Reviewer:
This manuscript investigates the 2019 Brazilian fire season under non-extreme climate conditions using a two-step CTE-LW/SW inversion framework constrained by TROPOMI and MOPITT XCO observations. The topic is scientifically relevant and timely. The inversion design is carefully structured, and several sensitivity experiments are conducted to assess robustness. The study provides valuable insights into regional fire CO emissions and bottom-up inventory discrepancies. The manuscript is generally well organized and supported by data. However, several methodological aspects require further clarification and quantitative support before publication.

We thank the reviewer for their kind words and for taking their time to evaluate the manuscript. We address the suggested clarifications and corrections below (in blue text).

Major Comments
1. Identifiability of the Two-Step Inversion Framework
The manuscript applies LW to optimize background CO and SW to optimize fire emissions. However: How does the LW step avoid absorbing part of the fire signal?
Thank you for this important question. As with any atmospheric CO inversion, we cannot absolutely guarantee that fire signals are not partially absorbed by other source terms; however, our framework is explicitly designed to minimise this aliasing through:

- *pre-selecting observations for assimilation:* The LW/SW is designed to decompose the inverse problem in two distinct spatio-temporal scales. In the long-window, we assimilate only background flask observations, which are rarely, if ever, near fresh fire plumes, and therefore not dominated by (Brazilian) fire-influenced CO signals. We acknowledge that background stations like ASC may pick up some fire signals every now and then from Africa; however, from the South American fire emission perspective, these would be part of the background signal.
- *LW state vector design:* The coarse and smooth spatial and temporal structure of the LW state poorly represents the sharp, episodic character of fire emission enhancements (and errors therein). Consequently, even if a residual fire signal were present in the background observations, the LW state vector has limited capacity to absorb it, and any aliasing would be confined to broad continental-to-latitudinal patterns averaged over seasonal timescales.

As a consequence of these choices, we largely rely on the prior fire emissions to capture inter-annual variability (IAV) of the fire signal. Even if we would alias some of the large/slow scale (continental, multi-month) variability in the fires , that is not our main concern, because our primary focus remains resolving the high-frequency temporal and fine-scale spatial variations of fire emissions for tropical South America's 2019 dry season in the SW step. Furthermore, for the case study presented we find that the background CO field in the domain is not strongly biased, so we can be confident that aliasing for this period/region is weak. We recognise that this assumption needs careful re-examination as we extend the framework (e.g. simultaneously cover multiple regions or apply it to different years) and consider this an important point of attention as we further develop/apply the framework.

Does LW optimization alter OH or NMVOC production patterns regionally?
The long-window inversion is given freedom to adjust the latitudinal structure of secondary CO production fields (including contributions from NMVOC oxidation), any scales more local than these are not explicitly scaled. Because OH concentrations directly govern the secondary CO production, any regional bias in OH will manifest as a bias in secondary CO production and will, to the extent that the observations inform on this, be aliased into the LW corrections of those fields. In practice, this means that the LW step can implicitly compensate for north-south gradients in secondary CO production and OH biases, even though OH itself is not an explicit state variable.

Could the SW step compensate for residual background bias?
We consider this risk to be limited, because compensation for residual background bias in the SW step would require that: (1) the bias is detectable in the set of super-observations assimilated in the short window; (2) those observations then must be within no more than ∼6 days downwind of a fire source region; and (3) the residual background bias co-varies spatio-temporally with the fire emission signal. When there is no such covariance in background/fire residual, the SW state vector has no pathway through which to absorb this.

The manuscript mentions a remaining ∼20 ppb positive bias in the free troposphere but argues it does not dominate SW results. This claim requires quantitative sensitivity tests. Please clarify information flow and provide supporting diagnostics.

Thank you for bringing up this point. Our statement was based on the finding that the SW ensemble spread in free-tropospheric observations was negligible. Upon reflection, we recognise that we did not address the possibility of a free-tropospheric bias projecting onto the fire emission estimates through our X_{CO} residuals ($y - H(x)$) in the SW assimilation step.

We conducted a sensitivity experiment (explained below) to quantify the potential impact of a free-tropospheric bias which demonstrates that this bias would have minimal impact on our reported fire emission estimates. Under a conservative scenario, we estimate the resulting change in fire emissions to be ∼10% or ∼4.7 TgCO. This falls well within the ±15 TgCO structural uncertainty we report and would not alter our conclusions. We have clarified this information flow and its implication for the conclusions in Appendix A3.

Description sensitivity experiment:
We applied a uniform perturbation of -20 ppb to the TM5 model profiles between ∼600 and ∼150 hPa. Although the Manaus profiles reach until ∼5km, we assumed in this experiment that the bias persists to the upper troposphere as conservative upper bound. The perturbation was applied to all assimilated X_{CO} within 10° of the Manaus profiles (2.6°S, 60.2°W), in total > 40k samples. Recalculating the X_{CO} (using the column averaging kernels) for TROPOMI, resulted in a change of about 10% in X_{CO} ($\frac{(X_{CO}^{FT-20} - X_{CO}^{base})}{X_{CO}^{base}} \times 100$) as demonstrated in Figure 1 by their kernel density estimates that we report per $\sqrt{R}$ quantile. We found comparable numbers for MOPITT.

Assuming linearity of the observation operator $H(x)$ with respect to the state vector x, this 10% ΔX_{CO} translates to a 10% change in the posterior fire emission estimates. Given that the observed free-tropospheric bias was predominantly positive, this suggests that our posterior fire emissions may be underestimated by ∼10% if such bias would be present over the whole domain, up to the tropopause, for the entire period.

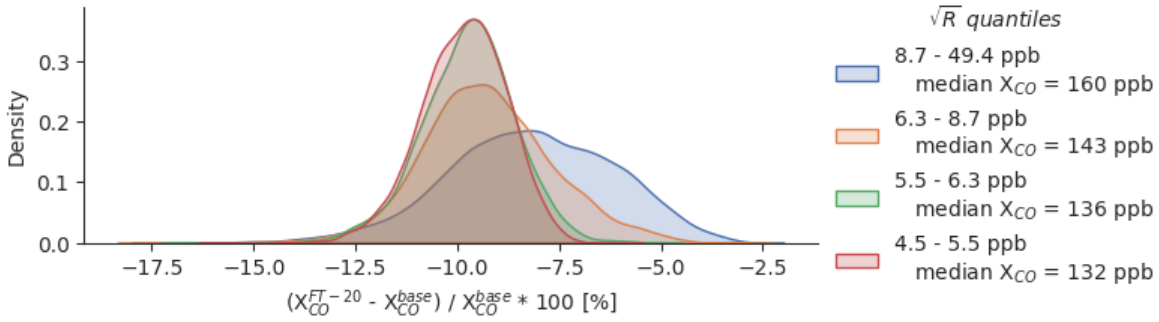

Figure 1: Kernel density estimates per $\sqrt{R}$ quantile with the change in X_{CO} (%) when subtracting 20 ppb for assimilated TROPOMI samples in the GFED5.1_TROPOMI inversion. The legend show the range of the $\sqrt{R}$ quantile and the median X_{CO} for that quantile (before subtracting the 20 ppb).

2. Justification of the X_{CO} ≥125 ppb Filtering Threshold
The chosen threshold is used to isolate fire plumes, but: How sensitive are total emissions to this threshold? Could low-intensity or diffuse fires be excluded? Please provide emission estimates under different thresholds (e.g., 100, 125, 150 ppb). Also include statistics of excluded observations (<125 ppb) to demonstrate that the filtering does not bias total fire emissions.
Thank you for the suggestion, we conducted an additional set of (GFED5.1 + TROPOMI) inversions with the 100 ppb and 150 ppb thresholds and updated Section 2.5.4 in the manuscript accordingly. Figure 2 demonstrates that our total emissions are not very sensitive to this threshold, even at the biome scale. The (max. - min.) range in posterior total Brazil fire emissions for the simulation period between the TROPOMI inversions with different thresholds (1.1 TgCO, or 2-3%, of the Brazil season total) is smaller when starting from different priors (1.7 TgCO). Between the different threshold runs, the Caatinga emissions are most affected in the 150 ppb threshold. This is logical as these columns usually are relatively clean and therefore get filtered out first, which limits the constraint on the Caatinga's fire emissions.

Please find the mean absolute errors of the excluded observations in Table A2 of the manuscript's Appendix A2 for 125 ppb threshold inversion and in Table 1 for the new runs. These indicate that the statistics for the match to non-assimilated obser-

vations does not degrade considerably after assimilation, consistent with minor changes in emissions.

Regarding low-intensity or diffuse fires, we agree these are likely under-constrained, as we discuss in Sections 4.2 and 4.3 of the manuscript. However, we hypothesise that the observation threshold is not the main limiting factor on constraining diffuse fire emissions, given that the inversion currently is limited in its capacity to scale these fires. This is because we apply multiplicative scaling factors with prior magnitude-dependent covariances, limiting the freedom to change emissions: zero/small prior emissions can only result in zero/small emission changes (original manuscript lines 478-484).

We updated Section 2.5.4 to clarify the impact of the threshold on the emissions.

Table 1: Mean Absolute Error over samples that were not selected for assimilation.

		MAE	
Experiment	n	Prior	Posterior
100 ppb	319730	5	6
125 ppb	497780	6	7
150 ppb	556250	7	7

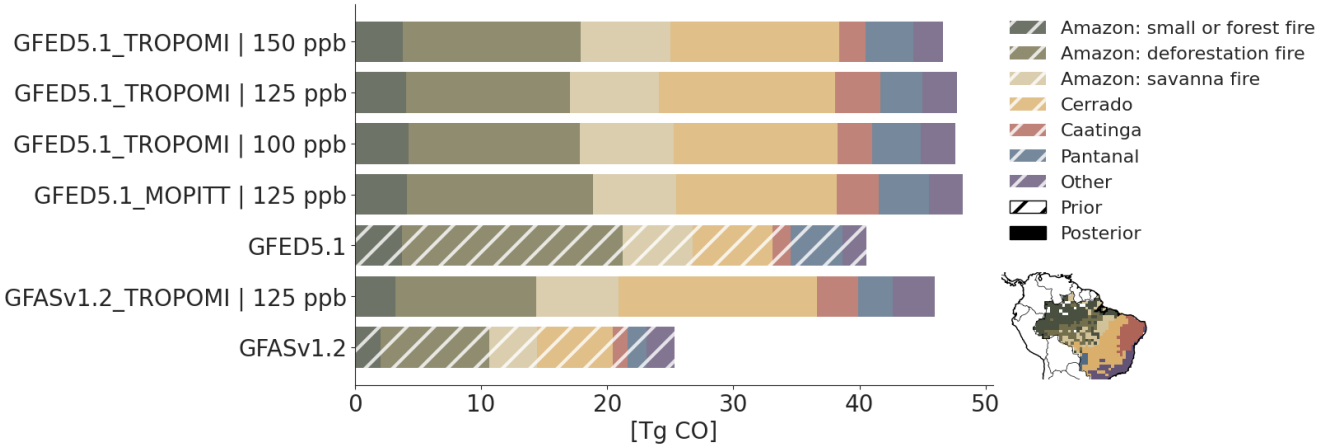

Figure 2: Prior (hatched) and posterior (solid) CO fire emissions summed over the simulation period (Jul 22 – Nov 30 2019 for Brazil, stacked by biome (indicated in the inset map). The figure is an extension of manuscript Fig. 4 with two sensitivity inversions using data selection thresholds of 100 and 150 ppb.

3. Observation Error Inflation Strategy

TROPOMI errors are inflated (transport error doubled + 4 ppb added) to achieve $\chi^2 \approx 1$. While common in data assimilation, further clarification is needed:Show χ^2 distributions before and after inflation (not only the mean). Explain how the "$\times 2 + 4$ ppb" parameters were chosen. Provide sensitivity tests to alternative inflation factors. Importantly, TROPOMI native uncertainty ($\sim$3 ppb) is much smaller than MOPITT ($\sim$7 ppb). Inflation changes their relative weighting. Please discuss whether the reported convergence between TROPOMI and MOPITT inversions is partly influenced by this weighting adjustment.

We understand the reviewer's reasoning regarding the potential impact of observation weighting. Isolating the impact of the observation error covariance choices (R) on flux convergence would require an extensive series of new inversions for both TROPOMI and MOPITT wherein we set different R choices, even then it would remain challenging to isolate the impact of R from intrinsic differences in the observations sets themselves on the flux convergence. Nonetheless, we agree that it could be a valuable exercise for future studies. For now, we clarify below why inflating R for the TROPOMI inversion was a necessary and logical step to facilitate a fair comparison with the MOPITT inversion in the context of this study.

An inversion must be balanced in terms of the actual model performance relative to observations ($y - H(x)$) and the anticipated errors ($HPH^T + R$). A poor balance could lead to over/under-fitting, make the inversion extra sensitive to noise in the obser-

vations, and possibly even lead to unstable solutions. Through either P or R the balance between prior weight vs. observation weight can be influenced.

We would like to underline that the R choice is not entirely free. Importantly, we need to consider here that R, besides instrument precision, must account also for forward model transport errors and the capacity of these models to represent the observations. As noted in Section 2.5.4 of the manuscript, TROPOMI and MOPITT "see" the atmosphere differently. TROPOMI's greater sensitivity to the lower troposphere exposes it to different transport modelling challenges that MOPITT largely avoids. Specifically, boundary-layer mixing dynamics, plume dispersion closer to the source, and turbulence-induced variability are features that our 1°×1° transport model (TM5) may not fully resolve or that are simply more difficult to capture. Assigning TROPOMI superobservation errors only, would underestimate the true model-data mismatch because it does not account for the aforementioned extra challenges of representing these observations.

Elaborating on Section 2.5.4, our inflation strategy consisted of:

1. *A uniform baseline structural error (+4 ppb):* to account for the bulk of representation errors associated with boundary-layer dynamics across the entire domain.

2. *A plume dispersion error (×2 column gradient):* this term scales dynamically, following the simulated column gradients, to represent plume dispersion uncertainty.

To address the reviewer's comment, we present sensitivity tests with various additive structural errors and a larger multiplicative dispersion error for R and illustrate how these could impact the distribution of χ^2_{inno} that represents the actual vs. observation error balance (Fig. 3). Note that we take the prior models for "MOPITT" and "With inflation" from the GFED5.1_MOPITT and GFED5.1_TROPOMI inversions respectively. The additional sensitivity tests are calculated from the GFED5.1_TROPOMI prior model and not from new inversions. The Figure demonstrates that R inflation ("With inflation"; median 1.15 ppb) yields a more balanced χ^2_{inno} than without ("Without inflation"; median 3.12 ppb). The remaining tests show the χ^2_{inno} distributions for the individual components and that the combination of both is more effective at bringing the χ^2_{inno} of the bulk towards one.

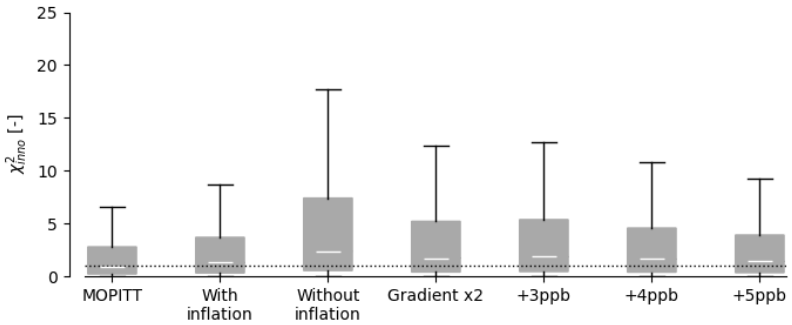

Figure 3: χ^2_{inno} distributions comparison. The χ^2_{inno} values were calculated using the prior model from the GFED5.1_MOPITT inversion (left-most box) and the GFED5.1_TROPOMI inversion (remaining boxes).

4. CO-to-C Conversion Assumption

CO emissions are converted to carbon emissions using a fixed CO/CO_2 ratio. However, emission factors vary by fuel type and biome. Please provide sensitivity analysis (e.g., ±20% in CO/CO_2 ratio). Quantify the resulting uncertainty in carbon estimates. Clarify the limitations of using CO-only inversion for total carbon quantification.

Thank you for this helpful suggestion. We re-calculated the total carbon emission (here: C from CO and CO_2) using 20% lower and higher $\frac{CO_2}{CO}$ and updated manuscript by including the following (lines 492-498):

To include the uncertainty arising from the assumed prior emission factor ratios, we performed a sensitivity analysis perturbing the assumed $\frac{CO_2}{CO}$ ratios by ±20% (Fig. A7). This widens the estimates by ∼55 TgC in both directions and results in a range of 216–334 TgC. The downside of CO-only inversions for total carbon emission quantification is that we must rely on the accuracy of the emission factors that vary by fuel type, fuel moisture content, and meteorological conditions.

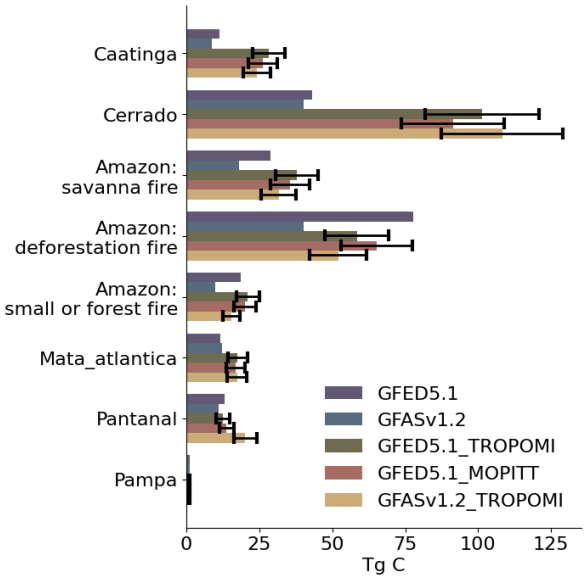

Figure 4: Carbon emissions (C from CO and CO_2) per biome. Posterior carbon emissions were inferred from CO using the prior inventory's $\frac{CO_2}{CO}$ emission factor ratios at the location/time of the emission. The error bars represent the min/max range for a hypothetical 20% error margin on the CO_2/CO emission factor ratios. This Figure was added to the manuscript in Appendix A7.)

5. Underestimation over Cerrado/Caatinga

The manuscript reports posterior emissions nearly twice the inventory values in this region. While fuel pool explanations are discussed, the attribution remains qualitative.

Please consider: Showing spatial posterior/prior ratios. Providing innovation statistics by biome. Clarifying whether transport redistribution contributes to the regional signal.

Thank you for these helpful suggestions. We agree that the attribution of the large emission increments over the Cerrado and Caatinga remains largely qualitative. While providing a full quantitative attribution falls outside the scope of this manuscript, we have expanded our discussion (Section 4.2) building on your suggestion.

To address whether the emission increases over the Cerrado and Caatinga could be an artifact of transport redistribution (e.g., compensating for errors originating in other regions) we evaluated the spatial cross-correlations in our posterior emission ensemble. Figure 5 demonstrates that the posterior emissions for the Cerrado generally have very weak correlations with other biomes, confirming that the inversion is capable of disentangling these regional signals from one another.

We note two exceptions to the weak correlation: First, a temporary strong correlation with the Caatinga occurs in mid-October. This coincides with a shift from the typical south-east trade winds to more easterly winds which likely causes the Caatinga air masses to mix with Cerrado to the extent that the inversion cannot separate the signals. Second, we observe occasional anti-correlations with savanna-type fires within the Amazon biome. This is an encouraging result: while our prior covariance matrix (P) structurally links these fires to the rest of the Amazon biome, the assimilation of TROPOMI X_{CO} seems to provide sufficient independent signal to break this prior assumption, which dynamically links their behaviour with Cerrado fires instead.

Finally, per the reviewer's suggestion, we computed the residuals and grouped these by biome (Table 2). We note that these should be interpreted with some caution, as grouping by underlying biome borders neglects transport. The table demonstrates that assimilation of TROPOMI reduces the mean absolute error across all major biomes, including a robust reduction over the Cerrado and Caatinga. This puts confidence in the inversion at the biome-level.

Table 2: Innovation statistics grouped by biome over assimilated observations. Including the mean absolute error (MAE) and mean uncertainty reduction (MUR) which is calculated as $\frac{1}{N}\sum_{i=1}^{n}\left(1-\frac{\sigma_i^{posterior}}{\sigma_i^{prior}}\right)$, with σ representing the standard deviation over the respective ensemble of realisations for each X_{CO} sample. The nature prior represents a single tracer forward run from the long-window solution (at 1°×1°), whereas the cycle prior is influenced by all previous short window cycles through mean state propagation. Based on the GFED5.1_TROPOMI inversion.

Biome	n (×1000)	MUR [-] (w.r.t. prior)	MAE [ppb] (nature) Prior	(cycle) Prior	Posterior
Amazon	43.7	0.89	13.0	12.6	7.8
Caatinga	0.3	0.65	14.3	9.4	6.4
Cerrado	14.9	0.80	15.9	10.6	7.0
Atlantic Forest	3.4	0.81	12.6	11.0	8.0
Pantanal	2.4	0.82	19.7	16.7	10.7

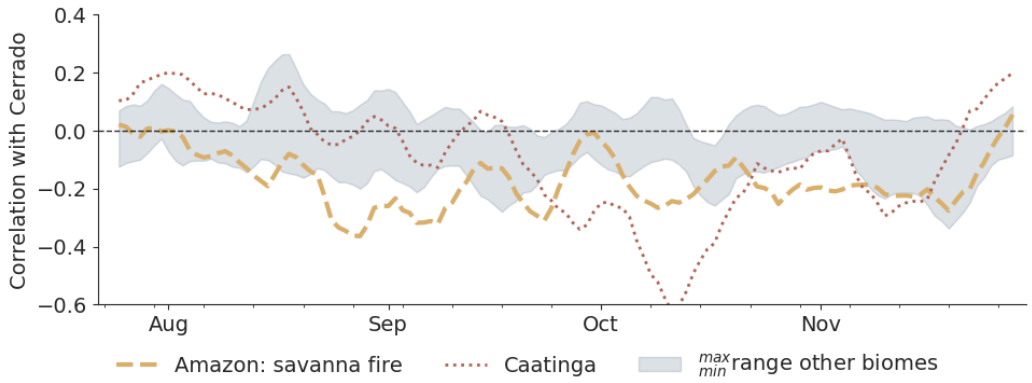

Figure 5: Six-day rolling average correlation between ensemble of posterior Cerrado emissions and that of other biomes. The shaded region represents the minimum-maximum range of correlations across other biomes/fire types in Brazil (incl. Amazon: other dominant fire type, Amazon: deforestation fire, Pantanal, Atlantic forest). Based on the GFED5.1_TROPOMI inversion.

Minor Comments

1.Clarify how the two-month MOPITT data gap in 2019 affects results.

We have added the following to Section 3.3 of the manuscript

Including this unconstrained period into our final multi-inversion posterior mean pulls the ensemble mean estimate toward the GFED5.1 baseline by up to 2.7 TgCO (difference between GFED5.1 and the mean of the two inventories for this period).

2.State earlier in the methods that the system cannot constrain super-local plumes due to resolution limits.

We included the following cautionary note in methods Section 2.4:

Note that our inversion system is not designed to resolve emissions from individual fire plumes, amongst others due to the 1°×1° resolution of our transport model and scaling factors, as well as biome level error correlations.

3.Clarify the origin of NMVOC and CH_4 pre-scaling factors.

The pre-scaling factors for secondary CO production from NMVOC and CH_4 oxidation were partly chosen to align our prior global budgets with established top-down estimates (e.g., Hooghiemstra et al., 2012; Zheng et al., 2019; Naus et al., 2022) and partly informed by visual inspection of prior and posterior CO mole fractions from initial forward runs. We have clarified this in the manuscript (Appendix A1).

4.In the conclusions, better define the applicable scale (regional totals vs. event-scale emissions) to avoid overgeneralization.

We appreciate the suggestion and revised the conclusion accordingly.

Technical / Presentation Issues
1.The "thick semi-transparent black line" in Figure 1 is difficult to identify; please adjust visibility.
We removed it to improve the figure's readability, assuming that the term *arc of deforestation* is generally known.

2.Provide bias and RMSE metrics for ATTO validation.
We have updated manuscript Figure 3 to include the requested metrics and added the following to Section 3.1.2:
Although the short-window inversion yields improved validation scores for most cycles (Fig. 3), from time to time, deviations from this pattern are to be expected for a non-assimilated site, and they underline the importance of aggregating posterior results before interpretation. In particular, we find a local reduction in performance in the last cycle which can be attributed to Ensemble Kalman Smoother (EnKS) edge effects, where the final cycle is constrained by only a limited set of observations. When combined with state propagation and a poorly represented fire in close proximity to the ATTO site, this can affect performance metrics at a local scale. However, the influence of these edge effects on domain total emissions remains minimal, as is evident from Fig. 5.

3.Add y-axis labels in Figure 5 for clarity.
Thank you for the suggestion. Note that the y-axes are shared across the ridges and the ridges are equally spaced by 0.3 TgCO day^{-1} to help interpretation. We have clarified this in the caption of the Figure.

References:

Hooghiemstra, P. B., M. C. Krol, P. Bergamaschi, et al. "Comparing Optimized CO Emission Estimates Using MOPITT or NOAA Surface Network Observations." Journal of Geophysical Research: Atmospheres 117, no. D6 (2012). https://doi.org/10.1029/2011JD017043.

Naus, S., L. G. Domingues, M. Krol, et al. "Sixteen Years of MOPITT Satellite Data Strongly Constrain Amazon CO Fire Emissions." Atmospheric Chemistry and Physics 22, no. 22 (2022): 14735–50. https://doi.org/10.5194/acp-22-14735-2022.

Zheng, B., F. Chevallier, Y. Yin, et al. "Global Atmospheric Carbon Monoxide Budget 2000–2017 Inferred from Multi-Species Atmospheric Inversions." Earth System Science Data 11, no. 3 (2019): 1411–36. https://doi.org/10.5194/essd-11-1411-2019