Review of the manuscript

On the estimation of stratospheric age of air from correlations of multiple trace gases

submitted to EGUsphere by F. Voet et al.

S. Chabrillat, BIRA-IASB, November 2024

General Comments

This study is an exhaustive and potentially very useful contribution on the estimation of stratospheric Age of Air (AoA) from coincident retrievals of several long-lived tracers, using the compact correlations which can be established in model space between AoA and such tracers. It convincingly shows how these new AoA derivations would have significantly reduced uncertainties compared with the standard derivations which rely only on SF_6 measurements. It relies on an advanced chemistry-transport model which is well suited to the task, as demonstrated by many earlier papers on AoA and the Brewer-Dobson Circulation (BDC). The methodology is explained in detail, even though its terminology deserves some additional attention. The estimation of uncertainties has been carefully considered, which greatly enhances the interest of this work. The method is applied to both synthetic and actual balloon retrievals, allowing in-depth discussions of the merits of this approach - even though they sometimes feel hastily written and should have been reviewed for consistency.

Overall, I wholeheartedly recommend publication after minor revisions addressing the comments below (especially MC1). I have one additional concern though: the availability of the data. The whole paper is very focused on the creation and usefulness of Look-Up Tables (LUT) between the AoA and tracers, including the associated uncertainties. These tables were derived from CLaMS simulations which have been run from 1979 until 2022, and their applicability is demonstrated from 2011 onwards. Since they seem easy to apply to several existing remote-sensing (and maybe even in-situ) observational datasets, they could be of great interest to the whole community studying the BDC - allowing new comparisons and advancing the whole field. But this cannot be done with the submitted manuscript, because the correlations are shown in the supplement in a graphical format (and only for 4 days of 2011) while they are not published in a numerical format.

I believe that these tables should be publicly downloadable, either in their raw form (mixing ratio bins versus mean and 1-sigma of AoA) or as polynomial fits (coefficients and polynomial expressions). This requires only 12 tables per day, and could be done either as a zip-file supplement or by uploading the dataset to a data repository and documenting the resulting DOI. If for some reason the tables are difficult to derive on other years than 2011 (and August 2021, 2022), that single year would still be useful. If only the 4 specific dates chosen for in-depth study can be made available, that could still allow application to several dozens of ACE-FTS profiles. If none of this can be done, then the final statement on Code and data availability should be corrected, specifying that the CLamS model data **and derived LUT discussed in this paper** may be requested by e-mail to the authors.

Major comments

MC1. What is exactly the upper limit of the validity of this method, and why? No results are shown above 25 km altitude. I understand that AoA-tracer correlations break down somewhere above this limit, but it would be useful to show where and how exactly. More specifically: from Fig. 5 and A4, I would have expected a a sharp increase in the uncertainty at some level around 25 km in the polar regions of Figure 6 (or maybe Figure 8 or 9?). Does it happen higher up? If yes, it would be good to extend upwards at least that figure in order to highlight this fact. If no, this would indicate that something is missing in the estimation of uncertainties as they should reflect the loss of correlation (i.e. usable information) above some altitude.

On the same topic, I am worried by Figure 10: below 21 km there are large differences between the AoA by CLaMS and by "GLORIA new method". This is good, showing that the new method preserves the information contained in the GLORIA retrievals. But above that limit, the dashed and greenish profiles converge with the solid magenta profile and all three reach a suspiciously good agreement. This is actually mentioned in the discussion (lines 435-437):

> "*...the correlations used in the proposed new age calculation method are based on model simulations and not necessarily reflect the actual atmospheric conditions. This could also, to some degree, cause the agreement between the "GLORIA-B new method" AoA and the "CLaMS clock tracer" AoA at the upper end of the height scale.*"

Could it rather be due to the loss of correlation between AoA and tracers for AoA>4 years (e.g. Fig. 3-4), leading to the "GLORIA new profile" losing any information from the 6 retrievals and simply reflecting instead the model-based correlations for such "old" AoA? That would be unexpected from the in-depth discussions about uncertainties (Fig. 6-9). I believe that to elucidate this point, it would help to plot figures similar to fig. 1-9 but for the two dates of the GLORIA soundings.

MC2. The Upper Boundary Condition (UBC) for SF_6 plays an important role in this study where it is located at the stratopause (55km). This is explained (lines 145-147) as

> *"The upper boundary condition for times outside of the measurement period was created by parameterizing the depicted seasonal cycle of each latitude with a sinusoidal least square fit and adding it to a shifted tropospheric tropical time series (taken from the lower boundary of SF6)."*

If that parameterization is as simple as described, I have some concerns. The Semi-Annual Oscillation already plays an important role at the stratopause, and it is modulated by the Quasi Biennal Oscillation (see e.g. Garcia et al., 1994). Or maybe that your sinusoidal least square fit includes several frequencies? Please clarify, and insert a reference if this fit was described in more detail in earlier work.

Could you also be more specific with the UBC for HCFC-22 ? Line 155 states

> "*An open upper boundary condition has therefore been defined for HCFC-22.*"

What does this mean? In my favorite (Eulerian) CTM, anything can happen at the upper boundary if no UBC is specified. Maybe a zero vertical gradient (i.e. null flux) is specified?

Tthese concerns could also be addressed by showing the timeseries of the (modelled and observed) SF_6 as well as the (modelled) HCFC-22 at the uppermost level.

MC3. What are the limitations of deriving LUT with only one model driven by one reanalysis dataset? Lines 443-444 state

> "*It should be noted that this slow bias of the ERA5 circulation in the comparison presented here is independent of the age calculation method used*".

…and I agree: a biased circulation in the driving reanalysis should impact the long-lived tracers and the clock tracer in the same manner, thus delivering the same correlations as a "perfect" reanalysis. On the other hand, the previous paragraph states (lines 435-436):

> "*...the correlations used in the proposed new age calculation method are based on model simulations and not necessarily reflect the actual atmospheric conditions.*"

…and I also agree with this, but the two sentences seem to contradict each other. Hence I recommend to discuss these "existential questions" a little further in Section 4.

Specific Comments

Original text is copied in *italics*, suggestions for corrections are typed in **bold**.

SC1. Terminology: some recurring words lead to confusion and should be replaced throughout the text:

- If I understand well, the midpoints of the AoA/vmr bins are interpolated with a polynomial fit which uses the Savitzky-Golay-Filter to create smoother LUT. The result is variously described as "*interpolated series*", "*interpolation*", "*smoothed series*", "*polynomial fit*" or "*Savitzky-Golay-Filter*". Please use a consistent term throughout the text and in the figures ("polynomial fit" is fine) and provide a reference on the Savitzky-Golay-Filter or its usage for this type of procedure (line 206).
- The CLaMS AoA derived from the lag-time of the clock tracer is written very often as *the "true" AoA* or *the* "*exact*" *AoA*. Even with the quotes, this is misleading since this AoA is biased by ERA-5… In most cases no adjectives are necessary as it is perfectly clear which AoA you are using as reference. But if an adjective is necessary, I recommend to write systematically "*the* ***actual*** *AoA from the clock-tracer*".
- From lines 211 to 308 there are 5 attributions of AoA spread in a mixing ratio bin to "*natural variability*". This is misleading for the same reason: this variability arises from the model simulation and its driving reanalysis and is necessarily natural… Hence I recommend to replace "*natural variability*" by "***model*** *variability*" or "***model/reanalysis*** *variability*".
- "*one sigma environment(s)*" → "**one-sigma range(s)**"

SC2. Abstract, line 6:

"*this method works well* **in most of the lower stratosphere** *up to a height of about 25 km*"

(in order to account for the white areas in Fig. 6 to 9).

SC3. Abstract, line 8-9:

"*The* **multi-species** *age calculation method is evaluated in a model environment and compared against the* ~~*true*~~ ***actual*** *model age* ***from an idealized clock tracer***."

SC4. The introduction is well written, but please cite some previous papers using compact correlations between AoA and long-lived tracers to derive an estimation of the AoA, and explain how their aims differ from those in your study. To the best of my knowledge, the first such paper was written by **Linz et al. (2017)** and the latest one was written by **Dubé et al. (2024)**. Even though the latter is still a preprint, it is important to cite it because it also uses CLaMS.

SC5. Section 2.1: please provide some details about modelling of vertical circulation in CLaMS. I guess that these are radiative heating rates between the isentropes, but do they come from ERA-5 or from an internal part of CLaMS? This is important as in each case there could be biases in these heating rates (which could be mentioned later in the discussion of the results).

SC6. Legend of Figure 2: the details of this Figure are already explained in its lower text box and in the text, so I believe that it is not necessary to repeat them for a third time in this Legend (also because all terminology issues arise there - see SC1). A shorter legend could be:

> "***Figure 2.*** *Schematic representation of method used to estimate the* **AoA corresponding to measured mixing ratios and its associated** *uncertainty* **(see text for details)**."

SC7. Line 219: " *...added together through means of (Gaussian) error propagation...*"
→ do you mean that they are added as a Root Mean Square Error (RMSE) ?

SC8. Line 268, first sentence introducing Fig 3: please clarify here that you are not yet showing the binning hemispheric approach described in the earlier section 2.2.

SC9. Line 270:
"*Additionally, the equivalent latitudes of the model points* **(Nash et al., 1996)** *are color-coded…*"

SC10. Lines 312-317: this procedure is written in a rather obscure manner, even though it seems quite simple. Please re-formulate, e.g. explaining that prior to the application of these LUT you want to compare the 6 total uncertainties for each species among each other and also with the total combined uncertainty.

SC11. Line 337: "*(compare* **with** *Fig. A4)*".
This figure A4 is compared multiple times with Fig. 5, 6, 7. I think that it should be moved from the Appendix to the main text.

SC12. Lines 355-356:

> "*The pseudo-measurements were created by adding normally distributed random noise to the mixing ratios of the six trace gases for all air parcels on the four considered days.*"

These synthetic measurements are not completely realistic as they do not take systematic errors into account. This is a real concern and should be mentioned somehow, because actual measurements would be biased w.r.t. modelled mixing rations since there are biases in the ERA-5 circulation.

SC13. Line 362: " *The thick black lines represent the zonally averaged tropopause from the ERA5 reanalysis data for each day*". Is this the thermal (T gradient) or dynamical (PV+theta) tropopause?

SC14. Lines 365-367: "*Even the northward intrusion of air with AoA below one and a half years into the layer of air with AoA between one and a half and two years at roughly 70°N and 14 km height in July is clearly present in the results of the correlation method (compare Figs. A4 (c) and 7 (c))*"
This region is not obvious to identify at first. Consider helping the reader by drawing a bounding box around it, at least for these two figures.

SC15. Lines 385-386: you mention Figure 8 but shouldn't that be Fig. 9 there ?

SC16. Lines 385-388:

> "*(Note that the scales of the colorbars are different and that the contours show different things in the two figures)*"
>
> This is the straw that broke the camel's back. I was wondering since Fig. 8 about the added value in showing absolute differences with color coding and relative differences with contour labels, which is quite confusing and was not done for fig. 6. Note that fig. 6 also could have shown absolute errors in color shading and relative errors with contour labels, but I am not complaining that this was not done. On the contrary, I really think that the color scales and/or contour labels should be changed in Fig. 8 and 9 to simplify them and allow direct comparison with fig. 6 !

SC17. Lines 395-401: please synthesize in a few sentences

> "*...the standard convolution method, as described in Garny et al. [under review], and the subsequent correction for SF6-depleted air from the mesosphere introduced by Garny et al. (2024).*"
> …including the limitations of this method and this correction. This is important because these limitations play a role in the discussion of Fig. 10 while the review by Garny et al. is still being reviewed.

SC18. Figure 10: the dashed "*greenish*" lines and the greenish shadings are not really visible, especially on top of the blue shadings. Consider using black lines and black horizontal error bars instead (possibly for a subset of the levels in the case of horizontal error bars).

SC19. Lines 430-432: " *Another possibl~~y~~***e** *explanation ~~could~~* **would** *be ~~that~~ some issue with the instrument during the Timmins flight lead***ing** *to the retrieval of systematically too low SF6 mixing ratios. Perhaps the unintended descen~~d~~***t** *of the gondola down to 22 km mentioned in sect. 2.3 could have caused such an issue.*"
→ … while not damaging the retrievals of the 5 other species, leading to the very different values by "GLORIA new method" ??

SC20. Line 462: "*The lookup tables for the remaining cases can be found in Figs. S17 to S28 of the supplement to this article*" → these are not LUT, they are plots of the LUT (see MC1).

SC21. Figure 11(f): "*...by means of gaussian error propagation)***.** *Contours: Relative difference between color-coded values and AoA from clock tracer.*"
This is very unclear… What do these contour lines mean exactly? Are they necessary? They are not discussed in the text…

SC22. Section 4.2: this discussion on the stability of the correlations is quite confusing:
"*Since these depletion processes do not fundamentally change over time, the future correlations of the five mentioned trace gases with AoA will likely be similar to the way they were in 2011, the year considered in this study*".
But GLORIA-B used LUT made specifically for its flight dates, so this question about the stability of the correlations is moot: the model correlations can be computed for the days of the measurements, i.e. in much more recent years tha 2011, and this was actually done! It looks to me like this section was written before the comparison with GLORIA-B and not removed afterwards.

SC23. Line 574: " *Also, such datasets could be used to study exchange processes between the troposphere and the stratosphere and therefore help to better constrain new emissions of prohibited substances like CFC's*".
This sentence is not clear. There are many other good reasons to sudy troposphere-stratosphere exchanges.

Typos, wording etc.

- Line 18: "*... Brewer-Dobson circulation (***BDC; see e.g.** *Holton et al., 1995; Butchart, 2014~~; BDC, see e.g.~~*)
- Line 78: "*... if the***y** *can be retrieved…*"

- Lines 198-205:
  "*~~For simplicity sake~~* **For visual clarity**, *only the last three of these mixing ratio bins are shown in the figure. The histogram illustrates the distribution of AoA within a given mixing ratio bin. The blue area in the histogram highlights the one sigma environment around the mean AoA value of the distribution (mean value ± one ~~sigma~~* **standard deviation***). Similarly, the blue area in each of the three depicted mixing ratio bins corresponds to the one-sigma ~~environment~~* **range** *around the mean value of the respective AoA distribution inside. Such one-sigma ~~environments~~* **ranges** *are calculated for each one of the one hundred fifty mixing ratio bins. Subsequently, a midpoint for each bin with the sample mean AoA as the x- and the middle of the bin range as the y-coordinate was then defined. Th***is~~e~~** *~~thus~~ constructed set of midpoints constitutes a sort of look-up table that can be used to interpolate a***n** *AoA value…*"
- Line 216: first sentence actually belongs to previous paragraph.
- Lines 372: "*The absolute difference ~~seems to~~ reach***es** *its highest values…*"
- Lines 381-382: "*These standard deviations are a quantification of the spread of the AoA difference in zonal direction. They can be used to estimate the uncertainty of AoA derived from individual measurements at different ~~latitudes~~* **longitudes** *for any ~~longitude~~* **latitude.**"

- Line 401: first sentence actually belongs to previous paragraph.
- Line 410: "*The magenta shading around the "GLORIA-B new method* ~~*method*~~*" AoA represents…*"
- Line 425: "*...however, any conclusions drawn from the* ~~*actual*~~ **corresponding** *values are hardly meaningful*"
- Lines 451-472: please replace all occurrences of "3d" with "3-D".

Additional bibliographical references

Garcia, R. R., T. J. Dunkerton, R. S. Lieberman, and R. A. Vincent (1997), Climatology of the semiannual oscillation of the tropical middle atmosphere, *J. Geophys. Res.*, 102(D22), 26019–26032, doi:10.1029/97JD00207.

Nash, E. R., P. A. Newman, J. E. Rosenfield, and M. R. Schoeberl (1996), An objective determination of the polar vortex using Ertel's potential vorticity, *J. Geophys. Res.*, 101(D5), 9471–9478, doi:10.1029/96JD00066.

Linz, M., Plumb, R. A., Gerber, E. P., Haenel, F. J., Stiller, G., Kinnison, D. E., Ming, A., and Neu, J. L.: The strength of the meridional overturning circulation of the stratosphere, Nature Geoscience, 10, 663–667, https://doi.org/https://doi.org/10.1038/ngeo3013, 2017.

Dube, K., Tegtmeier, S., Ploeger, F., and Walker, K. A.: Hemispheric asymmetry in recent stratospheric age of air changes, EGUsphere [preprint], https://doi.org/10.5194/egusphere-2024-1736, 2024.