Response to referee comments on "Satellite quantification of oil and natural gas methane emissions in the US and Canada including contributions from individual basins"

We thank the two referees for their careful reading of the manuscript and the valuable comments. This document is organized as follows: the Referee's comments are in *italic*, our responses are in plain text, and all the revisions in the manuscript are shown in blue. **Blue text** here denotes text written in direct response to the Referee's comments. The line numbers in this document refer to the updated **WORD** manuscript **with tracked changes**.

Referee 2

This work by Shen et al. uses TROPOMI XCH4 retrievals to quantify oil/gas emissions at high GEOS-Chem resolution of 0.25 by 0.3125 degree. It follows the optimal estimation framework with GEOS-Chem as forward model developed in the TROPOMI-Permian study by Zhang et al. (2018), GOSAT/in-situ-global study by Lu et al. (2021), GOSAT/in-situ- North America study by Lu et al. (2022), and TROPOMI-Mexico study by Shen et al. (2021). This work also presents significantly improved spatial resolution of emission estimates and detailed basin-level emission comparison with previous quantification by airborne measurements. It is recommended for publication in ACP after addressing the following issues.

General comments:

Some languages are still reminiscence from previous version of the manuscript. For example, "see Methods for more details", whereas there is no "Methods" section. The first three paragraphs of section 3 is largely redundant.

Response. Thanks for pointing this out. We have removed the 2^{nd} and 3^{rd} paragraphs in our old manuscript and combine related contents with the Method part.

It is suggested to clarify and maybe expand section 2.6 on the ensemble uncertainty analysis, which seems to be an advance from previous GC analytical inversion studies from the group. How does this "posterior uncertainty" compare with the classic posterior uncertainty calculated from equation 4? Where does the number 2400 come from at line 178? Posterior emission at each grid cell drawn from Gaussian PDF 100 times and then multiplied by 24?

Response. Thanks for pointing this out. We have added these contents to clarify our ensemble uncertainty analysis.

Line 252. The posterior covariance matrix $\hat{S}$ describes the error within the choice of each set of inversion parameters, and the ensemble allows us to explore the uncertainty arising from the selection of these inversion parameters. We use the Monte Carlo method to estimate the posterior uncertainty from the ensemble. For each of the 24 members, we generate 100 samples from the posterior distribution, which yields 2,400 samples in total for each grid cell. We report error statistics on the inversion results as two standard deviations (2σ), corresponding to the 95% confidence level.

Section 5 gives a novel way of quantifying TROPOMI's capacity to constrain basin-scale emissions. However, the posterior uncertainty threshold of 30% seems arbitrary and presumably closely related to the prior uncertainty assumed (50% here). Equation 8 may be misleading as it shows that the posterior relative uncertainty is driven by basin- total emissions and satellite coverage only. Consider adding prior uncertainty as a predictor, or using the relative reduction from prior error to posterior error.

Response. We tried to add prior uncertainty as a predictor, and we don't find it can significantly improve the performance of our linear regression model (Equation 8). This is because posterior uncertainty is less dependent on prior information when the satellite density is high. This also means our posterior uncertainty is not closely related to the prior uncertainty.

The referee has made a good point about the relative reduction of errors, and we add a new supplementary figure.

Line 420. We also estimate the relative error reduction from prior to posterior estimates, and our results show that the uncertainty decreases by an average of 40% (0-80%) across the 19 O/G basins (Fig. S15).

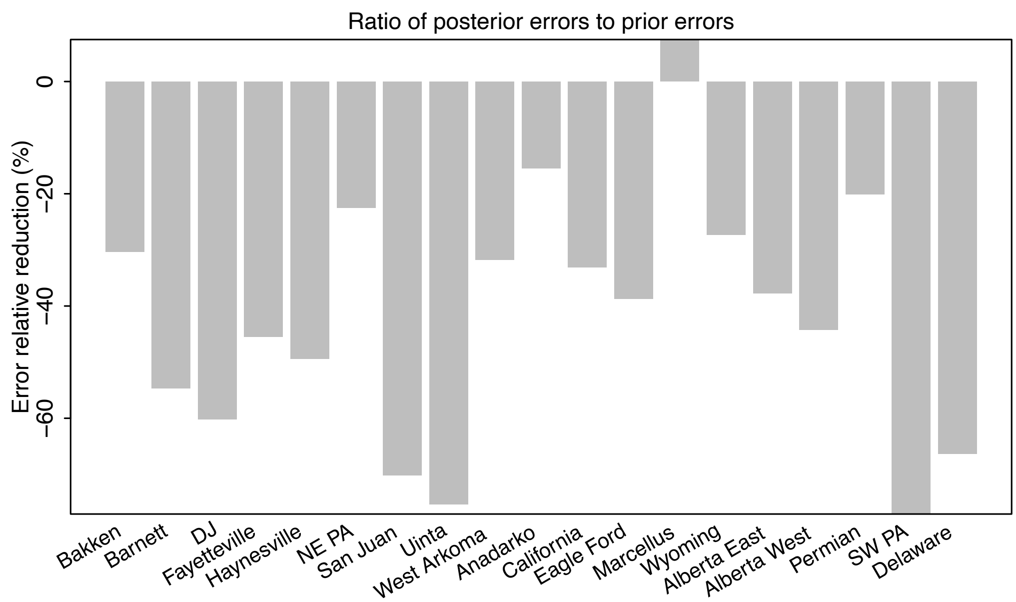

Figure S15. Ratio of posterior errors to prior errors for the 19 basins. The basin-scale prior errors are calculated using a similar approach as described in Section 2.6.

Specific comments:

Lines 112-115: is this the "specification of boundary conditions" mentioned in lines 91-92? It is not very clear what the "vertical fields" are, and how GC CH4 fields are corrected exactly. Was GC CH4 scaled every day, so that the mean of the boundary grids of the North America domain matches the mean of TROPOMI pixels within a buffer zone?

Response. Now we say this in text.

Line 127. Following Shen et al. (2021), we correct the local boundary conditions on a daily basis by scaling to the ratio of TROPOMI and GEOS-Chem columns averaged over the neighboring ±1,000 km and ± 15 days.

Line 122: it implies that the time variation of CH4 emission at each grid cell is not considered. Is constant emission assumed throughout the study period? Please clarify.

Response. Now we make it clear in text.

Line 115. Emissions from the oil and gas are assumed to be constant throughout each year with not seasonality.

Line 135: this equation implies that the model is linear so the jacobian (K) can be calculated only once. A linear forward model is essential for this framework given the cost of generating K. It is suggested to add a reference or calculations to justify that GC CH4 simulation is linear.

Response. Thanks. Now we say this in text.

Line 131. This involves constructing a Jacobian matrix K that describes the sensitivity of model XCH4 to each emission state vector element. The construction is done by conducting sensitivity simulations in GEOS-Chem for the inversion period perturbing individual state vector elements in turn, and this is readily done on a high-performance cluster as a massively parallel problem.

Line 168. Methane sinks from oxidation and uptake by soils are included in GEOS-Chem but we do not optimize them here since they are irrelevant in nested model simulations where the loss of methane is by ventilation outside the domain (Varon et al., 2022).

Line 177. The relationship between emissions and methane concentration (XCH4) is strictly linear since the sinks are not optimized (Varon et al., 2022).

Line 136: it is suggested to provide more information on how K and S_o are constructed and how the computational challenges are solved, given the very large number of observations (7e6).

Response. Here we make it clear in text.

Line 181. Both S_A and S_O are taken as diagonal**, and we use γ to avoid overfitting**.

Line 206. The regularization term γ is intended to account for unresolved observational error covariances in the inversion and thus to avoid overfit to observations.

Line 154: suggest to comment on the rationale of choosing this over the L-curve approach in some previous works.

Line 207. Following Lu et al. (2021), we choose γ such that $(\boldsymbol{\hat{x}} - \boldsymbol{x_A})^{\boldsymbol{T}} \boldsymbol{S_A}^{\boldsymbol{-1}} (\boldsymbol{\hat{x}} - \boldsymbol{x_A}) \approx n$ where n is the number of state vector elements, as would be expected from a chi-square distribution with n degrees of freedom. This yields γ in the range 0.1-0.4 with a best estimate of 0.2 (Fig. S7). We previously found a similar range of γ using the L-curve method in a previous regional inversion of TROPOMI data for eastern Mexico (Shen et al., 2021).

Line 169: it is "sigma_i,nation" that refers to the error standard deviations.

Response. Now we say.

Line 223. $\boldsymbol{\sigma_{i,nation}}$ refers to the error standard deviations on the national totals obtained from Maasakkers et al. (2016) and Bloom et al. (2017).

Line 182: suggest emphasize that this is "topography-corrected" XCH4 in the text.

Response. Now we say.

Line 261. The data shown in Fig. 1a are corrected for topography following Kort et al. (2014), but this correction is not used in the inverse analysis because the GEOS-Chem forward model accounts for topography.

Line 220: is 20% the sum of the first 9 largest O/G basins' uncertainty, or the mean of them?

Response. Now we say:

Line 277. The average posterior uncertainty is 20% (2σ) for the first 9 largest O/G basins

Line 269 and Figure 4: is R2 from the fitted ordinary least squares line between this work and inventory/field work? Different fitting lines may give different R2. Pearson correlation coefficient might be more proper here.

Response. Thanks for pointing this out. Both methods give very similar R^2.

Line 279: update the "Methods" section.

Response. Done.

Line 280-281: please clarify/confirm what is "error variances weighted by the corresponding error covariances".

Response. Now we say.

Line 409. The AK sensitivities (diagonal terms of the AK matrix) measure the ability of the inversion to quantify the true emissions independently from the prior estimate (1 = fully, 0 = not at all).