We would like to thank the reviewer for her/his review. We addressed each comment below, our answers and changes to the manuscript are highlighted in red. Additionally we added a reference to Anderson et al. (2024) that describes the GFFEPS emission prediction system (it was not available previously).

GENERAL COMMENTS

The manuscript makes many references to 'wildfires' and 'forest fires' (in the title) whereas the method does not select any fire types but covers basically everything. In this case I would think that 'biomass burning' would be more appropriate since it also covers controlled fires, and savannah fires.

Thanks for this suggestion, we agree that biomass burning is a better term. We have made the changes (from wildfire -> biomass burning and forest fires -> biomass burning, where appropriate) throughout the manuscript as well as in the title.

TROPOMI XCO is mentioned is a good source of independent information to test fire inventories, but the study does not actually perform such a test. Without this we still do not know which of the inventories in Figure 8 performs best. FRP is used to extrapolate TROPOMI information, but any of the other approaches could have been used for this also. It is an implicit choice that has been made without motivation. Better would have been to first test all the inventories against TROPOMI at the collected set of locations and then select the method that is most consistent with TROPOMI as a method for extrapolation.

We are a little unsure exactly what the reviewer is referring to here. But we assume they are suggesting that we use all the various inventories mentioned in our paper in the forecast model, and then compare the resultant CO with TROPOMI. This is an interesting study, but it by itself would be a very large effort and it is not the study we are engaging in. Using the output of such a study - say the inventory that produced the output that agreed best with TROPOMI - to scale up our emissions to regional and annual totals is problematic since these inventories are gridded, sometimes by hour and sometimes by day, and on varying spatial scale. this makes it very difficult to compare, and hence scale up, our emissions which are snap-shots for individual fires. Our method of using FRP to scale up or extrapolate is better suited since TROPOMI emissions and MODIS FRP can be directly linked in space and time.

In some places errors are assumed to cancel out in the mean, without proper justification. One example is the distinction between flaming and smoldering fires. Since the TROPOMI scenes that are used are quite a narrow selection of conditions for which the approach works best, it is unlikely to be representative of all fires. For example, the smoke generation in incomplete combustion is likely to be filtered out. Sampling biases are also likely to occur just because the size of fire and the meteorological conditions that are selected. I do not require to account for all this, but to spend a few sentences in the discussion to make the reader aware of such shortcomings in the current version.

As suggested, we included the following to the conclusion section:
" There is much uncertainty in this method as a single EC is assumed for each biome that is based on emission estimates at around 1:30pm local time and thus the TROPOMI scenes in general represent a small selection of conditions, it is unlikely to be representative of all fires or fire stages. The EC can also be affected by sampling biases because the size of fires and the meteorological conditions that are selected."

Another example is the error due to neglecting the vertical sensitivity of TROPOMI. No correction is made, whereas the error is always positive and therefore a bias. For this reason, it is insufficient to account for it as an uncertainty. I accept that a proper solution would require vertical profile information which can be solved later, but a 17% upward correction would be better than no correction.

Depending on the height of the plume and the averaging kernel, the effect of applying the AVK can be positive or negative. In our responses to reviewer 2 where an enhancement was where the AVK was low (near the surface) the VCDs increased, Rowe et al. (2022) found that on average the VCDs decreased when applying the aircraft observed profile. So depending on the shape of the profile and

averaging kernel the effect can increase or decrease the VCDs. Thus we are not applying a factor to correct the TROPOMI VCDs. We agree that the error is of systematic nature rather than random and added it to the random uncertainty rather than adding it in quadrature, this increased the total uncertainty to 57%.
We added the following to the text as well as changed Table 1.
Section 2.1: " Validation against TCCON measurements around the world showed that the TROPOMI CO columns have a high bias of about 10% (Sha et al., 2021) ."
Section 2.3: " The total uncertainty of the satellite-derived emissions (see Table 1), is based on the systematic bias and random uncertainty. The random uncertainties consist of the wind speed (~ 10%), the effect of the altitude used for the wind speed (~20%), and the uncertainty of the method itself (based on the relative difference between the true and fitted emissions of 34% after applying the above mentioned quality filters). The uncertainty of the wind speed caused by the uncertain altitude of the plume is based on the mean difference of the wind speed when comparing the winds 50\,hPa above and below the aerosol layer height. The uncertainty of the wind speed is based on \citet{gualtieri2022} who found approximately 0.5m/s for the 90% confidence interval for ERA5, with the average wind speed of approximately 5m/s (for our dataset), we assume a 10% uncertainty for the wind speed. These errors are added in quadrature, leading to a total random uncertainty of 41%. Additionally, the TROPOMI CO VCDs (comparison to TCCON) are biased high by about 10% (Sha et al., 2021), and not accounting for the averaging kernel correction due to the lack of profile observations for the fires will add another 6% (Rowe et al, 2022) (Table 3, difference between accounting and neglecting the averaging kernel). While the total emissions (or VCDs) could be scaled by the systematic bias, the effect of the averaging kernel correction depends on the profile as well as averaging kernel shape, also see averaging kernel analysis in Sect. 2.1 and Fig.B1). Adding the systematic and random error leads to a total uncertainty of 57%."

SPECIFIC COMMENTS

Line 73: The TROPOMI resolution is 7x5.5 km2 as mentioned later
Corrected.
Figure 1: Why use Gaussian functions rather than the observed enhancement over the background? Do not see the added value. I wonder also if the width of the integration box in panel d is wide enough for the tail of the plume. Tail dispersion outside the box could explain part of the apparent emission reduction with distance.
The Gaussian function is not used for the background correction. The Gaussian function is used to find the plume width, the correction for the wind direction, and correction of any offset from the centre (in case the estimated location from weighted MODIS thermal anomalies is not quite correct).
This is explained in the manuscript l. 151-170.
We have included the following sentence t make this clearer this in the manuscript:
" New improvements with regards to the plume rotation and plume widths are included, and illustrated in Fig. 1 where Gaussians are fitted across the plume to be able to automate the estimation by determining the plume width, correct the wind direction and correct the centre location of the fire."
Fig. 1 caption: " (b) simple Gaussians are fitted across wind in 4-km wide boxes up to 40 km downwind of the fire to find the plume width, correct the wind direction and fire location"

Line 167: An upwind mixing ratio as background can be quite inaccurate if the plume is superimposed on a non-uniform background. This might also cause apparent temporal variations in the emissions.
We tested various methods to determine the background CO concentrations including Gaussian fit, using the 90th percentile and the upwind CO concentrations. These were also tested with the CTM simulation (see Section 2.3) and we found that using the average upwind concentrations was the most stable method with the best results. If the background is too large the results are filtered.

Line 257: The error calculation missing 10% VCD uncertainty from the TROPOMI retrieval. (the 17% is only from the neglected Ak)

We considered scaling our result to account for this systematic effect but in the end decided it was best for the data user to apply such a correction, either the generic 6% value from Rowe et al. (2022) or 10% from Sha et al. (2021) or perhaps one better suited to their specific application.

In Rowe et al. 2022 the difference between applying the AVK (quality assured and using AVK: 10+-15%) and not applying the AVK (quality assured 16+-15%) is approximately 6% see (Table 3), in the 16% included in this study the 10% bias from the TCCON study is included and rounded up. However, instead of assuming this is a random error we included it as a systematic error and added it to the random uncertainty (rather than adding it in quadrature). This increases the total uncertainty to 57%.
We added the following to the manuscript and increased the total error (from 44 to 57%) throughout the manuscript:

Line 348: "red dot" io "orange dot"?
Corrected.

Line 357: 'this bias' is which bias?
Changed to: " The cause of the differences between the TROPOMI-derived and GFFEPS emissions are being investigated"

Line 369: What could be learned from AUST having a very poor correlation?
The model does currently not capture the emissions from Australian eucalyptus well.
We added the following: " Other regions, like AUST have a very poor correlation, slope and RMSE, indicating a need to improve the modelling of that region, which is currently still under development, such as improving the emission factors, correcting the fuel consumption and combustion completeness for eucalyptus (Anderson et al., 2024)."

Line 416: Increases instead of reduced this error.
Reduces is the correct term here. The smoldering CO emissions are very high, but the FRP for smoldering is very low thus the increase is not as significant. We changed the term to " has a smaller effect on the total emissions" to avoid confusion.

Figure 6: "blue being the outliers" You mean purple? Not clear what 'density' means here (I suppose frequency?). What is 'S'? The plot does not look like R=0.7. The axis are plotted such that we are essentially only looking at outliers. This should be improved.
The word outlier is not used correctly, this has been corrected.
The caption has been corrected to: "The black line indicates the slope ("s") of the best fit. … (yellow being the most frequent and purple being single points)"

Table 2: With rank you mean the fraction of total RFP?
Thanks for pointing this out, we clarified this more:
We added the following to the caption in Tables 2, 3 and 4: " rank of importance in terms of total FRP with 1 being the biome contributing the most to the emissions (``Rank'')"
In the text: " rank is number of importance with respect to the total annual FRP from GFAS with 1 being the biome contributing the most (for the 2019 base year)."

Line 566: 'by this' is by what?
Somehow this sentence ended up in the wrong location, it has been moved a few lines down and corrected:
" We also found that the FRP is strongly influenced by thick smoke which can influence these types of top-down emission estimates, and leads to an underestimate of fire emissions for fires with thick smoke (typically large fires). The directly-derived TROPOMI CO emission estimates are not impacted as much by the by the smoke and have been used here to verify and analyse this issue for individual fires."

Line 586: 'are at least' what?
Corrected to: " The TROPOMI-FRE top-down emissions uncertainties (based on the uncertainties of the ECs) are at least 50% for regional emissions or 25% for global emissions."

TECHNICAL CORRECTIONS

Line 181: The sentence starting with ‘Any comparisons to emissions…’ has been duplicated
Corrected
Line 345: GFFEPS io GFEEPS
Corrected.
Line 370: Table D2)
Corrected
Line 444: ‘evergreen’
Corrected