Reply to community comments.

We thank the reviewers for the interest in our paper and the generous comments. We are glad that this paper has been useful for graduate coursework. Reviewer comments are provided below in black with our responses in blue.

This review was prepared as part of graduate program course work at Wageningen University, and has been produced under supervision of dr. Ingrid Luijkx. The review has been posted because of its good quality, and likely usefulness to the authors and editor. This review was not solicited by the journal.

The paper by Yu et, al. entitled "A high-resolution satellite-based map of global methane emissions reveals missing wetland, fossil fuel and monsoon sources" presents a quantification of the 2018-2019 global methane budget, based on space-borne TROPOMI observations. Methane emissions are derived from the TROPOMI observations by coupling multiple 4D-Var adjoint inversions with a newly developed spatial downscaling approach. This enables the identification of previously missing or underestimated methane emissions from fossil fuel and wetland sources.

This research presents a new downscaling method that is applied to convert the GEOS-chem model output to a 0.1º × 0.1º resolution, using combined spatial information from the TROPOMI observations and from the prior estimates. This method enables very specific allocation of emission hotspots. In the study, OH is used as an additional constraint in the inversions, as recommended by Saunois et al. (2020). This advances previous studies, which often co-optimized methane sources and sinks by using methane data alone. The results section of the manuscript is well-written and discusses all the source areas in-depth, also suggesting possible underlying reasons for the found underestimations in prior inventories. The research reveals some interesting results regarding emission hotspots that were missing from prior inventories. That being said, I do have some remarks that could be addressed before publication.

We thank the reviewer again for the positive comments.

1) If I understand it correctly, the aim of the research is to quantify the 2018-2019 global methane budget and determine missing and underrepresented emission sources. However, the authors mainly present how much the prior estimates are underestimated compared to their findings, making the results section an evaluation of their one specific chosen set of prior inventories. This approach results in a high dependency of the research on the choice of the specific prior estimates. If other prior inventories were chosen, the underestimations and hotspots that the research now revealed would likely be very different, because for instance, as was pointed out in the introduction, two of the most commonly used anthropogenic emission inventories (EDGAR v5 and GEPA) are uncorrelated at a 0.1º x 0.1º resolution. To overcome this issue, I would suggest to shift the focus of the results section from the discrepancies in the specific prior to the obtained absolute values of the methane budget.

We do include discussion of the changes relative to the prior inventories as these are widely used and their strengths and shortcomings are therefore of broad interest. However, we also already include discussion of the absolute flux amounts as the reviewer requests. For example, Table 2 and Section 4 list and discuss the derived sectoral fluxes both globally and by specific country. Section 5 then states the absolute wetland flux amounts for every region discussed (globe, Amazonia, central Africa, South Sudd, North America). Absolute flux amounts are also provided at relevant points throughout Section

6. In our view this provides a balanced discussion between the derived adjustments and the fluxes themselves.

A good addition would then be a comparison of these results to independent measurements, such as the ObsPack or TCCON observations, or a comparison to other studies that also use inverse models to characterize the methane budget, such as Saunois et al. (2020).

We already compare the results to ObsPack and TCCON observations, as follows:

"Ground-based methane column (XCH4) observations from the TCCON network (GGG2014 (2014)) show comparable improvements over the prior for both the fixOH and optOH solutions (and for their individual member inversions), with 71% (from -12.9 ppb to 3.8 ppb) and 66% (to 4.3 ppb) mean bias reductions, respectively (Figure 3, Table S1). However, global in-situ measurements from ObsPack (near-real time version v2.0; (2021)) reveal a 93% (from -13.8 ppb to -0.9 ppb) absolute mean bias improvement for the fixOH framework compared to just 39% (to 8.4 ppb) for optOH (Figure 3, Table S1). Figure 3 further shows that the optOH solutions overcorrect the prior negative bias with respect to ObsPack, providing a first piece of evidence for a methane sink underestimate in this inversion."

We have now updated our discussion of the derived national methane budgets to include comparisons with both Worden et al. (2021) and Qu et al. (2021):

"Eight of the ten nations in Table 2 (China, India, US, Russia, Brazil, European Union, Pakistan, Indonesia) are likewise identified by Worden et al. (2022) as among the top ten anthropogenic emitters globally. Our inferred anthropogenic fluxes for the US and China agree well (within ~10%) with the GOSAT-based results from Worden et al. (2022) and with the GOSAT+TROPOMI results from Qu et al. (2021). Anthropogenic emissions derived here are likewise within 10% of the Worden et al. (2022) results for India and the European Union, with both studies lower (20–50%) than Qu et al. (2021). Our results for Russia and Iran are 21–28% higher than the GOSAT-based estimates, mainly reflecting oil, gas, and coal emissions, and ~40% lower for Brazil, mainly due to livestock. Emissions for Pakistan and Indonesia agree to within 1% for the TROPOMI- and GOSAT-based results (Worden et al., 2022). However, we find here that anthropogenic emissions from Bangladesh (7 Tg/y versus a prior of 4 Tg/y) are 3× higher than the GOSAT estimate (2 Tg/y), while adjacent emissions from Myanmar (4 Tg/y) are half the GOSAT estimate. Worden et al. (2022) conclude that the GOSAT-derived emissions for Myanmar are anonymously high due to impacts from their prior assumptions; we attribute much of that flux to Bangladesh and show later that it mainly arises during the South Asian monsoon."

We have also added new comparisons to other prior top-down studies as suggested:

"Our derived global wetland fluxes are ~20% higher than previous GOSAT-based estimates (145-148 Tg/y: Ma et al., 2021; Zhang et al., 2021), with similar latitudinal distribution to that found by Ma et al. (2021)."

and

"The 2019 global anthropogenic methane emissions obtained here are modestly (12%) higher than GOSAT-based results for 2010–2018 (336 Tg/y; Zhang et al., 2021), with both results pointing to higher-than-predicted biotic emissions (consistent with isotopic constraints; Nisbet et al., 2016)."

2) The authors take an ensemble mean of the 4 different inversion formalisms to calculate the emission corrections, while their previous research showed that some of them perform better for different purposes (Yu et al., 2021a). The different allocation of emission hotspots that are found through the different inversions are already nicely discussed in the results section, but the emission corrections are subsequently still calculated as the multi-model mean. I would like to see a more in-depth discussion of why the authors chose this approach, and why for instance the classical SF inversion is not left out here, since it is highly biased towards areas where the prior estimates of the emissions are high, and therefore likely makes the calculated underestimations of the prior estimates too small (Yu et al., 2021a). The BI inversion approach provides the best spatial distribution of all inverse approaches, while the EE inversion performs best in recovering large missing sources (Yu et al., 2021a). In the calculations of the hotspot emissions that are missing from the prior inventories, it is therefore probably better to use the EE inversion instead of the ensemble mean. Table S1 could also be used in this discussion, since it summarizes the performance of the different inversion formalisms, while the statistics presented here are currently not used in the text.

We thank the reviewers for the interest in our previous work (Yu et al., 2021a). We chose to use the ensemble mean as our base-case solution as each individual inversion strategy has unique advantages. For example, in Yu (2021a) we found that the SF inversion exhibited the best performance for large sources already present in the prior, while the BG inversion yielded the highest spatial correlation with the true fluxes, and the EE inversion had the best performance for identifying missing sources. Furthermore, our previous OSSE was performed at ~25 km and the employed inventories had large spatial disparities at that scale. In this study, our inversions were performed at 2°x2.5° (~200 km) and the spatial inconsistencies are reduced at this larger scale.

3) The section of the development of the novel downscaling method could be more extensive. Since a new method is presented here, it's very important that it is well-described. First of all, I would like to see the argumentation on why there was a need for a new downscaling method, and why previous downscaling methods were not suitable. Yu et al. (2021b) could be consulted, who present a nice review section of related work on spatial interpolation and downscaling of airborne pollutants.

We have now added new text comparing our downscaling approach with other methodologies to better motivate its use, as suggested:

"Compared to existing emission downscaling methods that rely on prior and posterior error covariance estimates (Cusworth et al., 2021), or are based solely on satellite data (Liu et al., 2021), our approach is unique in combining the prior emission information (and its uncertainty) with the oversampled TROPOMI observations themselves. Variable weighting between these terms permits greater influence from the observations when the prior emissions are more uncertain. The method thus assumes robust prior error estimates, a caveat that also applies to Cusworth (2021) and similar methods."

Further motivation is provided earlier in Section 2.6, and this has been slightly expanded:

"We present here a new method to spatially downscale the satellite-derived emissions for potential use in models. This downscaling is necessitated by the fact that the current GEOS-Chem adjoint model does not have global simulation capability at finer than 2° × 2.5° resolution. Furthermore, each of the 2-year inversions performed here required >12,000 CPU hours (>80 days on multiple processors) to converge, making higher-resolution optimizations computationally impractical. However, the inventories employed as prior, as well as the TROPOMI observations themselves, contain information at much finer scales (e.g., 0.1° × 0.1° and 7 × 7 km^2)—and thus contain additional high-resolution constraints that are neglected by the 2° × 2.5° inversions. We therefore leverage this information to spatially downscale the optimized emissions to 0.1° × 0.1° via …"

Also, since the downscaling method is presented as novel, a proper evaluation of its accuracy is very important. I therefore wonder why the authors chose to perform the OSSE only for one area, for the duration of one month and at a resolution of 0.25º x 0.3125º, and subsequently chose to use a 0.1º x 0.1º resolution in their further research based on this OSSE. The representativeness of this one OSSE for the whole research should be better discussed and possibly expanded, since the validity of the research is dependent on this outcome.

This OSSE evaluation approach was selected for several reasons. First, the GEOS-Chem model does not have global 4DVar optimization capability at 0.25°×0.3125° and does not have any optimization capacity at 0.1°. Even if it did, the computational cost would be far too high to run a global 4DVar evaluation at either scale. Second, the North American domain was selected as it includes all relevant source types and because we had simulation output available from Yu et al. (2021a).

To address the comment about accuracy we have now added a new downscaling bias reduction analysis, as follows:

"Figure S8 shows that the downscaled OSSE solution reduces the prior bias by 17%–56% for sources exceeding 1000 kg CH4/box/day (accounting for 99% of the domain-wide emissions) when not subject to transport error. In the presence of transport error, the downscaling method has limited success for the very largest sources (>2×105 kg/box/day), but nevertheless exhibits strong bias reduction (21%–50%) for sources between 1×103–2×105 kg/box/day (96% of domain-wide emissions)."

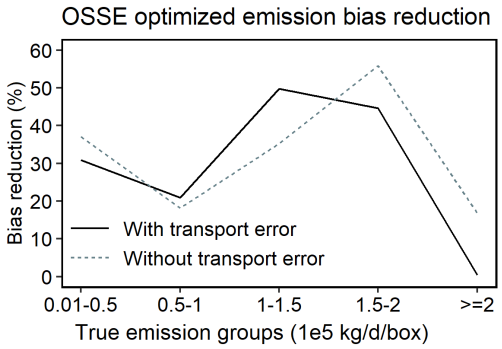

Figure S8. Downscaling bias reduction as a function of emission magnitude, based on 1-month Observing System Simulation Experiments (OSSE) over North America (see main text for details).

4) After the results section, I would suggest to include a section where the uncertainties in both the TROPOMI data and the prior estimates is discussed, since the research is very dependent on both, and therefore also dependent on errors in the data. Also, the methods could be further discussed in this section, such as implications of the downscaling of the optimized emissions, and the use of the different inversion formalisms.

Uncertainties in the TROPOMI data are already discussed in Section 2.1 based on both instrumental specifications and validation statistics versus GOSAT and TCCON observations. Our treatment of prior uncertainties is then discussed in detail in Section 2.4. We prefer not to duplicate that information in a separate section. The different inversion formalisms are already included in the discussion section based on the fact that we are using them to define our uncertainty range and to identify areas with consistent adjustment direction across inversions. The downscaled results are used later in the paper to map the advance and retreat of East Asian emissions.

5) In my view, the knowledge gap could be further specified in the introduction. The novel aspect of the methods is already highlighted well by stressing the importance of including OH constraints, which many previous studies did not include. However, a section on prior knowledge about hotspots and emission sources that are often underrepresented in prior estimates is missing, including how the research is still of added value to this. Hu et al. (2018), who used TROPOMI to map methane column concentrations for instance also observed the underestimated hotspot of the Sudd wetlands and Venezuela. Lu et al. (2021) performed an inversion study using GOSAT data and also revealed missing spots in observational data, but on a far coarser resolution than this study. I suppose that the authors mainly add to this because of the far higher resolution of the TROPOMI data they use, combined with the downscaling method, making it easier to pinpoint emissions to more specific locations.

Rather than expand the introduction (which would end up duplicating information provided later in the paper) we have instead expanded our comparisons to previous findings in the results section, as described above and in our replies to the other reviewers. We have added a citation of Hu et al. 2018 in the Sudd discussion as suggested.

6) The authors nicely present the main underrepresented sources and missing hotspots in the conclusion, but a section with the further implications of these findings is missing.

We believe that the implications are sufficiently covered in the results and conclusions and that an additional section is not necessary.

In the last section of the conclusion, some recommendations for future research are given (lines 561-564), but the statements include no references confirming that the addition of datasets of CO, methyl chloroform and formaldehyde would indeed improve future inversions. Also, the novel downscaling method is not mentioned in the conclusion, while this method is probably also relevant for further research.

We have added some references supporting the use of CO, MCF, and HCHO, as follows:

"We address the issue in this work by employing validated OH fields from a chemical transport model, but future methane inversions can benefit from incorporating additional datasets (e.g., CO, methyl chloroform, formaldehyde) as constraints on the methane sink (McNorton et al., 2016; Rigby et al., 2017; Turner et al., 2017; Wolfe et al., 2019)."

The new downscaling methodology is in fact mentioned in the first paragraph of the conclusion section.

Minor comments

Title: The current title is appealing because it directly mentions the new findings, but in my view, it does not cover the whole scope and innovative aspect of the research. I would consider changing the title to something like: "A high-resolution global map of methane emissions inferred from an inversion of TROPOMI satellite data reveals missing emission hotspots and previously underestimated sources."

We thank the reviewer for this suggestion, but we prefer our current shorter title.

Line 46: For a better overview of the previous research, I would elaborate here on what the conflicting reasons are for methane increase apart from the emission increase over tropical regions, such as an increase in emissions in the energy sector, an increase in wetland emissions, and a decrease in mean OH (McNorton et al., 2018).

We thank the reviewer for this suggestion. In the introduction we are aiming to provide a robust overview of the state of science while still being concise. We believe the current text provides a suitable broad-level overview and does discuss the uncertainties arising from OH.

Line 89: Please include the Sentinel-5 precursor/TROPOMI Level 2 Product User Manual Methane as a reference for requiring quality filter > 0.5: https://sentinel.esa.int/documents/247904/2474726/Sentinel-5P-Level-2-Product-User-Manual-Methane.pdf/1808f165-0486-4840-ac1d-06194238fa96

We have added this citation as suggested.

Line 96: Apart from mentioning the slope, please report the R2 as well here as a measure for agreement (R2 = 0.67).

We have added these values as suggested.

Line 117 - 128: Please elaborate on why these specific prior estimates are chosen, and perhaps also elaborate on how these datasets are constructed (by models/measurements)?

We selected these inventories as they are commonly used and represent the current state-of-the-science. We refer the reader to the cited papers for more detailed descriptions.

Line 118: Why did the authors chose to use the UNFCCC inventory from 2016? The new version from 2019 might be more representative for the study period.

At the time we performed the inversion, that new version was not available. The global difference between these two versions is small for gas (24 Tg/y vs. 22 Tg/y) and coal (31 Tg/y vs. 33 Tg/y)

emissions, while oil sources decreased from 42 Tg/y to 26 Tg/y, mainly due to Russia. We have now added a discussion of this point to Section 6 as follows:

"Indeed, subsequent revisions (year-2019; Scarpelli et al., 2022) to the UNFCCC-2016 inventory used here have strongly reduced fossil fuel emission estimates for Russia (e.g., 21 Tg/y from oil in year-2016 vs. 2 Tg/y in year-2019) due to updated emission factor assumptions."

Line 153: Please give a reference or explain why 50% uncertainty in the remaining sources is chosen.

We have added a citation as requested.

Line 162: It would be good to explain here how the OH sensitivity study is exactly performed, and specifically state where in the formula of the cost function the different uncertainties are used.

We have clarified that the prior uncertainties for OH are included in Sa.

Line 187: I wonder why the authors chose the values of 10% and 90% for the weight of the prior and the background respectively. Yu et al. (2021a) used 50% and 50% in their example of this background increment inversion formalism. Is this determined with sensitivity simulations similar as in the OG inversions? Please explain.

Our previous OSSE was performed on a 25 km grid. At this resolution, the spatial distribution of prior emissions is highly uncertain. In this study, our inversions are performed at 2°x2.5° and the spatial distribution of prior emissions have higher fidelity, justifying the reduced weighting of the flat prior.

Line 271: "Our 2019-2018 … growth rate acceleration": please elaborate on the implications of this statement on the findings that are presented in this paragraph.

We have revised this text as suggested:

"Our 2018–2019 analysis timeframe also spans an El Niño, which has been tied both to global OH decreases and to methane growth rate acceleration (Anderson et al., 2021; Turner et al., 2018a)—further complicating a differentiation between the fixOH and optOH solutions."

Line 314: Could the authors further explain here why the locations in the boxes of figure 2b were chosen for the analysis? This is probably because TROPOMI observations differ from the prior estimates in these areas. But when looking at the map, I see that this is for instance also the case for northern Italy and the Southeast US. Why are these areas not discussed?

We selected areas with substantial methane emissions that cover a range of source types and reveal significant model-measurement discrepancies. In the interest of length, we are unable to examine every global region.

Line 317: If I understand it correctly, the average yearly source and sink values for the years 2018-2019 that are presented here are not based on two full yearly cycles. The timeframe of the analysis only spans from 05/2018 - 10/2019. However, figure 4 indicates that the sources and sinks show seasonal variation. To retrieve yearly average values for the sources and sinks, these values can't be just averaged over a 1.5 yearly cycle. I would recommend to take these average values over one full yearly cycle, for instance from 10/2018-10/2019.

Thank you for the suggestions. We have now clarified in the manuscript that "annual values discussed later are for 11/2018–04/2019 plus the average of 05–10/2018 and 05–10/2019."

Line 412: I would move the explanation of figure 5c to line 396, since that is where the figure is first mentioned.

We thank the reviewer for the suggestions. At line 396, we explain the emission corrections with comparison to previous studies. In line 412, we further discuss the underlying emission drivers. We think the current layout is more clear and have left it as-is.

Line 442: Since these missing hotspots are one of the main outcomes of the research, the authors could consider to give their more exact locations, instead of only mentioning the countries.

We have added an additional reference to Figure S16 for a clearer view of these locations.

Line 447: I wonder how the hotspots can be missing in the UNFCCC inventory and show up in the EIA, since it seems like the UNFCCC is based on the national activity data from the EIA (Scarpelli et al., 2020). Is this because the authors used the UNFCCC data from 2016, and these activities were maybe still unknown at that time? Please explain this here, or as I mentioned before, consider using the updated UNFCCC inventory from 2019.

The precise reasons that these sources would be missing from UNFCCC but present in EIA are not apparent to us and would need to be the subject of future work. As noted the updated UNFCCC version was not available at the time we performed our inversions, but we have added discussion of the UNFCCC updates to the manuscript.

Figure 4: Please consider to make figure 4a-d larger, since the dots are very hard to see. Figure 4e is currently not referred to in the text. Also, I wonder why only the FixOH emission is shown here, and not the loss. I would either remove the fixOH emission from this plot, or include the loss as well.

We have rotated Figure 4 to enlarge it and improve visibility as suggested. We have also added a reference to Figure 4e as requested. The fixOH loss is shown in Figure 4e since it is the same as the prior, this is now clarified in the figure caption.

Figure 5: Figure 5a and 5c show information from previous research, while figure 5b shows main findings of the research. I would therefore suggest to make figure 5b a separate figure.

Figures 5b and 5c are in fact both based on findings from this research, and we included 5a to provide context for interpretation.

Figure S9: In my opinion, this figure could also be included in the main text, since it shows well how the outcomes of the four inversion formalism differ, and how the inversion ensemble is constructed.

Thank you for the suggestion. In the interest of length, we choose to include Figure 4 in the main text and keep S9 in the SI.

Specific comments

Line 17: Please remove "CO" here, since CO is not used as a constraint.

We have revised this text as suggested. "Employing remote carbon monoxide (CO) and hydroxyl radical (OH) observations with independent methane measurements for evaluation, we infer from TROPOMI a global methane source of ..."

Line 43: "the importance of" can be left out here.

We prefer to leave the phrasing as-is.

Line 229: Write abbreviation of OSSE out in full.

It is written out in full at first use, at the start of Section 2.4.

Line 335: The total emissions of China mentioned here (60 Tg/y) is different from the number in table S2 (61 Tg/y). Please make this consistent.

This has been corrected.

Line 342: "Europe Union" > "European Union".

Corrected.

Table 2: "Russian" > "Russia".

Corrected.

References

Hu, H., Landgraf, J., Detmers, R., Borsdorff, T., Aan de Brugh, J., Aben, I., ... & Hasekamp, O. (2018). Toward global mapping of methane with TROPOMI: First results and intersatellite comparison to GOSAT. Geophysical Research Letters, 45(8), 3682-3689.

Lu, X., Jacob, D. J., Zhang, Y., Maasakkers, J. D., Sulprizio, M. P., Shen, L., ... & Ma, S. (2021). Global methane budget and trend, 2010–2017: complementarity of inverse analyses using in situ (GLOBALVIEWplus CH 4 ObsPack) and satellite (GOSAT) observations. Atmospheric Chemistry and Physics, 21(6), 4637-4657.

McNorton, J., Wilson, C., Gloor, M., Parker, R. J., Boesch, H., Feng, W., ... & Chipperfield, M. P. (2018). Attribution of recent increases in atmospheric methane through 3-D inverse modelling. Atmospheric Chemistry and Physics, 18(24), 18149-18168.

Saunois, M., Stavert, A. R., Poulter, B., Bousquet, P., Canadell, J. G., Jackson, R. B., ... & Zhuang, Q. (2020). The global methane budget 2000–2017. Earth system science data, 12(3), 1561-1623.

Scarpelli, T. R., Jacob, D. J., Maasakkers, J. D., Sulprizio, M. P., Sheng, J. X., Rose, K., ... & Janssens-Maenhout, G. (2020). A global gridded (0.1× 0.1) inventory of methane emissions from oil, gas, and coal exploitation based on national reports to the United Nations Framework Convention on Climate Change. Earth System Science Data, 12(1), 563-575.

Yu, X., Millet, D. B., & Henze, D. K. (2021a). How well can inverse analyses of high-resolution satellite data resolve heterogeneous methane fluxes? Observing system simulation experiments with the GEOS-Chem adjoint model (v35). Geoscientific Model Development, 14(12), 7775-7793.

Yu, M., & Liu, Q. (2021b). Deep learning-based downscaling of tropospheric nitrogen dioxide using ground-level and satellite observations. Science of The Total Environment, 773, 145145.