Review of manuscript "Towards an improved understanding of wildfire CO emissions: a satellite remote-sensing perspective" by Griffin et al.
Revised version.

This manuscript describes a methodology to evaluate CO emission coefficients (EC) that may be used to estimate CO biomass burning emissions from observations of the fire radiative energy (FRE). This method relies on an inversion of CO emissions from TROPOMI satellite observations of total columns CO for a large set of fires detected in different biomes.
A first section describes the methodology for CO emissions inversion as well as an analysis of the corresponding uncertainty. The derived EC are then presented. Finally, the method is applied to the 2003-2021 time period based on the GFAS FRE database.

This study brings a valuable contribution to the efforts on improving the estimates of biomass burning emissions, for which the uncertainty is still estimated to be about a factor of 2. The approach chosen is original because it targets ECs and not just final emissions, so that it may be used in other emission inventories and time periods not covered by CO observations. I think this work is worth publishing in ACP, and that the authors did a significant effort to improve the manuscript and add new material to respond to the reviewers' comments. However, I also think that major revision is necessary before publication, which I'm confident the authors will be able to address.

Major comments.

I think the title should be revised to be closer to the added contribution of the manuscript, i.e. information on CO emission coefficient derived from TROPOMI observations. It is not clear if emissions are improved, which does not prevent the manuscript from making an interesting contribution.

Emissions inversion:
Section 2.3, uncertainty evaluation using synthetic data:
The authors were able to construct a database of 208 fires between May and September 2019, and only 105 remain after the filtering. It's not clear to me how these fires are representative, especially for regions in which the fire season is not during boreal spring and summer and for large fires (both very low wind and high wind may be favorable to large fires, e.g. during heatwaves).
Several criteria are listed for successful retrieval (l. 230—240 & FRP > 1000MW): are all inversions successful if these criteria are met?
Section 2.5, evaluation top-down vs bottom up:
In this section, the authors state that about 5000 fires are analysed (l.336); but approx. 4000 fires according to the legend of Fig.4… What is the exact number?
How many fires were detected during that time period, in total, and how many above the threshold of 1000MW (which is already very high)? What fraction of fire detection may be analysed using this method?

EC estimates:
This estimate relies on a landuse map to attribute specific fires to different biomes. For this purpose, they use the GLC2000 database. How is this quite old land use map representative

of vegetation in 2019-2021? I think this adds serious doubts to the results presented since it has been demonstrated that vegetation attribution is also a very large contribution to uncertainties in emission estimates (e.g. Turquety et al., GMD, 2020). There are many more recent land use classifications that may be used (e.g. from MODIS).

I think that table 2 is the most important result of the manuscript. Since the same approach is used in other work (e.g. GFAS), it would be necessary to compare results to previously published values. I understand that the authors mention that a coarser classification of biomes degrades the correlation coefficients, but a comparison is necessary in order to understand the potential added value of this work in estimating uncertainties on EC used in the literature. If I understand correctly, GFAS uses a conversion coefficient to estimate dry matter consumed from the FRE; and then derives emissions using tabulated emission factors. An equivalent emissions coefficient for CO could be estimated (or at least an order of magnitude).

Uncertainty analysis:
I appreciated finding an analysis of the method's uncertainties and of the uncertainties on the final emissions, which is a very difficult exercise.
It is evaluated through:
Section 2.3: academic case study with synthetic data allow an estimate of uncertainty on CO emission's inversion to 42%.
Section 3: detailed comparison of inversions with bottom-up inventory GFFEPS
Section 5: intercomparison of annual emissions for 2003-2021 using 5 other emissions inventories.
Throughout the paper, the authors estimate an uncertainty to 40% or 42% (a consistent number would probably be better). However, I think it is strongly underestimated.
This estimate assumes that the only uncertainty in the calculation of the EC values is the uncertainty on the inversion but I don't think that this is fully demonstrated. For example, the authors mention that an inversion is not possible for large wildfires because of multiple plumes. These wildfires are likely to emit a very large mass of CO. May this filter (and others) induce a bias? Does the fraction of fires with successful inversion depend on the biome?

The authors discuss the difficulty of such exercise due to overpass times and plume transport (p. 14). Would the very classic approach of comparing plumes simulated using a CTM may be more adapted in this case? It allows to compare the resulting enhancements regardless of under-constrained parameters (like diurnal variations in this case).

The authors mention that there are no uncertainty estimates for other inventories which is not true. Many publications use these inventories to simulate BB plumes using chemistry-transport models, that are compared to atmospheric observations. Some studies are included in the publications describing the inventories. It is commonly assumed that uncertainties on BB emissions are a least a factor of 2.
Different recent papers present an intercomparison of BB inventories and discuss uncertainties, e.g. :

- Wiedinmyer, C., Kimura, Y., McDonald-Buller, E. C., Emmons, L. K., Buchholz, R. R., Tang, W., Seto, K., Joseph, M. B., Barsanti, K. C., Carlton, A. G., and Yokelson, R.: The Fire Inventory from NCAR version 2.5: an updated global fire emissions model for climate and chemistry applications, Geosci. Model Dev., 16, 3873–3891, https://doi.org/10.5194/gmd-16-3873-2023, 2023.

- Pan, X., Ichoku, C., Chin, M., Bian, H., Darmenov, A., Colarco, P., Ellison, L., Kucsera, T., da Silva, A., Wang, J., Oda, T., and Cui, G.: Six global biomass burning emission datasets: intercomparison and application in one global aerosol model, Atmos. Chem. Phys., 20, 969–994, https://doi.org/10.5194/acp-20-969-2020, 2020.

Lastly, many conclusions are vague (e.g. l.338, 355). Although they are mostly supported by figures and tables I would have appreciated a few summary numbers.

Global inventory and intercomparisons:

Recent intercomparison exercises should be mentioned and discussed (see previous comments).

The authors find the highest agreement in trends with GFAS which may not be surprising since both approaches rely on the same FRE database. These are not independent estimates. Since the vegetation is assumed constant, the trends obtained reflects the trends in FRE.

To better understand the differences obtained for some regions in total emissions (e.g. BONA, SHAF, EQAS), it would be important to compare the values of EC used in GFAS with the values used in this study (see previous comment).

Minor comments.

Throughout the manuscript: check remaining typos, extra spaces in front of points, etc. I think it would be more accurate to talk about "emission inventory" rather than "emission budget"?

Abstract:
l. 4: remove 'More recently' since first emissions inversion were performed before the approaches based on FRP…
Main conclusions should provide some key numbers, as well as uncertainties.

Introduction:
l. 54, 75: laboratory measurements and field experiments.

Section 2.5 should introduce all emission inventories used in the manuscript.
What is the horizontal resolution of GFFEPS?

Section 3: 2nd sentence should be revised.