mentioned by the reviewer regarding the methods in this work sharing similarities with previous studies (e.g. Omara et al. 2018,2022,2024), each of these previous studies had a different scope and presented different data outputs than our work. Our work differs from these previous studies by Omara et al. by i) estimating full methane emission distribution curves (i.e., not just total methane emissions) across multiple spatial scales ranging from the entire CONUS to the aerial remote sensing campaign survey regions, ii) presenting both the cumulative, and grouped, emission distribution curves for major oil/gas emitting facility categories, which allows for the clear distinction between emission distribution curves from different facility categories and a relative assessment of their total contributions to CONUS annual emissions, iii) presenting detailed comparisons to prior work (i.e., satellite and remote sensing studies) on the distribution of methane emissions across spatial scales and at different emission rate thresholds, which involved the additional analysis of data provided by other studies, and iv) a revised methodology that utilizes new ground-based facility-level measurement data and a separate approach for accounting for lit and unlit VIIRS flare detections. As we discuss in the Main Text, characterizing the full methane emissions distribution curves, that is, the contributions from individual facilities emitting below or above an emission rate threshold is crucially important for effective methane emissions mitigation. For example, any methane measurement platform with an established limit of detection and/or 90% probability of detection, could reference the emission distribution curves we present in our work, and determine a rough approximation of what percentage of total methane emissions their measurement method technique could capture. Our work represents the first comprehensive attempt to develop such an emissions distribution curve, using empirical measurements collected from ground-based measurement methods and robust probabilistic models to characterize the facility-level distributions for the full US upstream and midstream oil and gas methane sources. As part of this work, we estimate emissions from major upstream and midstream oil and gas methane sources, including well sites, natural gas gathering and transmission compressor stations, natural gas processing plants, and emissions from natural gas flaring facilities, accounting for methane emissions from both the lit and unlit flares.

- In order to improve the clarity of our methods, and to better illustrate differences between our methods/results and other studies, we have added the following text in the manuscript.
    - [page 6-7] ***"We calculate annual methane emissions from all facility categories (i.e., six production bins of production well sites, T&S compressor stations, G&B compressor stations, and processing plants, and VIIRS flare detections) using a multi-step probabilistic modeling approach adapted from multiple studies (Omara et al., 2018, 2022; Plant et al., 2022) (Fig. 2). Briefly, for each individual facility and VIIRS flare detection in the CONUS for 2021, we***

estimate an annually averaged methane emission rate using empirical measurement data, and consequently the cumulative distribution of methane emission rates from the aggregation of these individual emission rates. Each emission rate estimate is indexed according to the corresponding replicate (n=500), and we use these repetitions to determine uncertainty for the cumulative methane emission distribution curves. The detailed steps of this process for all facility categories and VIIRS flare detections are described below.”

- Revised Figure 2 added:



  - *“Figure 2: Flowchart describing the facility-level estimates, with steps colored according to the specific process and data being used. We note that methane emission rates for flares are calculated using a separate approach from that of production well sites and midstream facilities. Processing plants and T&S compressors are excluded from the determination of whether a facility is a top 5% emitter due to a lack of available empirical measurement data.”*

- New text added in the Methods concerning VIIRS flare detections

  - [pages 8-9] *“For all VIIRS flares detections, we use the total reported volumes of gas flared for 2021 from flares detected using the VIIRS instrument (Elvidge et al. 2016) multiplied by the observed flare destruction efficiencies and percentage of unlit flares from Plant et al. (2022) to calculate annual methane emission rates from this source. As previously stated, our empirical measurements are largely located outside of oil/gas basins where the majority of VIIRS flare detections are located (i.e. Permian, Eagle Ford, and Bakken), but we cannot discount the possibility that there are instances of double-counting flares measured via our ground-based*

empirical data and those detected by VIIRS. For each VIIRS flare detection, we randomly determine whether it is an unlit or lit flare based on the basin-specific percentages of unlit flares reported by Plant et al. (2022). If a flare is determined to be lit, we use the corresponding basin-specific observed destruction removal efficiencies as reported by Plant et al. (2022) multiplied by the corresponding annual total volume of gas flared and convert to an emission rate. The basin-specific observed destruction removal efficiencies are estimated through a fitted normal distribution using the mean and standard deviations modeled from the 95% confidence intervals presented in Plant et al. (2022). If a flare is determined to be unlit, we use a destruction removal efficiency of 0%. For VIIRS flare detections located outside of the Bakken, Eagle Ford, and Permian basins, we used the total CONUS averaged flaring efficiencies destruction removal efficiencies of 95.2% (95% confidence interval: 94.3 – 95.9%) and percentage of unlit flares of 4.1% as reported by Plant et al. (2022).”

General Comments:

1) Methodology: What is the benefit of using a bootstrapping approach? Is the bootstrapping solely to provide confidence intervals, or is there an additional benefit?

The bootstrapping in this work is used for developing a probability distribution of a given facility emitting below the method LOD (i.e., 0.1 kg/hr), or being a top 5% emitter (in some cases). It is one of several ways in which we incorporate different facets of uncertainty into the estimates. The benefit of utilizing a bootstrapping approach is to include uncertainty associated with the chance of a facility being above/below the method LOD and a top 5% within the modeled outputs, which is then reflected in the estimated emission rate distributions. We would also point to a previous response regarding the stochasticity of facility-level emission rates for oil/gas facilities.

- [Previous response] “Multiplying an average methane emission factor by the number of facilities can produce a rough estimate of total methane emissions, but is not a suitable approach for characterizing facility-level methane emission distributions, which must account for the stochasticity in facility-level methane emissions profiles and related uncertainties (for references that discuss this stochasticity: https://www.nature.com/articles/ncomms14012, https://pubs.acs.org/doi/full/10.1021/acs.est.6b04303, https://pubs.acs.org/doi/full/10.1021/acs.est.8b03535) Developing robust methods for characterizing such distributions at the basin- and national-scale is the focus of this work. While we do present estimates of total emissions estimates at the national/basin/aerial spatial scale, these are a by-product of our methodology and not the main findings, which are the detailed distributions of individual facility-level emission rate and the large majority contribution of total emissions linked to an aggregate of smaller emitting sources (i.e., the distributions presented in Figures 3, 5, and 6).”

- We have added some new text in the Methods that describes some of the reasoning behind these methods.
  - [page 8] ***"Next, we remove the empirical measurements below the LOD and use bootstrapping with replacement (n=1,000) on the above LOD empirical measurements to determine the probability of an emitting facility being in the top 5% (i.e., 95th percentile or above of empirical measurement data) or bottom 95% (i.e., 95th percentile or below the empirical measurement data) of emitters, except for processing plants and T&S compressors which had too few measurements (n=20 and n=50 respectively) to distinguish between the top 5% and bottom 95% of emission or loss rates. This pseudo-random selection of a top 5% emitter within each facility category accounts for the functional definition of abnormally large emissions (i.e., super-emitters) that can be observed in all facility categories (including well sites in different production bins) (Zavala-Araiza et al. 2015, Brandt et al. 2016)."***

My criticism is that many of the same results and conclusions are achieved without this analysis or less complex approach. Conclusions 1 and 2 can be drawn solely from the prior EF distributions. Conclusions 3 and 4 require knowing the number of facility types and production rates (taken from Enverus, OGIM database) but also do not require the monte carlo bootstrapping. Same critique for sections 3.1, 3.2, and 3.3.

We have included the following response as an expansion to an earlier comment by the reviewer below which we believe addresses the main concerns:

- Multiplying an emission factor by the number of facilities would not be able to provide individual facility-level emissions, which is our focus, but rather an aggregated total of emissions without any information on how much methane is being emitted above/below a given emission rate threshold. While we do present estimates of total emissions estimates at the national/basin/aerial spatial scale, these are a by-product of our methodology and not the main findings, which are the detailed distributions of individual facility-level emission rate (i.e., the distributions presented in Figures 3, 5, and 6).

  If an emission factor (i.e., average emission rate) and the associated parameters (i.e., standard deviation of the mean) are applied to each individual facility to provide an independent emission rate, and this is repeated for all facilities in the CONUS which are then combined together, then this would form the base of our methodology. However, this approach would not produce an accurate distribution of emission rates because the emission factor would still need to account for facilities that are non-emitting, the fact that different facility categories (including different production ranges of well sites) emit at different rates, that the available empirical measurement data for well sites does not share the same production characteristics as the entire CONUS, and that the measurements used to derive this emission factor have inherent uncertainties. After accounting for these factors (and others) we begin to reconstruct the methodology used in our work, which we believe is essential to produce the findings we present.

As an example, the main conclusion of the authors (Conclusion 1, L712) is that 72% (70% as stated in abstract, L22) of total emissions are from facilities that emit less than 100 kg/hr. This is in fact buried in the last table of the supplement, which states the prior emission distribution and shows that 72.7% emissions facilities are from these "small" emitters. The posterior is unchanged from the prior, which is good since MC bootstrapping in this approach shouldn't change the center value.

In this instance, the table that is being referenced (Table S5) represents the posterior (i.e. the estimated individual facility-level emission rates), not the empirical data (i.e., the "prior"). The table's purpose is to easily highlight the information presented in Figure 3 in terms of different emission rate magnitudes and their associated contributions to total oil/gas emissions.

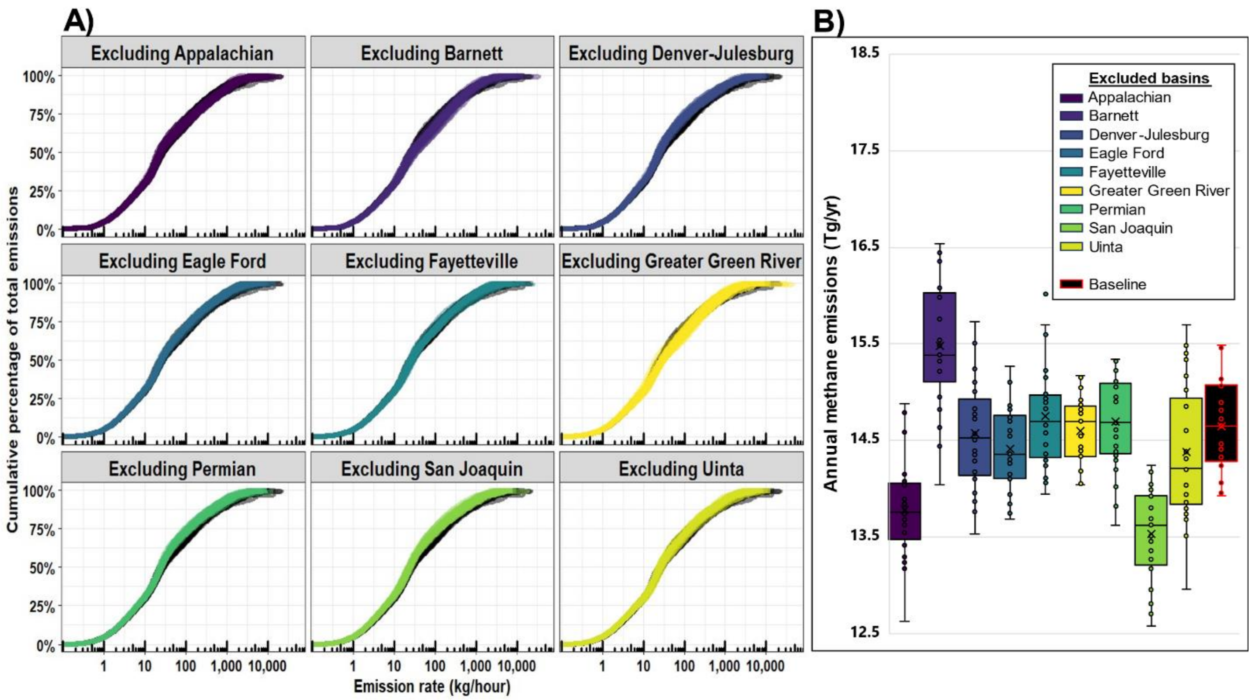
- We edited the caption of Table S5 for better clarity in SI to highlight that it is showing the resulting estimated emissions and does not represent the empirical measurement data, and have also moved the table to the front of the SI Tables.
  - ***"Table S1: Breakdown of total oil/gas methane emission for the CONUS in 2021 contributed from different magnitudes of methane emission rates with the corresponding percentage of total facilities responsible for those emissions. These results show a breakdown of the emission distributions curves presented in Figure 3 of the main text."***

A MC bootstrapping technique may be more interesting if applied to randomly select which EF studies to include. For example, if 6 of the 11 studies of facility category "Well Sites" listed in Table S1 were randomly selected for each simulation, then we might assess dependence based on regional dependence of studies, sampling/analytical methodology, etc. Indeed, regional differences are maybe observed, e.g. loss rates of 0.90% for Appalachian and Greater Green River regions (Omara et al, 2018) compared to >4.5% for San Joaquin and San Juan regions, but the variance within the regional populations precludes saying these loss rates are different (based on a Tukey test). Could the Tukey test be run on the *log_{10} (loss %)*, given that these appear to be lognormally distributed in Figure 1?

We agree with the reviewer that this would be an interesting sensitivity analysis to perform, so we conducted two additional tests (shown in the responses below). The tests examine the impacts of 1) Reducing the number of empirical measurement data to be used in the estimates and 2) Eliminating data from a given oil/gas basin/region (well sites only given limited data on regions for midstream assets. In order, the sensitivity tests show 1) reducing the number of empirical measurements only increases uncertainty bounds but does not affect the overall emission distributions or total emissions estimates 2) excluding data from certain regions does not generally impact our results for emission distributions or total emissions, even for the Appalachian where the majority of our empirical measurement data are located, with the analysis performed on all 9 basins varying the emission distributions by +/-3-4% and our total estimates by +/-6-7%, which is well within our stated uncertainty bounds.

We have since removed the Tukey tests due to this new suggested sensitivity analysis related to the impacts of excluding empirical measurement data from given oil/gas basins. We believe this revised approach better characterizes the uncertainties related to the spatial distribution of measurement data we use in our estimates, since it also includes additional factors such as the relative counts of facilities, differences in oil/gas production, and the number of empirical data available from each region.

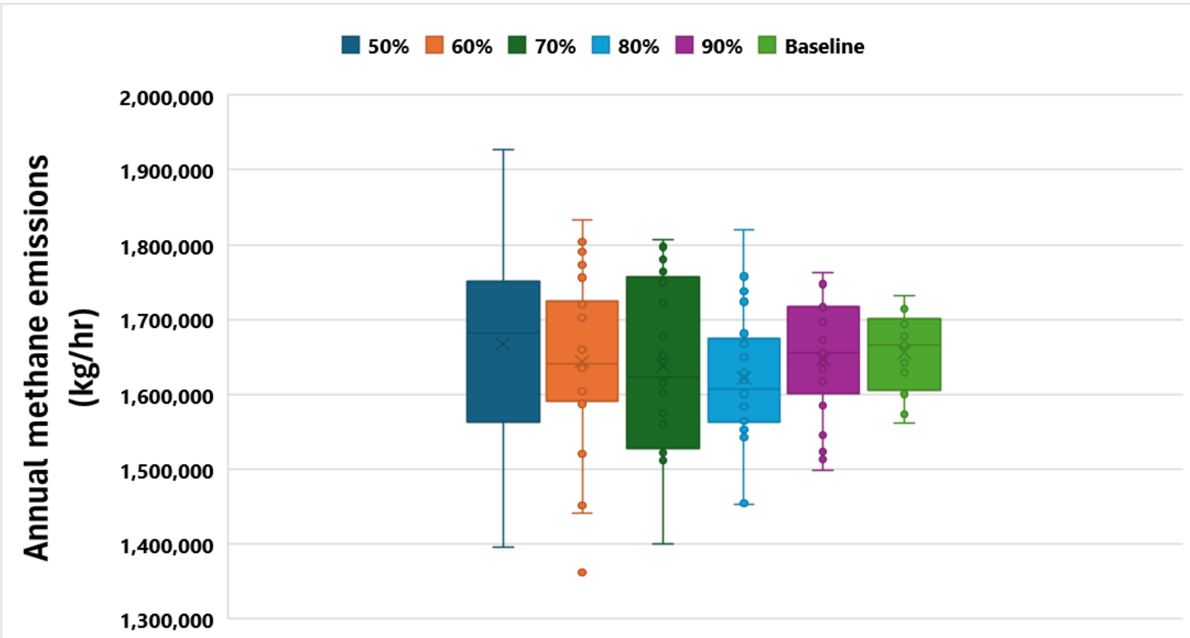
- The Tukey test figures have been replaced with a new Figure S9 displaying the resulting changes in total methane emission distribution curves (A) and total methane emissions for the CONUS (B):



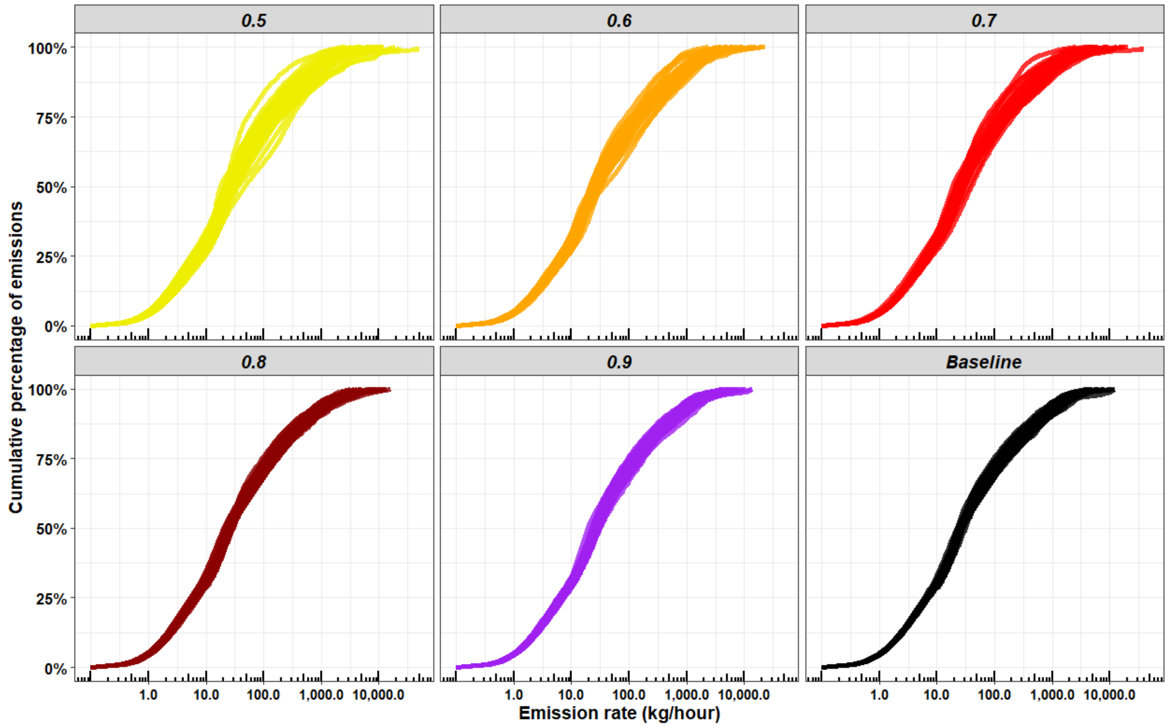
- *"Figure S9: A) Sensitivity analysis of the effects of excluding empirical measurements from a single basin showing the impacts on oil/gas methane emission distributions for the CONUS. 25 emission distribution curves are presented for each basin (colored lines) exclusion scenario with comparisons to the entire dataset of empirical data (black lines). B) Sensitivity analysis of the effects of excluding empirical measurements from a single basin showing the impacts on total oil/gas methane emission estimates for the CONUS. Each box and whisker plot contains 25 estimates of total methane emissions colored according to the oil and gas basin from which empirical measurements were excluded. The black boxplot with red outlines shows the baseline scenario, which has no empirical measurement data removed."*

- New text has been added in the Discussion section:
  - [page 28] *"In addition, there are variations in the number of production well site empirical measurements among oil/gas basins (Table S3), although a sensitivity analysis shows that excluding data from individual oil/gas basins does not significantly impact our results (Fig. S9)."*

- Below is a figure illustrating the effects of reducing the sample size of empirical measurement data on total methane emission estimates for the CONUS. 25 iterations are performed for each barplot. “Baseline” refers to the full suite of empirical measurements as we use in our manuscript. Note that we see increased variation in results from a 50% reduction of measurement data relative to our baseline, with no significant change in the average total methane emissions



- Below is a figure illustrating the effects of reducing the sample size of empirical measurement data on methane emission distributions for the CONUS. 25 iterations are performed for each emission distributions. "Baseline" refers to the full suite of empirical measurements as we use in our manuscript. Note that we see increased variation in results from a 50% reduction of measurement data relative to our baseline, with no significant change in the overall emission distributions



2) What is the 95% CI for the total national CH4 emissions?
The 95% confidence intervals for national CH4 emissions are 14.6 (12.7 - 16.8) Tg/yr. We have added these uncertainty ranges in the abstract.

3) Data Availability: Data should be made available in a publically accessible, reliable repository and linked, preferably, through a DOI per EGUsphere instructions.
We agree with the reviewer and are making all the data publicly accessible used for the emission distribution curves presented in Figure 3 (~350,000 rows by 500 columns) with coordinate/facility type/basin level data removed due to data sharing restrictions based on our activity data (i.e., Enverus).

These data are now available for download at Zenodo (link: https://doi.org/10.5281/zenodo.13314532), which now referenced in the Data Availability section.

Ideally, I would also prefer to see a table or reference section in the supplementary that has direct links, references, etc to the data from other studies used in this manuscript. This would be the data references in Table S1, plus Lan et al. 2015.

We have added this information to Table S1 (i.e. links in SI Table S1), including Lan et al. (2015) which was left out due to a clerical mistake and Goetz et al. (2015).

Specific Comments:

Would be useful to state what the LOD of Bridger GML is here.
We have included references to the Bridger LOD stated by Kunkel et al. 2023 in this section (i.e., 3 kg/hr).

“1,898 facility-level…” I am a bit confused since Table S1 only sums to 1866 observations.
We thank the reviewer for bringing this to our attention, we neglected to include the Lan et al. (2015) and new Goetz et al. (2015) measurements in this total. This has been fixed in Table S1.

“high-emitting intermittent are included” à “high-emitting intermittent sources are included”
Text has been corrected.

There appears to be a linearly decreasing relationship between the loss % and production rates for well-sites (facility category 5-9). Is this real? Is there a reason to include this in the facility-level model?
Yes, the measurements being shown in this figure are the empirical measurement data we use in our work. This decreasing relationship between production and production-normalized loss rates exists and has been shown/used in prior studies (e.g., Omara et al. 2018). However, the relationship between production and loss rates is weak (i.e., visible in a log/log plot), but useful for better constraining the extrapolation of emission rates to the full population of well sites in CONUS. A more detailed explanation of this relationship is explained in Cardoso-Saldana et al. (2020) (link: https://pubs.acs.org/doi/pdf/10.1021/acs.est.0c03049), but to briefly summarize: emissions from high-producing wells are a combination of production-independent leaks (i.e., fugitive emissions from leaks from pipes, flanges, etc) and production-dependent emissions (i.e., condensate flashing). As the production of a well drops exponentially over time, the associated emissions from production-dependent leaks also drop, whereas the production-independent emissions persist. We utilize this empirically observed relationship between facility level methane loss rate and production to constrain emission estimates for specific production cohorts, where, in general, loss rates are lower for higher producing facilities, and vice versa. This is an important component of our model as the distribution of well site productivity varies across basins.

“… gas flared for 2021 by Elvidge et al. (2016)… efficiencies from Plant et al. (2022)” Are these the correct references? It seems unlikely that Elvidge et al (2016) published gas flaring for 2021.
We thank the reviewer for pointing out this mistake. We have re-written the section to clarify that the Elvidge et al. (2016) reference is meant to provide background on the VIIRS detection instrument and is not being used to draw in actual gas flared values.

“…production well sites that we use in this work generally do not show significant…” à “… basin-to-basin, production well sites in …”
Corrections made

"… Ravikumar et al. (2019) From …" à "…Ravikumar et al. (2019). From…"
Corrections made

What do the error bars represent? 95% CI?
Yes, these bars all represent the associated 95% confidence intervals. We have added new clarifying text in the Figure 4 description.

"our results show the essentiality of expanding beyond solely on super-emitter mitigation". Some sort of grammatical correction needed.
Agreed, corrections made.

It would be nice to provide the sample size of these studies.
Agreed, we have included sample sizes for these studies in other countries.

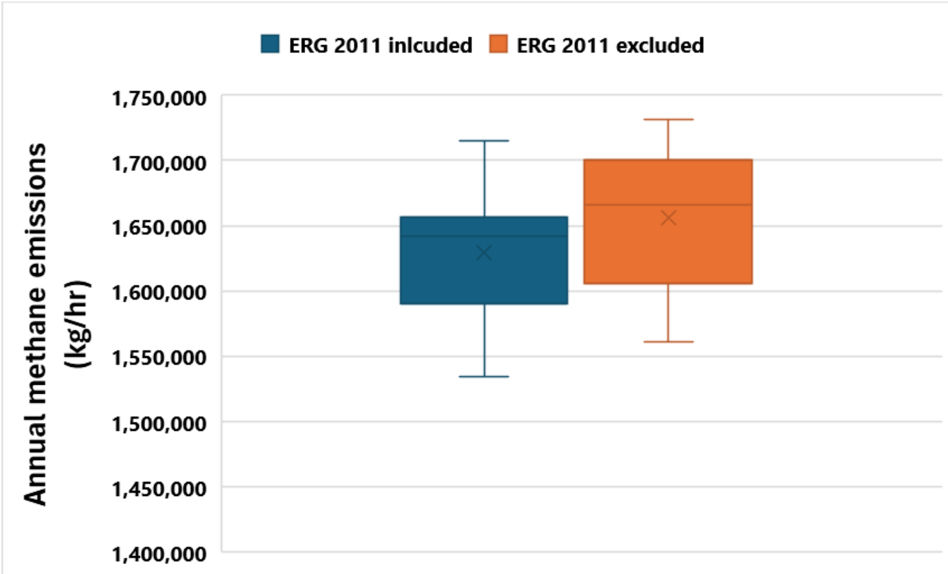
Appears to be missing a reference to Lan et el. 2015.
Lan et al. 2015 reference added, and number of measured facilities corrected throughout manuscript.

There are several other references used by Omara 2018 not included in this study. (Goetz et al 2015, ERG 2018).

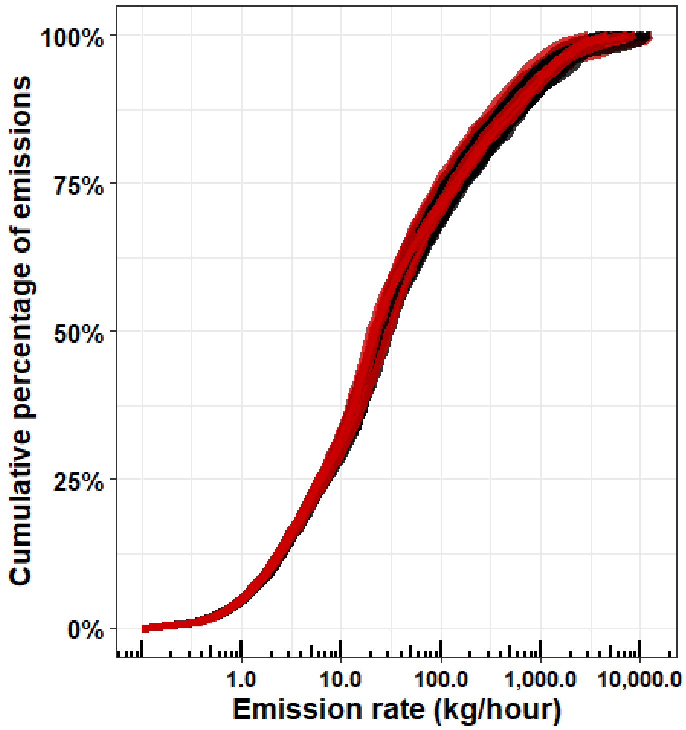
With regards to Goetz et al. 2015, we have since included these data into our dataset of empirical measurement data. The addition of these data (n=3) do not change our results and main findings in any significant way.

We only know of the ERG 2011 study, but we would be happy to investigate a more recent component-level study if another exists. For the ERG 2011 dataset, we decided to exclude it given that it is an older dataset (10+ years) and a compilation of component-level measurements, which we acknowledge in our work may underestimate total facility-level emissions given that there is no guarantee that all emitting components were measured. While we do include some component-level aggregation studies in our work, both of those studies (Riddick et al. 2019, Deighton et al. 2020) provide measurements within the past 10 years. However, we did perform a sensitivity analysis on the effects of including versus excluding the ERG 2011 data and found no change in our model results for both total emissions and the emissions distributions (see below).

- Comparison of total CONUS oil/gas methane emissions when including/excluding ERG 2011 empirical data for 25 estimates each.



- Comparison of emission distribution curves for total CONUS oil/gas methane emissions when including (red)/excluding(black) ERG 2011 values for 25 emission distribution curves each.



The total number of well sites for the Barnett basin is 32 wells less than the sum of the bins. I assume this is the 32 wells measured by Lan et al. (2015) that was not included in Table S1.

That is correct, this was a clerical error on our part. The mistake has been corrected to include the 32 wells from Lan et al. (2015), in addition to the new 3 measurements from Goetz et al. (2015) as previously suggested by the reviewer.