The MS[ACP-2022-749] by Varon et al., used satellite-based xCH4 observations and inversion modeling to constrain CH4 emissions from oil and gas production hotspot in Permian Basin, the largest oil production basin in the United States at a weekly scale. It's very important to quantify CH4 leakage in these fossil fuel production areas. In general, this MS is easy to follow and well written, considering a few similar studies have been conducted in the same region, some clarification is needed to be highlighted, especially for the most important improvement in this study as displayed below. And it can be accepted after addressing the following comments.

Main comments:

In general, there are a few published papers in the same region that used similar approach (i.e. Zhang et al., 2020), although this submission is conducted at a weekly scale instead of a monthly or annual scale, the authors still need to address the improvement of their study from previous ones, i.e. approach in the inversion framework? or found the weekly relationship between CH4 emissions and other activity indexes? or used more robust observations? or prior emissions and background?

Line 19-21 "The mean oil and gas emission from the region (± standard deviation of weekly estimates) was 3.7 ± 0.9 Tg higher than previous TROPOMI inversion estimates that may have used too-low prior emissions or biased background assumptions". It seems the inversion results are sensitive to prior emissions, have you tested or quantified this potential bias of using different prior emissions to your results?

Line 88-89, "19346 ± 13073 observations per week over our full inversion domain (96°–110°W, 25°–38°N), including 3062 ± 2314", The standard deviation of available data numbers in each week is so high and is comparable with averages, which indicates there are not enough data in some weeks, please address what the potential bias for emission inversion in these weeks with lower available observation numbers. And can you display the time series of available CH4 observation numbers in each week?

For some rainy or cloudy weeks, the available observation data can be sparse, leading to a large missing data gap in the study domain, and how will this situation affect inversion results for this large region? The reason to mention this comment is that your following analysis of the relationship between CH4 emissions and activity indexes ignored the influence of available data.

Line 92-94, "We use dynamic 3-hour boundary conditions from a global 4°×5° simulation corrected with spatially and temporally smoothed TROPOMI data as described by Shen et al. (2021). A one-month spin-up simulation starting from these boundary conditions is used for initialization". As we know that CH4 background uncertainty (bias) will be carried on to calculated CH4 enhancement, which is directly related to posterior CH4 emissions, what the bias of CH4 background in this study and potential uncertainty in deriving CH4 emissions?

Line 110, "It attributes 94% of Permian emissions to oil and gas activity, and we assume the same fraction for our posterior emission estimates.", Can you clarify what is the potential uncertainty of using the constant fraction of 94% to oil and gas activity in inversion results? As I

know most inversion studies have the ability to constrain posterior emissions from different categories.

Section 2.2, As displayed in the reference list, there are some other inversion studies in the same domain, (i.e. Zhang et al., 2020), it's better to illustrate what the main improvement of your study when compared with these previous studies, because it's very hard for audiences to remember and distinguish the method difference between all related studies.

Line 125, "mitigate boundary-condition errors (Shen et al., 2021; Varon et al., 2022).", have you assessed the improvement of CH4 background with observations?

Line 136-137, "The error covariance matrices and are assumed diagonal with uniform error standard deviations of 50% and 15 ppb, respectively", As I understand, the 50% uncertainty for prior inventory may represent regional averages not all grid cells in study domain, which can be much larger than 50%, the same as 15 ppb for observations and GEOS-Chem simulations, so whether the inversion results are sensitive to different values of 50% and 15 ppb, if you assign a slightly larger extent (i.e. 80%, and 20 ppb), how much will the results change?

In figure 3, the weekly emission changes can vary by 100%, indicating the potential bias or uncertainty of CH4 emissions at weekly scale can be much larger than 50%.

Line 190, It seems the use of proportion 94% will largely influence your results of CH4 emissions from oil and gas. I am curious why the inversion model cannot constrain CH4 emissions from each category?

Data displayed in Figure 5 for model simulated xCH4 and observation.

Overall, why tower based CH4 concentrations seem higher than simulations with both prior and posterior emissions (scatter plot is below 1:1 line)? Whether it indicates the posterior CH4 emissions are still underestimated? How about plotting time series of concentration?

Whether one of the reasons for the large difference between model simulation and tower observations is the vertical gradient in the lowest GEOS-Chem model? And what is the height of the lowest model level? or aggregation error for spatial resolution between the point scale and regional scale(25km)?

Line 256-257, "The mean satellite-inferred emission (0.72 Tg is 20% lower than the mean tower (0.88 Tg and Scientific Aviation (0.89 Tg estimates during the period of overlap", From the above concentration comparisons, I also agree that the satellite-inferred emissions are obviously underestimated.

Line 398-410, It's better to display the comparisons between atmospheric inversions and multiple linear regression with figures instead of only using tables and numbers.

Line 438, "assuming 80% methane content for Permian", whether the assumption of using 80% is reasonable, and what the general extent of this value in the study region?

Technical comments

Line 398, "summarizes the models and results" I just guess the first author forget to delete this sentence of comment from other co-authors.