We extend our gratitude to Reviewer 2 for their detailed review. In response to the feedback from both reviewers, we have implemented significant revisions to the manuscript. We offer detailed insights into these changes below, and for your convenience, we have also provided a PDF document illustrating the differences between the previous and current versions.

Our answers are highlighted in red font, while the reviewers' comments are in black.

This manuscript focuses on carbon monoxide (CO) emissions from wildfires; its goal is to produce a ~20 year global budget based on emission coefficients derived from TROPOMI and MODIS data:

- First, TROPOMI-derived emissions are calculated for the 2019-2021 period with an automated version of the flux method. To verify the robustness and uncertainties of the automated method, emissions are derived from synthetic CO column values generated with the GEM-MACH model and CFFEPS for the May-September 2019 period over North America. From this test, several filters (dealing with background CO, wind, plume geometry, etc.) are defined.

- Then, global TROPOMI-derived emissions are calculated for the 2019-2021 period, the filters defined in the previous step are applied, and CO emission coefficient values (EC) are calculated separately for 15 biomes: EC=E/FRP=emissions/fire radiative power. FRP values are from MODIS Aqua. Somehow

the GFAS FRP dataset (which results from assimilating both MODIS Terra and Aqua FRP values in GFAS) is used to provide "a guidance on total daily FRP that can then be combined with the derived ratio between TROPOMI CO emissions and MODIS FRP"; meaning unclear from the manuscript.

To make it clearer, we added:

"… by applying the derived emission coefficients to assimilated daily FRP based on MODIS measurements (available from GFAS)."

- Following, a global budget of CO emissions from fires is calculated for the 2003-2021 period based on TROPOMI-MODIS Aqua EC values and GFAS FRP values. Results are analyzed by region and by biome, and compared with respect to several emission inventories.

The manuscript should be improved to avoid repetitions, typos, correct wording and sentence integrity, provide explanations for acronyms/abbreviations when first used, etc. The text is hard to follow at times because of these issues. Some of the Appendices are never mentioned in the text. Expressions such as “many of the retrieved emissions”, “many outliers”, “aligns very well”, “the model agrees pretty well” should be avoided. Quantitative statements should be used instead.

We reviewed the manuscript, revising numerous sections to enhance its quality by eliminating repetitions, rectifying typographical errors, and refining the wording. Details can be found below, a pdf showing the difference between the last and current version is provided.

We specifically changed the expressions high-lighted by the reviewer:

“*Many of the retrieved* emissions are very close to the original emissions, however, *many outliers* can be seen, where most likely the retrieved values are below the originals.” -> “While a substantial portion of the retrieved emissions closely matches the original values, there are noticeable outliers, where the retrieved values are below the original emissions.”

“This shows that for certain regions (i.e. CEAM, NHSA, EURO, and MIDE) the *model agrees pretty well* with the satellite-derive emissions (taking slope, R and RMSE into consideration).” -> “Considering factors like slope, R (correlation coefficient), and RMSE (root mean square error), the model demonstrates strong agreement with satellite-derived emissions for specific regions, namely CEAM, NHSA, EURO, and MIDE (see Table B1).”

Some figures would benefit from more consistent axis ranges (e.g., Fig. 1 b and c), axis labels with both axis title and units (Fig. 5). Panels labels should be properly referred to in text (e.g., Fig. 5). For clarity, please add a map showing all the regions discussed in the manuscript (CEAM, NHSA, EURO, …). Consider including a figure illustrating the areal extent of biomes.

Fig1c: axis range has been adjusted to be similar to 1band a grid included.

Fig5: missing axis labels and units have been added.

Figures have been included showing the regions and biomes (Figs. A1 and A2), and refer to these within the manuscript.

The manuscript should justify why the TROPOMI averaging kernels are not applied to retrievals with qa_value <1. From the TROPOMI CO readme document: "We recommend using only data with a qa_value = 1 in case the averaging kernel is not applied. Data with a qa_value = 0.7 are of similar quality provided the averaging kernel is used to account for the vertical retrieval sensitivity in the presence of mid-level clouds. Quality assurance values of qa_value = 0.4 represent experimental data to be used with caution." Table 3 in that same document provides additional information regarding qa_values, cloud heights, and tau_aer values.

While the masking effect of smoke and clouds on FRP observations is discussed several times, the effect of smoke and clouds on TROPOMI retrievals is never mentioned. This is a very important issue, given the focus on TROPOMI observations acquired over active fires, with very smoky and (potentially) very cloudy conditions. Please discuss.

While for clear sky the TROPOMI CO averaging is around 1, the averaging kernel is affected by clouds. Near fires there is mostly smoke, which is treated as clouds in the TROPOMI retrieval algorithm. Smoke is made of small particles (like 0.1 um) and so should be all but invisible to light at 2 um. Cloud particles are much larger and so you cannot use the cloud Aks to estimate the AK with smoke, thus it would not be helpful to correct for the averaging kernel in the case of fire emissions as those might be wrong to begin.

Schneising et al., 2020 (https://doi.org/10.5194/acp-20-3317-2020) found the TROPOMI averaging kernel to be 0.95±0.05 in the boundary layer, and thus at the most will only account for a 5% uncertainty which is much lower than the overall uncertainty of the emissions estimate.

We also found a recent publication by Rowe et al., 2022 (https://pubs.acs.org/doi/full/10.1021/acsearthspacechem.2c00048) that investigates the uncertainty of the TROPOMI CO observations in smoke, we added the reference to our manuscript, showing approximately 10% higher CO than the aircraft measurements. This uncertainty is consisted with the uncertainty assumed in our error budget analysis (Table 1).

We included the following discussion in the manuscript: "For our analysis, we have utilized observations rated with a quality flag greater than 0.5, where 0 represents the lowest quality and 1 denotes the highest quality. This choice aligns with the recommended quality threshold (Apituley et al., 2018). Notably, when we investigate areas near active fires, the quality flag of the retrieval can be impacted by the presence of smoke. Consequently, including observations with a quality flag of 1 would result in the exclusion of a substantial number of data points, primarily due to the influence of smoke.
The CO averaging kernel from the TROPOMI observations predominantly registers

values close to 1 within the boundary layer, specifically around 0.95 with a narrow range of variability (approximately ±0.05) (Schneising et al., 2020). Nevertheless, the presence of clouds diminishes the sensitivity of the averaging kernel beneath them. It is important to note that smoke primarily comprises minuscule particles (~1 μm), which are invisible at the 2μm wavelength. The TROPOMI algorithm lacks the capability to differentiate between clouds and smoke, which is why attempting to correct the averaging kernel in regions near fires would introduce additional uncertainty to the analysis. Thus, the averaging kernel is not considered in this study, but is taken into account in the overall uncertainty (uncertainty of VCDs) of the emission estimate (see Table 1). Rowe et al. (2022) investigated TROPOMI CO in thick fire plumes and found agreement within 10 % with the aircraft observations, which has been used to estimate the overall uncertainty of the emissions."

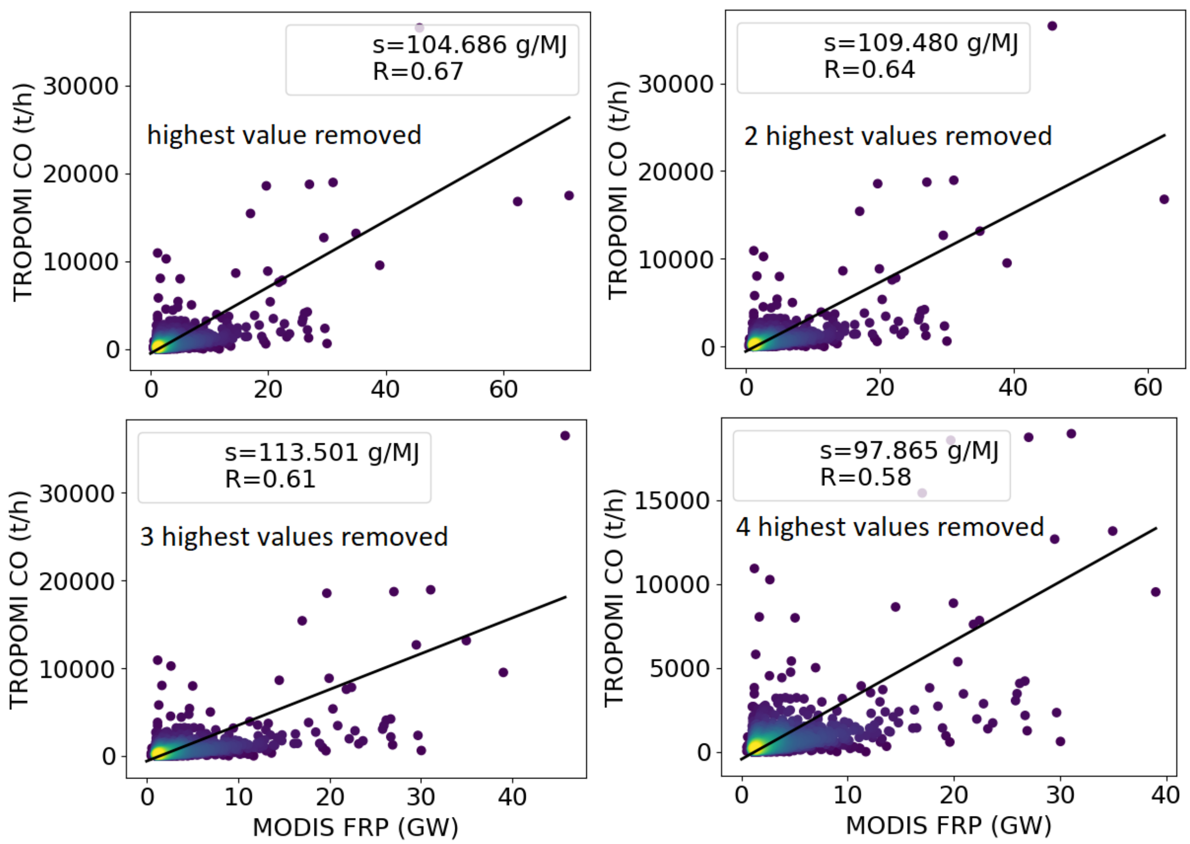
In lines 319-321, to explain discrepancies between emissions from TROPOMI and GFFEPS, it is suggested that GFFEPS values may be low because of missed fires due to thick smoke. Could TROPOMI values be high due to smoke? Clouds? Other factors?

GFFEPS relies on hotspot detection for the emission estimation (not missed fires, missed hotspots), as shown in Fig. 5 (bottom panel), some of the hotspots are under a layer of thick smoke meaning these are not accounted for in GFFEPS, reducing the emissions. The figure also shows that the TROPOMI observations are high over the fire and the smoke plume, and TROPOMI provides good quality data that meet the quality threshold. If anything, the thick clouds would lead to lower CO VCDs (not higher) as the averaging kernel would have lower sensitivity in the boundary layer below the cloud. The TROPOMI CO emissions approach has been validated for fires in North America using measurements from FIREX-AQ (Stockwell et al., 2022), the uncertainties associated with the emission direct estimate are discussed in section 2.3, showing an overall uncertainty of 40%. GFFEPS on the other hand is a new fire emissions prediction system and has not been validated, and neither does it have a range of uncertainties. We summarized the previous discussion and included this in the manuscript:

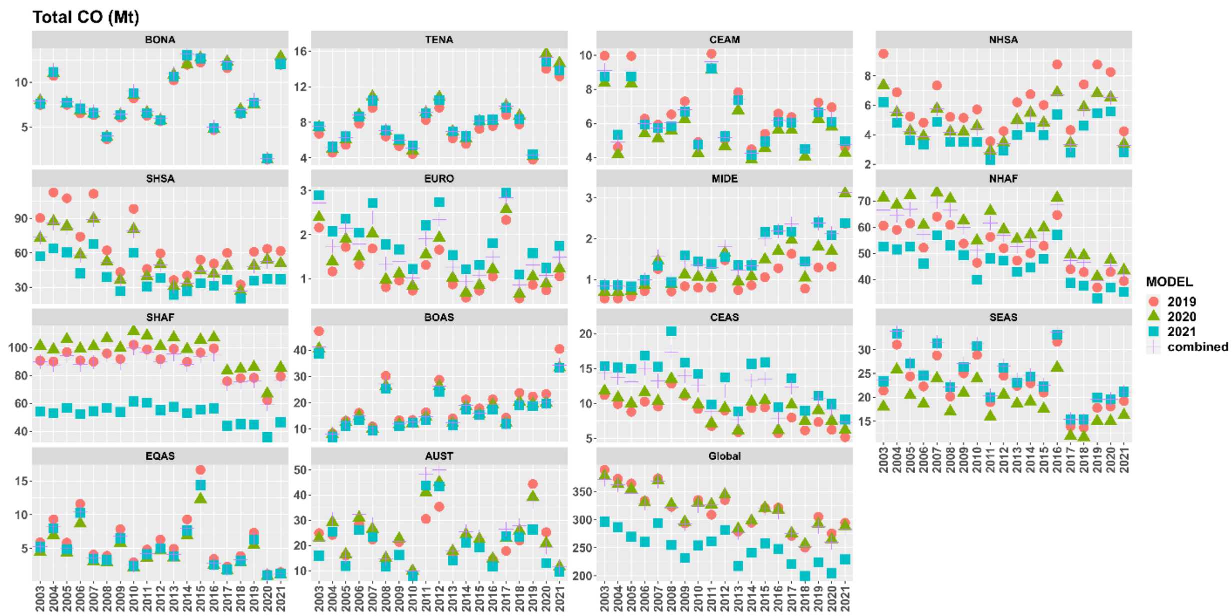
"This discrepancy is likely due to an underestimation from GFFEPS rather than an overestimation of TROPOMI emissions: TROPOMI offers high-quality data over fires and smoke plumes (see Fig 5) and potential cloud cover would likely result in lower CO levels detected by TROPOMI due to reduced sensitivity below the smoke plume rather than higher. It is important to note that the TROPOMI CO emissions approach has been validated (Stockwell et al., 2022) with a 40% overall uncertainty see Sect. 2.3), while GFFEPS still requires validation and associated uncertainty estimates."

Figure 6 shows that the correlation between CO and FRP may not be very robust, despite R being 0.70. How would the slope (i.e., EC) change if the most extreme outliers were removed one by one?

We removed the highest values (although it's hard to say that these are outliers, these might just be large fires for which the emission retrieval is often not successful or they might not be as common). The result is shown below. While R reduces as more values are removed by 0.12, it is still above 0.5. The slope increases when the highest values are removed, however, when all 4 values are removed the slope is back to the original slope shown in the paper. The increase of the slope is at maximum 16% and below the estimated uncertainty and range of values for other years.



Furthermore, we included the analysis for separate years to analyse the impact: while 2019 and 2020 are not too different, 2021 shows smaller emissions, primarily and most significantly for the SHAF region. The global totals are almost identical for 2019, 2020, and 2019-2021. The following figure is now included in the appendix and can be interpreted as an indicator of uncertainty for the total emission estimates.



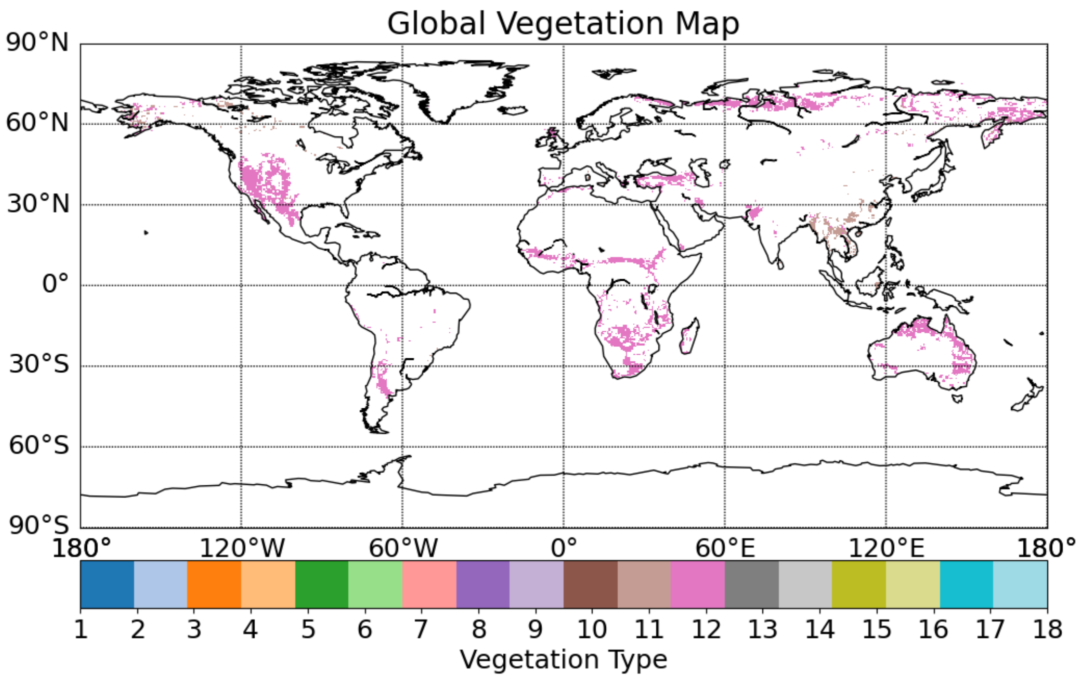
Line 379: “The emission coefficients vary between 120 and 39 g/MJ” Please clarify that EC=120 value corresponds to (11): shrub cover, closed-open, evergreen; EC=39 to (12): shrub cover, closed-open, deciduous. The most extreme EC values correspond to rather similar biomes, please discuss.

The error has been corrected: “The emission coefficients vary between 120 and 39 g/MJ, where the largest CO emissions relative to FRP are from shrub cover evergreen (11) and the lowest are from shrub cover deciduous (12).”

Even though both types are shrub, they are very different types. Type 11 is not very common (~0.5% of the Earth surface area) and predominantly occurs in central Asia, in some parts of Northern Canada and Alaska. Deciduous shrub is more common (~2.2%) and spread around the entire globe. This is showing that just using shrub as a vegetation classification might not be a good idea when estimating fire emissions. This can be seen in the figure below (only vegetation 11 and 12 are shown – otherwise it’s the same figure as Fig. A2)

We included the following in the manuscript:

“Even though both biomes are shrub, they are quite different biomes, based on their CO emissions and way hey burn as well their location and occurrence. Evergreen shrub (biome 11) is not very common (it covers approximately 0.5\,\% of the Earth’s surface) and appears primarily in Central Asia and in some parts of Northern Canada and Alaska (see Fig. A2 and Table A1). Whereas deciduous shrub covers approximately 2.2% of the Earth’s surface and grows globally (see Fig. A2 and Table A1). …. A simplified classification of forest, shrub and grassland is not appropriate based on our results. For example the ECco for different types of shrubs has both the largest and smallest emission coefficients and forests vary between roughly 49 and 95 g/MJ."



Lines 379-380: "where the largest CO emissions relative to FRP are from broadleaved evergreen tree cover (1) and the lowest are from cultivated managed areas (16)". How could readers see CO emissions relative to FRP? Similar issue in lines 474-475.

We included the following in the introduction where the term appears for the first time:

"Additionally, we determine biome specific emission coefficients (emissions relative to FRP), which are the CO emissions produced relative to the amount of heat energy released by the fire (FRP). This can provide insights into the efficiency of combustion, and help quantifying how emissions from a particular ecological region or biome are related to the heat energy generated by wildfires in that region. This information can be valuable for understanding the environmental impact of wildfires in different ecosystems and for developing strategies to manage and mitigate their effects. Furthermore FRP is often and more easily measured from satellites compared to CO, and determining a biome specific CO-to-FRP ratio can help to determine the daily total emissions of fires."

We added to the sentence in Sect 4 (near the former l.379-380):

“The emission coefficients vary between 120 and 39 g/MJ, where the largest CO emissions relative to FRP are from shrub cover evergreen (11) and the lowest are from shrub cover deciduous (12), meaning three times more CO is emitted from evergreen shrub (biome 11) compared to deciduous shrub (biome 12) for a fire that burns with equivalent heat energy.”

Line 401: “Most CO emissions are from evergreen forests (biome type 1), which also has the largest EC for CO”. Table 2 shows that biome 11 has the largest EC, please correct.

Thank you for pointing out this error, we corrected the statement to: “… which also has one of the largest EC for CO…”

Please justify why is biome 15 (“Regularly flooded shrub and/or herbaceous cover”) included in the analysis. Biomes 7 and 8 are excluded because “fires were not observed, namely: regularly flooded tree cover (7 and 8)”

No fire emissions were detected for these types of biomes (7, 8, 10, 19, 20, 21, 22) by TROPOMI. For regularly flooded shrub and herbaceous cover TROPOMI observed 127 fires. We added the following statement in the manuscript: “…fires were not observed by TROPOMI (and therefore no information is available on the CO emissions) …”

Biome # 15 (regularly flooded shrub and/or herbaceous cover; should this biome be included in the analysis, given that it is regularly flooded?) has the second largest EC (105 g/MJ). Please explain.

The biome is included because fires and CO emissions could be detected for this type of biome. With TROPOMI we were able to estimate emissions from 127 fires.

The manuscript assumes that EC values “do not change drastically over the years”. However, tables C1, C2, and C3 in Appendix C (never mentioned in the manuscript; please correct) show otherwise. EC values change between -50% and +150%, depending on the period analyzed. EC values change strongly among biomes; changes do not seem to follow recognizable patterns. How would the global budget change if a different set of EC values was used? Please quantify.

We estimated emissions with 2018, 2019, 2021 values, and included the comparison in the appendix (Fig. C1). Even though the EC change over the years the impact in the total emissions is not as significant. The largest difference appears for 2021 where the emission coefficients particularly for biomes 1-3 is about half as much as for the other years. This reflects to a 20% change in the emission totals.

We included a brief introduction in Sect. 4:

“The slope between the CO emissions (in g/s) and the FRP (in MJ/s) is EC (in g/MJ), the values of which are shown in Table 1 (and for 2019, 2020, and 2021 in Tables C1, C2, and C3, respectively).”

And added a discussion in Sect. 5.1:

“To assess the uncertainty of the total annual emissions of our estimates (TROPOMI-FRE), we also used emission coefficients derived from fires of individual years (2019 to 2021). Using emission coefficients from 2019,2020, and 2019-2021 combined did not impact the total emissions (see Fig. C1), only for 2021 the total emissions reduced by approximately 20%, due to overall lower ECco (for biomes 1-3, see Table C3). This shows that the uncertainty of our approach is at least 20%, but since the individual TROPOMI derived CO emissions have an uncertainty of 40%, we would expect the overall TROPOMI-FRE annual emission to have similar uncertainties on the order of 40%.”

What is the areal extent of each biome? That information could help readers understand what biomes (and what biome’s EC) may have a stronger effect on the global budget.

We included this information in Table A1. In units of km2 as well as coverage of the Earth's surface in percent.

Figure 9 shows emissions decreasing with time. An evaluation of the temporal trend in the global budget with respect to actual measurements (not inventories) is missing in the manuscript. Have wildfire emissions really decreased in the last two decades? While measurements of global emissions for all biomes may be unavailable, a literature search may provide useful information for specific regions/biomes.

All inventories presented rely on satellite observations: GFAS (and ultimately TROPOMI-FRE) are based on satellite observations of fire hotspots (MODIS). GFFEPS is based on satellite hotspot observations, and (satellite-observed) area burned. FINN 1.5 is based on MODIS and FINN2.5 is based on MODIS/VIIRS observations. GFED also include satellite observed area burned.

The MOPITT satellite has been observing CO since 2003 Satellites alone cannot obtain total CO fire emissions, looking at the total global CO concentration or vertical column density would not distinguish the source: anthropogenic vs wildfire. MOPITT observations of CO show decrease between 2002-2018 which is consistent with our findings (Buchholz et al., 2021).

We re-wrote the entire section 5.1 "CO emissions over the past two decade". We changed Figure 9 to include all inventories discussed in this paper (with the exception of GFFEPS that is currently only available for 2019). We included a Figure, showing the change of emissions for different regions from all inventories discussed in this paper (GFAS, GFED, FINN 1.5, FINN2.5). Additionally, we provided a table with the trends and indicate whether or not the trend is significant (for all 5 inventories, including our own).

We included a literature search finding that area burned and MOPITT CO are declining [Buchholz et al., 2021, Zheng et al., 2021, Giglio et al., 2013], which is consistent with our findings. However, we did not find a publication focussing fire emission trends for the same time period as our study.

Section 5.1:

"5.1 CO emissions over the past two decades

The inventories discussed in the previous section provide data for various past years, except for GFFEPS (currently only available for 2019). For our independent estimates, we relied on daily FRP data from GFAS, which is based on MODIS FRP, available from 2003 to the present. Under the assumption that the ECco values (as derived in Sect. 4) remain relatively stable over the years, we conducted an extensive analysis of the entire time series and calculated CO emissions spanning from 2003 to 2021 (refer to Fig. 10). To assess the uncertainty of the total annual emissions of our estimates (TROPOMI-FRE), we also used emission coefficients derived from fires of individual years (2019 to 2021). Using emission coefficients from 2019,2020, and 2019-2021 combined did not impact the total emissions (see Fig. C1), only for 2021 the total emissions reduced by approximately 20 %, due to overall lower ECCO (for biomes 1-3, see Table C3). The uncertainty of our approach is at least 20 %, but since the individual TROPOMI derived CO emissions have an uncertainty of 40%, we would expect the overall TROPOMI-FRE annual emission to have similar uncertainties on the order of 40 %. Furthermore, we also present data from the other four inventories for the same time frame. The results are visualized in Fig. 9. As expected, the emissions from wildfires in various regions across the globe exhibit significant interannual variability.

Notably, EURO and MIDE consistently report the lowest wildfire emissions throughout the entire time series and are barely noticeable in the figures. The predominant source of wildfire CO emissions is from SHAF and SHSA, followed by NHAF. This consistent pattern is evident for all the inventories analyzed.

To enhance the clarity of emissions identification and changes across different regions, we have depicted emissions by region in Fig. 10. The rate of change for this time period has been quantified for each inventory, and the results are presented in Table 3. Significant rates of change (with a p-value below 5 %) are highlighted in bold, while all other rates of change are statistically insignificant.

Globally, CO emissions are experiencing a decrease ranging from 5.1 to 8.7 Mt(CO)/yr between 2003 and 2021 across all inventories, with the exception of GFED. Notably, GFED does not reflect a global decrease due to the substantial increase in CO emissions within the BOAS region, amounting to 19.8 Mt(CO)/yr. This overall decrease is primarily driven by significant reductions in SHSA (ranging from 2.1 to 6.3 Mt(CO)/yr), NHAF (ranging from 0.6 to 7.6 Mt(CO)/yr), SHAF (ranging from 0.9 to 5.6 Mt(CO)/yr), and CEAS (ranging from 0.3 to 3.3 Mt(CO)/yr), all of which show statistically significant decreases across at least four inventories. In contrast, CO emissions from wildfires are on the rise in TENA, with an increase ranging from 0.2 to 4.1 Mt(CO)/yr. Additionally, emissions in the EQAS region exhibit an interannual cycle that appears to correlate with El Niño years, resulting in higher emissions across all inventories in 2006, 2009, 2014, 2015, and 2019.

These findings align with prior research. Giglio et al. (2013) reported a decreasing trend in the annual area of land burned since 2000, which corroborates our observed reduction in CO emissions. Moreover, Zheng et al. (2021), also observed a decline in burned area between 1998 and 2015 through satellite observations, but reported stable or only slight decreases in wildfire emissions. The satellite instrument "Measurement of Pollution in the Troposphere" (MOPITT) on board the TERRA satellite (Drummond et al., 2010) has been observing CO since 2000, (Buchholz et al., 2021) showed that MOPITT CO has been steadily decreasing by -0.50% per year between 2002 to 2018. No study examining fire emissions for the time period presented here currently exists to our knowledge."

Furthermore, we included some changes in the conclusion section to reflect the changes in section 5.1.

"Examining the trends over the past two decades (corresponding to the MODIS lifetime), it appears that global CO wildfire emissions have, on the whole, decreased. This decline is consistently observed across all inventories utilized in this study. However, this trend is highly region-specific, with the most substantial reductions occurring in SHSA, SHAF, NHAF, and CEAS. Conversely, wildfire emissions in TENA are on the rise. For all other regions, the variability within the past two decades has been too substantial to determine a statistically significant trend."

References:

Buchholz, R. R., Worden, H. M., Park, M., Francis, G., Deeter, M. N., Edwards, D. P., Emmons, L. K., Gaubert, B., Gille, J., Martínez-Alonso, S., Tang, W., Kumar, R., Drummond, J. R., Clerbaux, C., George, M., Coheur, P.-F., Hurtmans, D., Bowman, K. W., Luo, M., Payne, V. H., Worden, J. R., Chin, M., Levy, R. C., Warner, J., Wei, Z., and Kulawik, S. S.: Air pollution trends measured from Terra: CO and AOD over industrial, fire-prone, and background regions, Remote Sensing of Environment, 256, 112, https://doi.org/https://doi.org/10.1016/j.rse.2020.112275, 2021.

Giglio, L., Randerson, J. T., and van der Werf, G. R.: Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (GFED4), Journal of Geophysical Research: Biogeosciences, 118, 317–328, https://doi.org/https://doi.org/10.1002/jgrg.20042, 2013.

Zheng, B., Ciais, P., Chevallier, F., Chuvieco, E., Chen, Y., and Yang, H.: Increasing forest fire emissions despite the decline in global burned area, Science Advances, 7, eabh2646, https://doi.org/10.1126/sciadv.abh2646, 2021.

Lines 446-447: "certain regions see increased emissions (e.g. TENA, AUST)" (text refers to Fig. 9). Hard to know for sure, but it looks like neither TENA nor AUST show increasing emissions. Either provide a better figure to illustrate the statement or remove statement. Similar issue in line 498.

We re-wrote the entire section 5.1, see above. We included an additional figure that makes it easier to see the trends by regions by plotting the emissions by region normalized to 2003. We included the slope of the line of best fit in a table in the appendix.

Additionally, to the TROPOMI/FRE dataset we also looked at GFED, GFAS and FINN between 2003 to 2021.

We changed the conclusion section based on the changes from this section.

Lines 476-477: “for forests we determined ECs between 64 and 120 g/MJ” Please note that EC=120 is not for a forest but for biome 11: Shrub Cover, closed-open, evergreen.

Thank you for pointing out our typo, we corrected this in the manuscript:

“(e.g. for forests we determined ECs between 64 and 95 g/MJ)”

The manuscript refers to the need for multiple observations in a day (from geostationary instruments) in order to better understand fire evolution; TEMPO is mentioned but GEMS (which has been operating for a couple of years now) is not. Please discuss both.

We appended the sentence as suggested:
“Geostationary satellite sensors, such as TEMPO (covering North America), Geostationary Environment Monitoring Spectrometer (GEMS), or Sentinel-4 (covering Europe and Africa) will help to validate the diurnal pattern of emissions.”