

REVIEW egusphere-2025-858 (2025-05-16)

The manuscript titled “Quantifying CO emissions from boreal wildfires by assimilating TROPOMI and TCCON observations” by Voshtani et al. with reference egusphere-2025-858 is a highly valuable scientific contribution in atmospheric air quality modeling and emission inversion for wildfires. Overall, it is very well written and provides a complete investigation of the benefit of large-scale CO emission inversion using TCCON total column retrievals. The study utilizes the CHEEREIO data assimilation toolkit for CO emission inversion and considers different wildfire emission inventories during the period between May and September 2023. In the first part, the benefit of assimilating TCCON data in addition to TROPOMI retrievals is analyzed on a global scale with 3 different global wildfire emission inventories. The results are compared with the assimilated, as well as different types of independent observations. The second part focuses on North America where the regional CFFEPS emissions are also included in the comparison. This part also includes a comparison and evaluation of different TCCON retrievals of CO.

Overall, I suggest accepting the manuscript with the following minor corrections:

General comments:

- A) My first general comment is on the length of the manuscript and its broad range of investigations. This is not a criticism, but rather reflects the extensive information content in this study. But still, I would at least suggest dividing the results in Sec.4 into two sections including the global and North America evaluation, respectively. You may even consider dividing the paper into 2 parts because especially the evaluation of different TCCON retrievals in Sec.4.5 appears to have a quite different purpose than Sec.4.1-4.3 and thus addresses a different audience (which can also be seen e.g. in the summary). This would also allow for a more detailed description and evaluation of the different TCCON retrieval data in Sec.4.5 (see related smaller comments below).
- B) I would also highly appreciate a short introductory paragraph at the beginning of Sec.2 and 4 to guide the reader and provide an overview of the content of each subsection.
 - For Sec.2: It would be helpful to already provide the information which observations are actually used for assimilation and validation in the section overview. Otherwise the amount of different observations might be confusing. This could maybe also be included in the title of the subsections. It is also not very clear why the aircraft data are presented in Sec.2.4 while the other in situ data are in Sec.2.3.
 - For Sec.4: Please also describe that the global evaluation in Sec.4.1-4.3 includes the 3 global emission inventories, while the regional evaluation in Sec.4.4 and 4.5 additionally includes CFFEPS (and see general comment A above)
- C) Thought the results, there are several locations where the word “uncertainty” or “error” is used without a clear specification on the actual measure. Since different measures are used in the evaluation, it is important so use specific wording.
 - 1.53: how is the uncertainty quantified? Eg ensemble spread or difference to independent data?
 - 1.358: How is the "a posteriori" uncertainty estimated? from LETKF analysis scheme? According to the table description for Tab.3, it is the a posteriori ensemble standard deviation. Please specify in the text.
 - 1.400: (as above?): Please define in the text how the uncertainties are estimated, ie as posterior ensemble spread after the assimilation as stated in the description of Fig.3? This is particularly important to interpret the large reduction of uncertainty when additionally assimilating TCCON compared to a small change in the posterior emission estimates themselves (on a global scale). Please also define how the global and regional

uncertainty estimated are calculated from the spatially resolved ensemble spread (or equivalent).

- 1.586: The word "uncertainties" is unspecific in this case. Are you referring to correlations, biases? Or to the investigation of uncertainties as given by the ensemble spread in Sec.4.2? Please formulate more specific.
- 1.664, 1.666, 1.667: please specify "uncertainty" in terms of the quantities that are given in Fig.11 or state where this statement is coming from elsewhere.
-

D) From my knowledge figure descriptions like color-coding should only be given in the figure caption because it's hindering the reading flow in the text (eg. 1.560ff, 1.606ff, 1.630ff, 1.658ff, 1.692ff, 1.717ff, ..., please check also other locations, especially before)

Smaller comments:

1. Abstract: Wildfires play an important role on a wide range of scales, from very local impacts on individual hotspots to global scale. Please state more clearly in the abstract that this study focuses on the large-scale.
2. Sec.2.4: can you please explain a bit more in detail how the aircraft data was sampled? Are the sensors mounted on commercial aircraft or was there a measurement campaign?
3. Sec.3: Please give a bit more details on the ensemble generation:
 - Are the emissions from all sources perturbed or BB-only?
 - How was the width of lognormal-distribution (ie perturbation variance) for the emission perturbations determined?
4. Sec.4.1, 1.386-388: This suggests that you are including all global emissions in the comparison. Since your assimilated observations were limited to below 60°N (&S), excluding this region from the comparison would provide more isolated insights on the actual impact of assimilation. Did you look at this as well?
5. Sec.4.1: 1.412f: I don't understand the importance of this sentence in the given context. Why should the a priori emissions and error estimates be different between the experiments?
6. Sec.4.1, 1.427-429: Is this "improvement" w.r.t. to the posterior? Please state clearly right when mentioning the 1st time in L.427.
7. Sec.4.1, 1.455-447: You mentioned the impact of TCCON already before (eg L.443). If the previous comparison refers to a priori vs joint a posteriori, please explicitly state so. From L.431ff, I was assuming you were comparing a priori vs TROPOMI a posteriori above. If this is the case, then I don't understand the explanation with few TCCON stations in L.443ff. Please clarify.
8. Sec.4.1, 1.455ff: That's a reasonable conclusion, is it possible to verify that, e.g. with the OSSE where you only assimilate TCCON in Apx.B?
9. Sec.4.2: For understanding, I would suggest adding a note that the normalized error reduction (and maybe also the DOFS?) increases with increasing weight of observations compared to the model a priori, even for overfitting to the observations, and that a small error reduction could also be due to a good model a priori.
10. Sec.4.3: Does the validation use the same TCCON data as used in joint inversion? If so (as stated in L.539f) state explicitly right from the beginning (e.g. 1.528) because it is an important information that the validation data is not independent for the joint inversion.
11. Sec.4.4: The title is rather long which makes it difficult to extract the specific content if this subsection: North America. Please shorten the title to stress its specific purpose more clearly. (and see also general comments above)
12. Sec.4.4: Which timer period is used in the statistical evaluation in Fig.11? The whole period or some of the episodes? Please explain in the figure caption and the main text.
13. Sec.4.5:

- Is the InSb retrieval really independent from the original retrieval, even when from same instrument? Without deeper knowledge about the retrievals: If its about different averaging kernels, it's probably more an "valuable" additional information rather than an completely independent measurement in a mathematical sense? Please explain.
 - Could you give some details about the combined TCCON observations? Were the two products combined, and how, before using in the assimilation? Or were both used jointly in the assimilation? If so, how are where treated, e.g. are they assumed to be 2 completely uncorrelated datasets? And is this a reasonable assumption, and why?
14. Sec.4.5, l.701ff: To me, the differences appear to be quite small. I would suggest stressing this a bit more, e.g. by adding this information in the last sentence in l.712ff.
15. Sec.4.5, l.718ff: That's an interesting point. Did you look into how the different averaging kernels impact the a posteriori emissions? Do they induce in different spatial or temporal distributions of emissions?

Technical and formulation-related comments:

- l. 216: “aircraft directly samples the atmosphere” Is there a word missing? Do you mean aircraft measurements (then “sample” in plural)? The aircraft itself does not sample.
- l.226 & 228: The word "driven" is used in two different contexts here.
 1. I'm not aware of using word for assimilated observations in a model. Suggest replacing here.
 2. Probably suitable word when referring to reanalysis, however formulation remains unclear. Is MERRA-2 used for initialization? Suggest more specific formulation.
- l.230-232: Please add a specific literature reference of the model description here.
- l.293, Eq.(1): tilde over analysis error covariance matrix was not defined
- l.301ff:
 - Please specify more clearly which variables are referring to emissions (ie in model space), and which to concentrations (ie in observation space), eg for apriori profile, simulated CO profile, ...
 - the definition of M remains unclear: does it refer to the forecast model or emission-to-concentration mapping, or both? If it is mapped in observation space as denoted, shouldn't it include the operator H? Please clarify.
- l.306: “, and the observation and the prior error covariances”: It unclear where is this part of the sentence referring to? Are the observation and error covariances also scaled based on the obs increments? Or are they also passed on the LETKF processor? Please specify.
- l.337: Spelling error? Do you mean “updated by”?
- Fig.4: The subfigures are quite small, it's especially hard to see influences of TCCON assimilation with a low spatial coverage
- l.481f: “more reliable” compared to what? The a priori?
- l.588: “difference ”unclear: Do you refer to differences between TROPOMI-only and joint inversion or between the different prior emissions?
- l. 593: “underestimation” w.r.t. what? Please specify and refer to section above if applicable
- Fig.9: The in-figure legend is confusing: Is "ensemble mean" the a posteriori ensemble mean, and "model a priori" the a priori ensemble mean? From my understanding, both are ensemble means, right? Then it would be clearer to label "a priori" and "a posteriori" (for each emission model) in the legend and the information "ensemble mean" can be added in the figure caption.
- l.622: “even though the” What is the message of this part of the sentence? You already compare with the a priori of GFAS and CFFEPS before. So they are closer to TROPOMI compared to what? Maybe the a priori of QFED and GBBEPx? Please reformulate and make the specific message of the sentence more clear.
- Fig.12: in-figure labels are very small, hardly readable