We thank the reviewer for the kind words, and for taking the time to provide feedback on the manuscript, which has helped improved its quality further.

Line-by-Line Comments:

Line 2. By how much do these estimates vary by?

We choose not to become more quantitative here, as that would require references in the abstract. There are several examples that indicate how there might be large uncertainties in carbon fire emissions. A good example is given in Ramo et al. (2021), which used a high-resolution burned-area dataset, derived from Sentinel-2, to show that carbon emissions in GFED4s might be underestimated by up to 31% over Africa, since the coarser-resolution MODIS burned-area product used by GFED misses small fires. More generically, uncertainties in biomass burnt translate to uncertain carbon emissions. Clearly, there is a role for top-down studies such as ours to better understand and reduce uncertainties.

Line 8. A fourfold increase over wet years?

Yes, we have made this explicit now.

Line 5. I'm confused whether "3-daily" refers to 3 times per day or every 3 days. Please make this clear here.

We have adjusted phrasing to indicate we mean every 3 days (here and throughout the manuscript).

Line 24-25. "albedo changes" needs to be corrected to burned area or surface reflectance. GFED (van der Werf et al., 2017) is primarily based on the MODIS burned area product. Burned area classification is derived from changes in surface reflectance. The references the authors list refer to quantification of fire emissions rather than pure monitoring of fires. For the latter, the authors should cite papers that describe the active fire and burned area products, such as Giglio et al. (2016) and Giglio et al. (2018). I'm not sure what the authors mean by the products being "partly related." Related in what way? Being able to serve as the basis for emissions estimates?

We have rephrased to separate between fire monitoring and emission estimation (L23-26). "Partly related" referred to the fact that bottom-up emission estimates sometimes share assumptions (e.g., landcover classification or emission factors), but we have removed this phrase for clarification.

Line 71. Please state the spatial resolution of the ERA-Interim reanalysis product.

We have provided this now in L74-75.

Line 77. The authors say GFASv1.2 is provided at 0.5° spatial resolution here, but GFASv1.2 is provided at 0.1° spatial resolution.

Corrected.

Line 88. “three-day” and “three-daily” are used interchangeably. Stick with one or the other.

We have adjusted to three-day total for clarity and consistency.

Line 90. Why a 0.03 Tg threshold specifically? Is this a statistical cutoff?

This is a pragmatically chosen number that we found includes a large area still (24% of the inner domain area), but not all grid cells. In this way, it is somewhat arbitrary, but rather than the exact value chosen here, the important conclusion is that we can start from a completely flat prior in time that contains very little spatial information, and still retrieve results close to our GFAS- or FINN-based estimates. This insensitivity to the prior is rare in inverse modeling studies. We do not see how a slightly different cut-off choice would impact that conclusion.

Line 95. Why were the GFAS emissions outside the domain not averaged? How much does the interannual variability of the emissions outside the domain influence the results?

We wanted to determine how uniquely MOPITT data determines the fire emissions in our domain of interest, i.e., which part of the posterior solution comes from the prior, and which part from the satellite data. In the outer zoom domain (i.e., the global domain) we are reliant on the variability from bottom-up inventories, as we only assimilate surface observations. Moreover, we are not interested in the posterior fire emissions we derive there.
We considered the middle zoom domain also part of the boundary conditions, whereas we wanted to test the influence of the fire prior. However, this is a design choice, and we could equally well have run this sensitivity test with a flat prior in the middle domain. Given that we find strong constraints from MOPITT in the inner domain, we consider that we would similarly find strong constraints in the middle domain, and thus that boundary conditions would be little affected, but we have not performed this test.
In the FINN inversion, the fire emissions in all domains are of course changed relative to the GFAS/CLIM inversions, and here we find similar emission estimates as in the GFAS inversion. Additionally, we have tested the influence to boundary conditions by adjusting the error on surface observations (Section 3.3.3 and Supplement 1), which resulted in little change in emissions. This indicates that most emission information comes from inside the inner domain.

Line 115. Please explain why a factor of the square root of 50 is chosen.

The square root 50 factor was pragmatically found to increase the observational cost function of an inversion that assimilates both surface and satellite data by a factor two relative to an inversion that only assimilates surface data (Hooghiemstra et al., 2012). I.e., this indicates that satellite data and surface observations have equal weight in the inversion. Without this factor, the large number of satellite data dominate the observational cost function, and the surface observations (especially near the Amazon domain) are not fitted within one-sigma.
In hindsight, it might have been appropriate to the error even more, since we find that in the base inversions surface observations near the Amazon domain are not always well-reproduced, indicating that the satellite data still have too much weight in the inversion. However, when we reduce the error on surface observations by a factor 10 (giving them more

weight in the inversion), we match the surface observations better, we still retrieve a good match with MOPITT, and we find similar posterior CO fire emissions. Therefore, it doesn't seem that the posterior fire emissions are very sensitive to this factor.
Since this factor is so specific to the inversion set-up, we do not consider it worthwhile to include a recommendation in the manuscript.

Line 163-171. The authors should quantify the interannual variability, e.g. standard deviation. In general, the authors should be more quantitative in describing their results.

We agree with the reviewer, and we have substantiated the manuscript section with more quantitative descriptions of the results.

Figure 4. The black circles in 4a shows how much CO is added to the fires, but a scale/legend is needed here.

We have added a legend to Fig. 4a.

Figure 5. It's hard to see the purple "Observed" line in the top panel. Since the difference between "Observed" and "Simulated" are shown in the bottom panel, just showing the "Observed" line in the top panel might be a better approach.

We now only show the "observed" line in the top panel.

References

- Hooghiemstra, P. B., et al. "Interannual variability of carbon monoxide emission estimates over South America from 2006 to 2010." *Journal of Geophysical Research: Atmospheres* 117.D15 (2012).
- Ramo, Ruben, et al. "African burned area and fire carbon emissions are strongly impacted by small fires undetected by coarse resolution satellite data." *Proceedings of the National Academy of Sciences* 118.9 (2021): e2011160118.