We thank the reviewers for their additional suggestions for the manuscript. Please see below our responses to the comments. We have listed out the reviewer comments in black and the replies in blue.

Reviewer #1

I think the manuscript is improved, particularly in supporting and qualifying the ability of the posterior solution to distinguish between emissions from different sectors. I still think that this manuscript is suitable for ACP, but I remain concerned that the authors have not shown that the Bayesian inversion method improves the simulation of satellite observations that are not directly used in the inversion itself. Both reviewers raised this issue, and I do not feel that it was sufficiently addressed. The coal point-source comparison in Figure 11 is useful, but a large focus of the manuscript emphasizes improvements to area emissions (notably, livestock), which should be supported by some form of validation against withheld TROPOMI or GOSAT observations. I recommend major revisions to provide some form of independent validation prior to publication.

Figure 7 alone is not enough to convince me that the posterior emissions are an improvement over the prior. Because the same observations are used in the inversion and in the evaluation shown in Figure 7, the posterior is expected to reduce the mismatch for those observations. This does not show whether the posterior solution is closer to true emissions or if it reflects overfitting or compensating errors, particularly given that the strongest posterior sensitivities occur in regions with dense observations.

I do not find the authors' argument that withholding a fraction of observations as a validation dataset "would not help" because the Bayesian inference solution is regularized persuasive as currently stated. Regularization reduces risk of overfitting, but it does not make evaluation with withheld observations uninformative. In fact, independent validation showing that the Bayesian inference method improves simulation of withheld observations should support the authors' claims that the solution is properly regularized. Could the authors explain this argument more clearly?

Because the Jacobian matrices, K, have already been computed and the inversion is analytical, some form of independent validation appears possible without additional GEOS-Chem simulations. One option would be to withhold observations in time or by observing system (TROPOMI or GOSAT), re-solve the inversion with the remaining data, and evaluate the posterior against the withheld observations using the precomputed Jacobian matrices. Other independent tests of posterior predictive skill would also address this concern.

Thank you again for taking the time to review and help improve this work. We appreciate the suggestion of removing GOSAT observations for independent validation and have now included this in the manuscript. We find the contribution of the GOSAT observations to the inversion is negligible due to GOSAT's sparsity relative to TROPOMI (Figure 5) and the inversion results do not change.

The inversion yields improved agreement with GOSAT observations, as is now shown in Figure 8 and described in the text in lines 401-407:

"Figure 8 shows an independent evaluation against GOSAT observations for 2023. Prior and posterior simulations were sampled using the GOSAT observation operator for direct comparison with the observations. The posterior simulation shows a reduction in bias from 17.1 to 15.0 ppb; most of the remaining bias is due to the mean difference between TROPOMI and GOSAT observations (9.8 ppb; Figure 8, right panel). The RMSE decreases from 23.0 to 21.1 ppb, while the correlation coefficient increases from 0.60 to 0.64. The reduced major axis (RMA) slope also decreases from 1.37 to 1.24. The posterior emissions are therefore better able to simulate the GOSAT observations than the prior emissions."

Technical comment: $\mathbf{W}*$ in eq. 5 should be defined in the main text. I assume it represents a pseudo-inverse, but it could be confused with a conjugate transpose.

Thank you for noticing this. We now define $\mathbf{W}^*$ on line 356.

Reviewer #2

We thank the authors for their detailed responses to the reviewer comments. The revisions are clear, comprehensive, and address most of the concerns raised. However, I remain unconvinced by the justification for not using independent observations for validation.

Thank you again for taking the time to review and help improve our manuscript.

The authors argue that validation using the same observations employed in the inversion is acceptable because the Bayesian inference framework is regularized. However, the posterior emissions are still optimized to reduce the mismatch with those observations, so demonstrating reduced bias against the same dataset does not constitute an accurate evaluation. I understand that no additional in situ datasets are available over Colombia for 2023, but it is common practice in atmospheric inversions to withhold a subset of observations for validation. Given the lack of independent datasets, I recommend that the authors either (1) reserve GOSAT observations for evaluation only, or (2) randomly withhold 10–20% of the satellite observations across the analysis period and use them for independent validation. This would provide a more robust assessment of the posterior emissions and strengthen confidence in the inferred fluxes.

Thank you for this suggestion. We have followed your recommendation and removed GOSAT observations from the inversion so they can be used for evaluation. We find their contribution to the inversion is negligible due to GOSAT's sparsity relative to TROPOMI (Figure 5) and the inversion results do not change.

The inversion yields improved agreement with GOSAT observations, as is now shown in Figure 8 and described in the text in lines 401-407:

"Figure 8 shows an independent evaluation against GOSAT observations for 2023. Prior and posterior simulations were sampled using the GOSAT observation operator for direct comparison with the observations. The posterior simulation shows a reduction in bias from 17.1 to 15.0 ppb; most of the remaining bias is due to the mean difference between TROPOMI and GOSAT observations (9.8 ppb; Figure 8, right panel). The RMSE decreases from 23.0 to 21.1 ppb, while the correlation coefficient increases from 0.60 to 0.64. The reduced major axis (RMA) slope also decreases from 1.37 to 1.24. The posterior emissions are therefore better able to simulate the GOSAT observations than the prior emissions."