Paper title

Response to Anonymous Referee #1

Joshua L. Laughner on behalf of all coauthors

January 10, 2023

We thank the reviewers for their time in evaluating our manuscript, especially given its length, and are excited by the positive responses that the paper received. Below we respond to the individual comments. The reviewer's comments will be shown in red, our response in blue, and changes made to the paper are shown in black block quotes. Unless otherwise indicated, page and line numbers correspond to the original paper. Figures, tables, or equations referenced as "Rn" are numbered within this response; if these are used in the changes to the paper, they will be replaced with the proper number in the final paper. Figures, tables, and equations numbered normally refer to the numbers in the original discussion paper.

The comparison result from CO in Section 5.1 is a little bit disappointing, but not surprising given that CO is very variable.... I would like to see two extra comparisons that can be done quickly. One is the improvements of GGG2020 compared to GGG2014 after removing Armstrong and Lamont sites; The other is the comparison between GEOS-5 CO with other observation-assimilated CO simulations (for example, the CAMS global atmospheric composition forecast model, link below) over these two sites to see if the CO overestimate/underestimate is heavily model-dependent.

We have added these as Figs. S15 and S16 in the revised supplement. These comparisons show that the CO profiles are reasonable outside of urban areas or other locations with substantial fossil fuel combustion. These implications are discussed in two paragraphs added to Sect. 5.1:

> "Outside of urban or energy-intensive locations, the agreement between the new GGG2020 priors and co-located in situ profiles is much improved. Figure S15 compares RMSEs and mean prior vs. in situ differences for CO when Armstrong AFB, Lamont, and Orléans (another near-urban location) are excluded from the comparison. In that case, the RMSE reduces by about a factor of two or better at all levels except the surface in the new GGG2020 priors compared to the GGG2014 priors."

> "We compared CO profiles from the GEOS FP-IT product to the Copernicus Atmospheric Monitoring Service (CAMS) model to see if this issue of overestimated CO is common among models. The results for 2018 through 2022 are shown in Fig. S16. In general, GEOS FP-IT CO is dramatically greater than CAMS CO

in Los Angeles (at the Pasadena TCCON site). This is also true at Armstrong AFB, but to a lesser extent. In Paris, both models exhibit very high surface CO on some of the sampled days, though this was more frequent in the GEOS FP-IT CO profiles. At Lamont and East Trout Lake, both models had CO DMFs of similar magnitude (even with our factor of 1.23 scaling applied to the GEOS FP-IT data), with the main difference in vertical distribution. While the factor of 1.23 applied to bring the GEOS FP-IT CO in line with ATom observations (Fig. 12) definitely aggravates the GEOS FP-IT overestimate in urban areas, it improves the mean CO in more remote areas. In the future, drawing CO profiles from a model that better represents urban-rural CO gradients would improve the CO priors, but requires an existing model run that also covers the full range of times needed by TCCON (from 2004 on)."

Line 44: The description of the "1%" error in the shape of CO2 is a little bit ambiguous to me, and can be more specific, although you already have supplementary materials to explain that. One or two more sentences to explain that in the main text may help. Also, a change in the lower troposphere by 4ppm is not the same as changing the XCO2 by 1%. Similarly, the different scenarios in Figure S1 may represent different changes in XCO2, which may be confusing when compared to the retrieval error of ¡0.025% which is for XCO2.

We have edited this paragraph to more explicitly define what a 1% error in shape means in this context by specifying that it means the error in the prior relative to some truth changes from one level to the next:

> "The relationship between the shape error in the prior and the error in the retrieved column amount depends on the averaging kernels. For TCCON CO_2 retrievals, testing with synthetic spectra shows that a **4 ppm error in the profile shape (defined as the error in the prior compared to the true profile changing by ±4 ppm between the top and bottom levels)** leads to an error of $\leq 0.025\%$ in XCO_2 at solar zenith angles (SZAs) $\lesssim 60°$, and $\leq 0.125\%$ up to SZA $\approx 75°$. (Details of how this was quantified are given in Sect. S1.) This means that for typical SZAs observed by TCCON, an error of **about 4 to 8 ppm** in the CO_2 prior results in a retrieval error well below the 0.25% ceiling required for TCCON data."

As for the different scenarios in Fig. S1 representing different XCO_2 values, that was intentional. The results shown in Fig. S1 are intended to test the retrieval sensitivity to errors in both shape and total CO_2 column. We have made this explicit in the supplement:

> "Figure S1 shows the different prior profiles (panel a) and the resulting change in retrieved XCO_2 compared to the true profile (panel b). We defined two types of shape error: a "jump" where the CO_2 DMF increases or decreases suddenly at a specific altitude, and a "linear" error where the CO_2 DMF varies linearly with respect to pressure. For all shape errors, we defined a 1% error to mean that the DMF changes by 1% (4 ppm) between the top and bottom of the profile. Both the "jump" and "linear" cases each have three subcases that vary whether

the troposphere, stratosphere, or both have the error. **These various profiles represent different errors in both the shape *and* prior XCO_2 values. This was deliberate to test how the retrieval is sensitive to not only the error in shape but the total amount of prior CO_2.**"

Line 6: "improving the description of CO2, CH4, N2O, HF, and CO in the stratosphere", please rephrase.

Changed to:

"A particular focus of this work is improving the **accuracy of CO_2, CH_4, N_2O, HF, and CO across the tropopause and into the lower stratosphere.**"

Figure 12: Please add a unit (or is it just fraction/ratio?) for the x-axis of the lower panel.

This is the unitless relative difference between GEOS and ACE CO, we have clarified this in the caption:

"Comparison of colocated ACE-FTS and GEOS-5 FP-IT CO data. **The x-axis is the unitless relative difference, (GEOS - ACE)/ACE.** The y-axis..."

Figure 12: In the caption: The fit is a robust fit using a Tukey biweight function with no intercept. Please explain "Tukey biweight function". Also "intecept" is a typo.

Clarified as:

"The fit is a robust fit using a Tukey biweight function with no intercept, **i.e. using the `RLM` linear model with `M = TukeyBiweight()` from the Python `statsmodels` package (Seabold and Perktold, 2010).**"

References

Seabold, S. and Perktold, J.: statsmodels: Econometric and statistical modeling with python, in: 9th Python in Science Conference, 2010.