Anonymous Referee #1

The authors use a range of surface and satellite observations of methane to estimate methane emissions from 2010 to 2017. They also use a combined methane-carbon monoxide-formaldehyde inversion that also uses satellite observations of formaldehyde and carbon monoxide. The study describes a range of calculations that sometimes appear to be cobbled together without any particular logical flow, almost as if two groups have written this without any proper integration. Some of the calculations are also presented in a way that makes it difficult to gain any meaningful insights. The paper would greatly benefit from a robust revision, not least to ensure the authors' key messages are easier to understand. Below I outline my substantive and minor comments.

We sincerely thank the reviewer for the constructive comments. We have revised the manuscript accordingly to improve the overall logical flow. Please see responses to individual questions below.

Substantive comments

Line 54: here (or in methods) it would be useful to outline the caveats associated with the CH4-HCHO-CO method. The method assumes correct knowledge of the underlying chemistry, e.g. the fate of the methyl and higher peroxy radicals.

We thank the reviewer for pointing out this point. We have added more discussion regarding model uncertainties in section 2.2.1. "Here, we use a simplified chemistry scheme that assumes methane being oxidized into formaldehyde in a single step. We expect this simplification to have a relatively small impact on the inverse results of methane, given that all pathways of methane oxidation result in formaldehyde as an intermediate product. Besides, HCHO production from non-methane VOC oxidation is simulated upstream with a full-chemistry model, so that the correction on OH from the inversion will not directly feedback to the VOC oxidation. This should not be an issue as we optimize the production of HCHO instead of VOC emissions, but the impact of VOC on OH recycling is not accounted. Future studies using a full chemistry scheme to optimize methane and OH simultaneously would be helpful to diagnose potential impacts of this simplification on the derived methane lifetime."

Line 57: here (or in methods) is an opportunity to tell the readers about any differences in the vertical sensitivity of GOSAT, OMI and MOPITT and how they might impact the combined inversion results. Even if this is addressed in an earlier paper, an acknowledgment would be useful.

We have added information regarding the vertical sensitivities of the three satellite retrievals in the method section 2.1.2. "Satellite retrievals of the three species (CH_4, HCHO, and CO) we use here are generally sensitive to the entire vertical column with some differences toward the lower troposphere. GOSAT X_{CH4} retrievals using backscattered shortwave infrared (SWIR) radiance have approximately uniform sensitivity to methane at all pressure levels (Parker et al., 2015). OMI HCHO retrievals using ultraviolet (UV) radiance are sensitive to the entire column with some decline in the lowest atmospheric layers (Gonzalez Abad et al., 2015). For MOPITT, we use the multispectral total column CO retrieval products that combine near-infrared (NIR) and thermal infrared (TIR) radiances and hence have an enhanced sensitivity to the lower troposphere (Deeter et al., 2014). Such subtle differences in the vertical

sensitivities of the three retrievals as well as their different vertical profiles and lifetimes may influence the ways the observations of the three species inform about OH, which is another source of uncertainty in addition to the model and observation errors."

Line 63: this reader did not find anywhere in the paper any mention of the ability of this combined system to independently estimate CH4, HCHO and CO.

This information is documented in section 2.2.1. "This inversion system has been documented and evaluated by a series of studies focusing on tracers including CH_4 (Pison et al., 2009; Locatelli et al., 2015; Cressot et al., 2014), HCHO (Fortems-Cheiney et al., 2012), CO (Yin et al., 2015; Zheng et al., 2019), and CO_2 (Chevallier et al., 2005, 2010)."

Section 2.1: by using XCH4 from the proxy retrievals the authors are assuming XCO2. Irrespective of what XCO2 they use, this approach will introduce an error in the posterior emission estimates, which should be acknowledged. The resulting XCH4 data might very well agree within X% of TCCON data but this study is making statements about low and high latitude regions where there is barely any coverage from TCCON.

We agree with the reviewer that there are measurement errors irrespective of the retrieval methods. We have expanded the part about observation uncertainty in section 2.1.2 to provide more details regarding retrieval errors. "Here, we use GOSAT X_{CH4} proxy retrievals (OCPR) version 7.2 from the University of Leicester, which has been well documented and evaluated against various observations. The retrieval has a single-observation precision of ~14 ppb (~0.7%) and a regional bias of ~4 ppb compared to TCCON stations (Parker et al., 2015, 2020). This product is also consistent with other GOSAT methane retrievals (Buchwitz et al., 2017). However, we note that there is limited spatial coverage of TCCON stations to fully evaluate GOSAT observations in the high-latitudes and the tropics. "

Line 124: does the optimisation of CH4, CO and HCHO lead to a chemically consistent atmosphere? It would also be useful if the authors reported the methyl chloroform e-folding lifetime as a way of assessing the prior and posterior OH.

The optimization results in very small scaling factors on OH and hence does not change much of the prior OH field (INCA or TransCom) (Fig. 3), both of which have been documented by previous studies such as Patra et al., 2011, Naik et al., 2013, and Zhao et al., 2020.

Line 139: I was baffled by the diversity of uncertainties attributed to chemical production of HCHO production and OH. Please tell the reader where these values come from. Particularly for the low OH uncertainty, given that later in the study (line 145) the authors explain the large differences between OH fields.

The uncertainty assigned to OH (20%) was low compared to that of the HCHO production (200%). Previous studies find relatively small interannual variations in OH (Montzka et al., 2011; Nicely et al., 2018), and a relatively small prior uncertainty was used in many inverse studies (e.g. 10% on the hemispheric mean and 5% on individual years in Zhang et al., 2021). As for the large uncertainty of HCHO production, it accounts for error propagations from VOC emission estimates (based on MEGAN2.0) to the HCHO production simulated by LMDz-INCA. In retrospect, we acknowledge that the uncertainties

assigned to OH might have been too small, which could have resulted in the small posterior scaling factors on OH. Future studies exploring different characteristics of the prior error statistics will be very valuable.

Section 2.2.3: From what this reader understands, the focus of the work is on the 4DVar method. To address the difficulties associated with the ease with which the posterior solution can be characterised using this method, the authors have decided to include additional inversions. This somewhat muddies the water unless the authors can convincingly show both methods produce consistent emission estimates - not just zonal mean totals. For example, is Figure 6 consistent with the 4DVar system?

There are trade-offs between a variational inverse system and an analytical one. A variational system can handle large state vectors and hence our system can account for multiple species simultaneously (i.e. CH4-HCHO-CO), and at the same time, effectively reduce aggregation errors in both space and time (i.e. optimize gridded fluxes or scaling factors on a weekly basis). However, it cannot estimate the averaging kernels that describes the sensitivity of the posterior solution to the true state, as well as the error covariances of the posterior. An analytical system can estimate this useful diagnostic information, but it is limited by computational capacity so that some temporal and regional aggregations of the fluxes are needed to construct the state vector such that the response functions are dependent on the spatial-temporal pattern of the prior fluxes for each state vector.

Our major point here is to estimate the information content of available observations from surface stations or the GOSAT satellites to inform regional methane fluxes. As the available observations, the prior fluxes, and the transport models are the same, these estimates using the simple analytical inverse system (as shown in Figure 6) can provide relevant information to help us interpret the inversion results derived from the variational system that optimizes gridded fluxes of the three species simultaneous.

Line 209: this is a bold and unsubstantiated statement that appears with no prior warning, e.g. discussion in methods. I am sure the authors could come up with competing reasons for small inter-annual variations.

We have revised the text to make a proper transition from results to the discussion. “The resultant small interannual variations in the posterior OH field is in line with some modeling study that shows a high OH recycling probability and hence a weak sensitivity to emission perturbations (Lelieveld et al., 2016).”

Line 216: this is a critical point. Later discussions about OH do not appear to address this point.

We have added more discussion here. “In addition, we note that the uncertainties assigned to the prior OH concentrations might have been too small, and hence result in the small posterior scaling factors on OH. Future studies exploring different characteristics of prior error statistics will be very valuable.”

Line 222: this diversity in results is not addressed very well in the paper and does not bode well for using the alternative set of inversions (section 2.2.3) to help characterise the 4DVar solution. This reader is less concerned about the results using the surface data than the range of results inferred from the satellite data. These satellite inversions are consistent only by virtue of their large uncertainties.

For the inversions that assimilate GOSAT X_{CH4} observations, there are systematic differences in the resultant zonal emission magnitudes due to differences in (1) prior OH filed and (2) GOSAT only or multi-species constraints as shown in Figure 4b. Nevertheless, they all agree on the interannual variations

as shown in Figure 4a, which is what this paper primarily focuses on. The differences between S2 (GOSAT only) and S3 (Multi-species) given the same OH are relatively small in most of the zones except for the 0-30N band, which is mainly due to the scaling on OH in S3 as informed by HCHO and CO observations.

Section 4.1: what I find a bit odd is the authors' use of a four-year period (2010-2013) that includes a La Nina and a subsequent four-year period (2014-2017) that includes a large El Nino. Subtracting these two periods could potentially exaggerate the growth over the eight-year period, particularly over the tropics. Figure S12 shows the temporal changes in global methane emissions (at least I assume it shows the global values). An equivalent figure to accompany Figure 7 would be useful.

As suggested by the reviewer, annual changes in regional emissions are shown in Figure 8, while differences between the 2010-2013 and 2014-2017 period are shown in Figure 7 using the same regional mask. We have made the connections more apparent in the revised text. We agree with the reviewer that such La Nina and El Nino contrast may influence our interpretation about methane growth over the eight years. We have added some discussion regarding this point.

Line 288: do Gatti et al and Liu et al use consistent methods to calculate fire emissions? Otherwise, I am unclear how this statement is necessarily valid.

The two studies used different observations and approaches (Gatti et al. primarily used aircraft campaigns and Liu et al. used satellite retrievals of CO and CO_2). They both show the same temporal changes in fire emissions that are in line with our findings here, which we consider as independent observational support of our results.

Lines 294-298: this statement does not make sense as written. Are the authors suggesting that variations of XCH4 and wetland extent are consistent but land models that incorporate CH4 emissions are let down by imperfect representations of various hydrological processes? And that is why models do not capture XCH4 variations?

We have removed this comment following the reviewer's suggestion.

Line 298: the authors' qualitative statement is noted. They noticed a relationship between one study and another. I am certain they can do better than that.

Detailed analysis of regional drivers on methane emission changes from wetland would be an interesting follow up study. However, it may exceed the scope of the current paper.

Line 301: tropical African emissions of methane originate mainly from the Congo Basin? That is inconsistent with previous studies. The attribution cannot be "supported" by a statement that large peatlands exist in this region. This reviewer understands from Dargie et al that most of the central part of the basin is permanently flooded in which case why would methane emissions be increasing?

We concur with the reviewer and have revised the text according. "Our result of increasing tropical Africa wetland emissions is consistent with a recent regional inversion using GOSAT data at a high spatial resolution of 0.5°×0.625°, which find a positive trend of 1.5–2.1 Tg yr^{-2} in the region over 2010 to 2016, mainly attributed to wetlands in the Sudd in South Sudan (Lunt et al., 2019)."

Line 315: another study has estimated Chinese trends in methane are *likely* due to coal mining but is there any evidence in the multi-tracer inversion that this is true? Are the spatial distributions over China consistent with that conclusion? The authors take more time to interpret the Russian signal using spatial distributions. I encourage the authors to do something similar for tropical Africa and China.

We agree with the reviewer's comments and have removed the original discussion. "The sectoral breakdown of emissions from China suggests a substantial increase in anthropogenic sources from fossil fuel, agriculture and waste, adding up to an overall trend of 1.0±0.2 Tg yr^{-2} between 2010 and 2017 (Fig. 8). As stated above, this attribution relies on the relative contribution of different sectors from the prior information and does not account for structural changes in time. A recent inverse study focusing on Asian emissions from 2010 to 2015 derived nearly the same magnitude of emission trend for China (Miller et al., 2019), a continued increase is confirmed here beyond 2015 till the end of the record in 2017. In contrast, a global inversion that separate the mean anthropogenic emissions and trends in the state vector find a smaller trend in anthropogenic emissions over China (0.39±0.27 Tg yr^{-2}) for the period 2010-2018, and a trend of 0.72±0.39 Tg yr^{-2} focusing on the period 2010-2016 (Zhang et al., 2021)."

Minor comments

Line 41-42: the statements after the first dash makes little sense to this reader.

We have deleted it following the reviewer's suggestion.

Line 48: there have been a few studies to investigate the recent acceleration. I urge the authors to use primary references rather than Nisbet et al 2019 reference, which glosses over some of the underlying issues.

We thank the reviewer for this suggestion and have added more references including McNorton et al., 2018, Turner et al., 2019, Zhang et al., 2021.

Figure caption 1: remove parentheses around Thoning et al.

Corrected.

Line 68: 'We' should be 'we'.

Corrected.

Formaldehyde is referred to as CH2O and HCHO. Please be consistent.

Thanks for pointing this out. All changed to HCHO.

Line 162: posterior or a posteriori. Please be consistent.

All changed to posterior for consistency.

Equations 1 to 3: the convention is to use lower case bold for vectors and upper case bold for matrices.

Corrected.

Line 171: reference is a bit mangled.

Corrected.

Line 176: ...generally capture well. . . This reader fails to understand the meaning of this statement.

Revised as “In general, the global average methane growth rate is well captured by the posterior model states”

Line 183: anchor points rather than anchoring points?

Corrected.

Line 185: this statement assumes that variations in the column overhead can be related to changes in the underlying surface emissions.

The statement here compares the measurement precision and spatial coverage (including vertical sampling) between the two types of observations without discussion or assumptions about underlying surface emissions.

Line 290: typo. Anthropogenic.

Corrected.

Line 294: increase exponentially with temperature?

Yes. Changed accordingly.