We would like to thank the reviewers for their time and pointing out that additional discussions are needed about how our results compare with others such as the Global Carbon Project.
To summarize, our posterior results are consistent with previous published results about methane emissions based on GEOS-Chem and the GOSAT data (e.g. Maasakkers *et al*., 2019,2021; Zhang *et al*., 2021; Qu et al, 2021) as well as with a non GEOS-Chem inversion (Tsuruta *et al*. 2017) which projected emissions to biospheric versus anthropogenic. Our posterior fossil emissions are also consistent with recent reports of oil, gas, and coal emissions to the UNFCC (Scarpelli *et al*. 2022). In addition, one of the advantages of the Bayesian approach we demonstrate here is that we can test if uncertainties in the priors are causing the difference between our results and those in the literature and we can state unequivocally that it is not our choice our priors but instead a function of the satellite data and the inverse system used to relate surface emissions to these satellite data. While these comparisons were in the previous version of this paper we have attempted to make them more explicit in this version with direct comparisons in the abstract and again highlighting differences and similarities in the relevant sections (i.e. Section 3.1) and conclusions.

Reviewer 1

**Comment**: While the methods and aim are unique, several assumptions made for treatment of the priors, aggregating sectors, partitioning emissions sectors, and spatial disaggregation to country level are a concern. These assumptions lead to global posterior fluxes that are not consistent with earlier studies. For example, that anthropogenic fossil fuel emissions are 82 +/- 12 Tg $CH_4$ for 2019 are inconsistent with Saunois et al., 2020 (135, 121-164 Tg$CH_4$), the Global Methane Assessment, Schwietzke et al, 2016, etc., with the range of uncertainty not overlapping with any of these previous studies.

**Response**: As noted in the previous general response, we have added additional discussion in the abstract and Section 3.1 on comparisons between our results and those of Saunois *et al*. and Schwietzke *et al*. [edits found in new section "Comparisons to Top-Down Inversions from GCP" and "Fossil Emissions"]. Basically, we now directly compare to Saunois *et al*. for the top-down inversions presented there as well as added discussion comparing our results to Tsuruta *et al*. (2017). For fossil emissions, we now note the discrepancy with Schwietzke et al, 2016 and possible causes (upscaling isotopic measurements). We also contextualize against more recent bottom-up estimates (Scarpelli et al., 2022) and find our results are consistent with these estimates.

**Comment**: At the country-level, the underestimate in fossil $CH_4$ emissions is propagated to national scale comparisons between the inversion posterior fluxes and the inventories. The directional changes are inconsistent with findings of Stavert (https://onlinelibrary.wiley.com/doi/full/10.1111/gcb.15901) and Deng (https://essd.copernicus.org/preprints/essd-2021-235/) and Alvarez (https://www.science.org/doi/10.1126/science.aar7204).

**Response**: Thank you for highlighting the Stavert and Deng papers, which were submitted or accepted near or after submission of our paper such that we were unaware of these papers. We have added comparisons to these papers as well. Note that we can compare total anthropogenic from our Table 3 to the results shown in the Stavert and Deng papers and our results are actually quite consistent with theirs; especially for the GOSAT alone results shown in Deng *et al*.. The one outlier that we can find is India where we report much higher anthropogenic emissions than Deng *et al*. (although within the spread of in situ based estimates). We have added language to that effect in the revised manuscript.

To help the reader understand the differences, I would recommend clarifying the following components of the manuscript:

**Comment**: An explanation on how sectors are aggregated is needed, why is agriculture and fire emissions combined, this is not a standard way to show these sectors?

**Response**: We are working closely with a parallel effort to produce stock changes related to $CO_2$ which reports an "AFOLU" category which is similar to our agriculture and fires shown in Figure 5 as stated in the text. However, as noted several times in the text, the sectoral partitioning is performed using the priors shown in Table 3. To avoid confusion we add additional language when discussing Figure 5: "Sectoral attribution is based on the nine categories in Table 3; here we combine categories so that they are similar to what is being reported for the $CO_2$ based carbon inventories"

**Comment**: How was the decision made to split the sectors to those listed in Table 3? The partitioning to oil/coal/gas/seeps is likely much more uncertain than what is represented in the Table and overstates the capacity of the inversion system and confidence in the results.

**Response**: We chose those sectors which are the top-contributors to the global methane budget. A future version of this approach will add other sectors such as termites and biofuels which we chose (for now) to exclude as internal discussions suggested these components were too small and also highly uncertain such that even the location of their priors is likely not well known. As you note, an argument could be made for excluding seeps for the same reason; however the location of the seeps are likely well known even if the magnitude of their emissions is not. Updated language in Sections 2.3 and 3.1 have been added to indicate that care must be taken when interpreting our Seep emissions results.

**Comment**: Why is the geologic seepage prior of 32 Tg used? On line 286, 2 Tg are used – but on line 663, 32 Tg is used. This assumption is likely a source of bias in the FF posterior mentioned earlier.

**Response**: Our sectoral partitioning approach only depends on the total prior flux (**xa**) and not what makes up that prior flux. Also recall that we can swap in different priors (**za**) if there is evidence (i.e. observations) to suggest that we should. As there is no obvious reason to exclude the larger prior for Seeps of 32 Tg as discussed in Etiope *et al*., we used that value in our partitioning code. However, we have added additional discussion in Section 3.1 discussing how this could change our results on total fossil emissions as many of the seep locations overlap that of the coal and oil emissions. Note that adding the seep estimates in with the anthropogenic fossil emissions improves but does not make consistent our total fossil results with those from Schwietzke *et al.*.

**Comment**: Line 489 says 7 regions, line 494 says 8 regions. **Response**: Fixed

Comment: The spatial disaggregation is not clear – the inversion is made at 2x2.5 degrees, using information at 8 regions, then disaggregated to 1x1 degree. These steps are not clear.

**Response**: The reason for projecting fluxes to emissions at 1x1 degree is to 1) make a more accurate estimate of the country level emissions estimates as a 1x1 degree map better represents a countries borders and hence emissions and 2) projecting emissions to finer scale (e.g. 0.1 degree, which is the resolution of the EDGAR emissions) is computationally challenging and 3) demonstrate that the algorithm works at 1x1 degree and can be used to estimates emissions at any spatial resolution. For example, a future goal is to update the inventories at the reported spatial resolution of 0.1 degree using satellite based fluxes from both the large scale (e.g. TROPOMI, Methane-Sat) combined with high-res estimates from satellites such as Carbon mapper. Our algorithm shown here provides the approach for this combination; we have added language in Section 2.3 discussing these additional details.

Comment: A table for the global summary of posterior fluxes is needed.

Response: We do not think a posterior flux that is an analog of Table 3 is needed as Table 3 is intended to show the sectoral types and regions used as priors and as needed to ensure that we can project fluxes back to sectoral emissions in a computationally feasible manner. Our posterior emissions estimates are described in detail in the final table.

Comment: Which inventory is used in Table 3.

Response: Inventories and their references are discussed in Section 2.3 both in the text and in Table 2. We have added additional language pointing to 2.3 and Table 2 when discussion Table 3.

Reviewer 2

Major Comments

Title: The title ‘The 2019 Methane Budget...’ is a bit confusing/misleading. There is a broad and well established community effort known as ‘The Global Methane Budget’ and many readers may think this is an update from the same group. However, there is little to no overlap between the two author lists and the approaches are very different. I’d recommend changing the name to avoid confusion. In addition, highlighting ‘Each Country’ in the title seems inappropriate given that results show that the GOSAT data used constrain ~25% of the countries considered (Table 3 and discussion).

**Response**: We are not sure how to address this comment as the title accurately reflects our results and approach. We also respectively push-back on use of the words "The 2019 Methane Budget" in the title as we do not believe that the Global Carbon Project should have exclusive use of the words "Methane Budget". Our use of the words "Each Country" is also accurate; Bayesian integration implies that the posterior is a combination of prior and posterior, consequently, our results for each country are valid as long we have characterized our assumptions and corresponding uncertainties. The authors welcome alternative suggestions to the title as long as the suggested title captures the work presented in this manuscript.

**Comment**: Abstract: The abstract is long and could be shortened to focus more on key findings and less on details of methods. It's not clear if the authors want readers to focus on the findings or inherent limitations that come from uncertainty in the methods and input datasets.

**Response**: We do want the reader to understand the new Bayesian methodology, which in turn allows us to project top-down integrated quantities back to inventories at the desired resolution. We also want the reader to evaluate the results that are enabled by this new approach. Both are needed in the abstract as most scientific results do require either new data or new approaches. However, we have shortened the abstract and hopefully made it more focused.

**Comment**: L204 – Model errors, particularly related to representation of atmospheric transport, are critically important and should at least be mentioned in this section. Are transport errors characterized in this study? If not, what might the effect be? Recent papers on this issue for CO2 inversions (e.g. Schuh et al., 2019, **https://doi.org/10.1029/2018GB006086**) raise substantial questions that are particularly relevant to GEOS-Chem inversions. These should at least be discussed in the introduction and conclusions as major factors that could alter results substantially in the future.

**Response**: Thank you for highlighting this discrepancy! Model errors are discussed throughout the manuscript and a previous version had more discussion on model errors but this discussion was subsumed in Section 2.1. For example, model transport and chemistry is mentioned in Section 1.4 as one of the driving errors (McNorton *et al*. 2020) but then not elaborated upon. We have added (back) a paragraph on model/transport error in Section 1.4 and how it is mitigated with our approach.

**Comment**: L268 – Seasonal variations are assumed to be correct, which could have a large impact on the results. How do seasonal variations in the EDGAR v4.3.2 inventory used here compare with newer version like v5.0 and v6.0?

**Response**: We are not quite sure how to address this question as there seems to be one specific and one implied questions here. To address the specific question, it would be a substantial effort to update the fluxes and emission priors to use the latest and greatest EDGAR values; this will need to wait for a future analysis which we intend to do as part of our NASA Carbon Monitoring System efforts. The implied question appears to be about the role of seasonality in the inventories on our yearly fluxes and by extension the projected emissions. If that is correct, then seasonality does have an impact but its likely minor because the GOSAT data integrate the combined effects of emissions. GOSAT sampling also make it challenging to quantify seasonality in the fluxes although we will attempt this in a future analysis to determine its viability.

**Comment**: L274 – I'm particularly confused about the treatment of wetland sources. They are not included in the state vector, but are presented as distinct in Section 3 results (e.g. Figures 3 and 4).

**Response**: We have modified the language on wetlands and removed the statement "wetland fluxes are not treated as separate elements in the state vector" and moved this discussion to the next paragraph. To add some context, previous GEOS-Chem inversions estimates wetland emissions at very coarse spatial resolution; whereas the fluxes from the integrated emissions are estimated at 2.5 x 2 degrees; however we found that we could not easily project these fluxes back to emissions with the proposed new algorithm (Equations 6 – 8). Consequently, the GEOS-Chem inversion we use now combines all emissions (including wetlands) to fluxes at 2.5 by 2 degrees; we have added language to hopefully clarify these differences and what we use in the revised manuscript.

**Comment**: L390 – Is there a demonstrated improvement in independent data collected in areas that show more influence from nearby sources? It's true that surface data are primarily located outside of source areas, but some aircraft data like ACT-AMERICA do target areas of North America and some TCCON stations would be more relevant in this context. Could the authors provide more detail given the importance of independent validation of such results?

**Response**: Unfortunately, the spatio-temporal resolution of our fluxes are not as easily compared to flight data taken over short time periods.

**Comment**: Section 2.1 – More information is needed about the data used for the inversion. In section 1, the authors point out that inversions with GOSAT and TROPOMI data can provide different results due to biases in the data (attributed largely to TROPOMI). This underscores the importance of the dataset used. More details on what retrieval method is used, version, bias correction, etc. are needed to better understand how these factors affect the results.

**Response**: As noted in the paper (beginning of Section 2.1), the approach and results for the fluxes used in this paper are well documented in previous published papers (Qu *et al*., 2021; Zhang *et al*. 2021; Maasakkers *et al*. 2019, 2021). We do not think it is necessary to re-validate these results in this paper as this paper is focused on using those fluxes with the new algorithm to quantify emissions by sector. To make this more clear we modified the first couple of sentences in Section 2.1 to: "We estimate top-down fluxes based on the approach and results described in Maasakkers *et al*. (2021), Zhang *et al*. (2021) and Qu *et al*. (2021) and the reader is referred to these papers for a more extensive description of the approach and validation of these methane fluxes. To summarize,…"

**Comment**: L426 – Why is 1 degree spatial resolution chosen when the atmospheric model is run at coarser scale (~2 degrees) and many priors are available at finer resolution (0.1-0.5 degree)?

**Response**: As noted in previous comment, the reason for projecting fluxes to emissions at 1x1 degree is to 1) make a more accurate estimate of the country level emissions estimates as a 1x1 degree map better represents a countries borders and hence emissions and 2) projecting emissions to finer scale (e.g. 0.1 degree, which is the resolution of the EDGAR emissions) is computationally challenging and 3) demonstrate that the algorithm works at 1x1 degree and can be used to estimates emissions at any spatial resolution. For example, a future goal is to update the inventories at the reported spatial resolution of 0.1 degree using satellite based fluxes from both the large scale (e.g. TROPOMI, Methane-Sat) combined with high-res estimates from satellites such as Carbon mapper. Our algorithm shown here provides the approach for this combination; we have added language in Section 2.3 discussing these additional details.

**Comment**: Section 2.3 – The use of information at different spatial scales is confusing. The model is run at 2 by 2.5 degrees, prior information is then used to provide additional spatial and sectoral information (down to 1 degree). I am not clear how the 7 or 8 regions described in this section are used in the attribution (and the number is inconsistent within the discussion).

**Response**: See response from previous comment. Also, we fixed wording in text suggesting 7 regions.

**Comment**: L512-513 – The authors point out an important limitation of the utility of such methods – emissions are co-located and cannot be distinguished. How does this affect the intended use of this product in assessing BU inventories?

**Response**: This is an important point that is central to the whole manuscript and which the new algorithm (Equation 6 through 8) explicitly addresses. We have added language in Section 2.3: "Although there can be many emissions within a single grid box, uncertainty can still decrease for each emission type as shown in Equation 8, which shows that these correlations are quantified in the posterior covariance. *Uncertainty reduction of a particular emission therefore depends on the magnitude of the emission and its uncertainty, its correlations with nearby emissions of the same type (next section) and the magnitude and uncertainty of emissions within the same grid box*."

We also add a reminder statement in Section 3.1 to address this point: "*Uncertainty can decrease for emissions even when there is more than one type of emission in a grid box. As shown in Equation 8, this uncertainty reduction depends on the magnitude of the emission and its* *uncertainty, its correlations with nearby emissions of the same type (Section 2.3) and the magnitude and uncertainty of emissions within the same grid box*."

**Comment: Section 3** – How do the authors think that limitations in current satellite data (e.g. albedo biases, lack of data in cloudy regions) affect the results? Are these factors particularly confounding for estimation of certain sectors like rice cultivation and wetlands?

**Response**: Systematic errors are discussed in Section 1.4 but we have added the statement "likely related to poorly characterized surface albedo" when discussing differences between TROPOMI and GOSAT. We also added a sentence when discussing spatial resolution in Section 1.4: "The spatial resolution of the estimate in turn depends on the sampling, pixel size, measurement uncertainty, and lifetime of the gas." to address how sampling affects spatial resolution, which in turn affects smoothing error.

**Comment**: L629-630 – Can the authors provide a point of comparison for a different inversion system that is not based on GEOS-Chem?

**Response**: There are few recent published studies that use GOSAT data and a model that is not GEOS-Chem. However, we did find a paper by Tsurata et al, which is part of the GCP ensemble that we could compare against (see previous comments for Reviewer 1)

**Comment**: L698 – How are low albedo GOSAT data handled? Are they excluded or is there a difference in the retrieval method that should give one more confidence in the GOSAT based inversions. See previous comment on need for more specifics about GOSAT data used in the study.

**Response**: As noted in a previous comment, the fluxes used here have been extensively documented in previous manuscripts. Here we are using the fluxes to project back to emissions. However, we have addressed concerns about sampling and data errors in revised manuscript.

**Comment**: L718 – How robust are results for smaller geographic countries (e.g. Myanmar) given the limited resolution of the model used?

**Response**: Based on the last table, we obtain 2.7 DOFS for Mynamar; given that the prior emissions are all in the livestock sector we expect that the results are robust as the fluxes are well resolved and furthermore there is sufficient spatial resolution to resolve different regions of Myanmar. Of course the caveat to this is that fluxes are modified by random and systematic errors as discussed in Section 2 and systematic errors from model transport and chemistry are hard to quantify in regions with significant convection such as Myanmar. We have added a statement to that effect in the paper.

**Comment**: Section 3.3 – The table including DOF information is useful and the authors take care to note in the text that the inversion really only provides additional information in 58 of the 242 countries listed (and that information is still quite limited, with DOF < 2, in all but 31 countries). However, a casual reader could easily overlook this information and think that the article is claiming to provide satellite-based analysis over small countries/emitters, which I don't think is what the authors intend. Could table 3 be color coded to indicate confidence in the results – for example, green if DOF >2, yellow is DOF between 1 and 2, and red for DOF < 1?

**Response**: As discussed in 3.3, limited information is still information. For example, even a DOFS of 0.5 is still useful as it means that 50% of the estimate is from the measurement. We have added language to that effect in the abstract and re-emphasized it at the start of Section 3.3. We have also added color coding to indicate the different levels of DOFS. DOFS > 1 are green, 0.5 to 1 are yellow, and below 0.5 is red.

**Comment**: L27, 689, 735, 785, 814 – The authors state multiple times that this study is intended as a starting point, which is honest given the complexity of the task. But it is not clear what the next steps beyond a starting point would be. Should these data be used by policy makers involved in the Global Stock Take? What are the priorities for the research community in moving the current state of the art forward? Providing such context would help readers understand how they can best make use of the results presented here.

**Response**: These are good points. I think we can add that regions with limited bottom-up capabilities can use these estimates. Furthermore, we have added discussion that these remote sensing results are at odds with bottom-up estimates of wetland/aquatic emissions and isotope-based fossil emissions and instead emphasize that livestock / rice emissions are dominant drivers of the global methane budget. Essentially, the satellite data show increases in the priors in regions with livestock emissions (e.g. E. Africa, S. Brazil, India) and decreases in regions with fossil emissions (e.g. Russia) and wetlands (e.g. Canada, and Siberia). As these changes occur in regions with lower atmospheric transport errors (e.g. Jiang *et al*. 2013; 2017) it is challenging to attribute these differences to model error; hence renewed attention on the priors should be given here.

Minor/technical comments

L371 – Satellite names should be corrected and should be consistent with L806, L820. (fixed)

Fig. 3-4 – Color scale and size make these hard to see. Consider eliminating some panels to make larger and/or adjusting range of color bars. Response: We will discuss with co-authors and address if possible.

Table 3 – Since a number of different inventories are discussed and used at various points in the study, the authors should clarify which estimates are included here.

Response: These are shown in Table 2 and now restated in the Table 3 caption.