Response to Reviewer #2 for discussion paper

Comparison of mean age of air in five reanalyses using the BASCOE transport model

Chabrillat et al., ACPD, 2018

We thank the reviewer for his/her insightful comments. It appears that the version of the manuscript which was reviewed by this referee is the version first submitted to ACPD (on 4 April 2018) rather than the version finally published in ACPD (on 7 May 2018). Fortunately all comments apply equally to both versions. In our replies below the bold type is used to highlight text in the revised manuscript.

Replies to general comments

- *The principal conclusion is that the simulations of AoA obtained when BASCOE is constrained by different reanalyzed datasets differ substantially from one another. This is not at all unexpected given the differences among the reanalysis models.*

The reanalysis systems are based on different models but they assimilate very similar satellite datasets. Many users of reanalyses are neither modellers of stratospheric dynamics nor aware of the lack of observational information to constrain the BDC in the reanalyses. From feedback obtained at the 5th International Conference on Reanalysis (ICR5, Rome, November 2017), such users do not expect to see a spread between the reanalyses which is as large as between unconstrained GCCMs (Fig. 4). On Fig. 8 they easily understand that the uncertainties in the observational timeseries are large (due to sparse and irregular sampling) but they do not expect to see that the spread between the reanalyses is as large as these observational uncertainties.

A third highlight of this paper is the intercomparison of AoA trends between the reanalyses. Several reanalysis intercomparisons of diagnostics related to stratospheric dynamics have already been published and showed significant differences with respect to their trends (e.g. Abalos et al., 2015; Miyazaki et al., 2016). Yet for the AoA diagnostic, most recent studies rely on ERA-I with much interest in the latitudinal structure of its trends. We found that over the post-2002 period ERA-I is the *only* reanalysis to deliver opposite trends of AoA in the two hemispheres (Fig. 12, middle column). This is also an unexpected result.

- *The paper is well organized and clearly written, with some exceptions, the main one being that the procedure for computing AoA is not well explained. In particular, it is not clear whether AoA is calculated with respect to a reference level at the tropical tropopause or in the troposphere, and this introduces some ambiguity in the interpretation of the results.*

We agree with the referee that the handling of the reference level was problematic in the submitted manuscript. All our calculations used the surface both as the source region and to compute the time lag defining the mean Age of Air. But in order to better highlight the different transit times from the equatorial tropopause, Fig.1 and 3 were corrected a posteriori by subtraction of the time-averaged AoA at 100hPa, 10°S-10°N. All other figures used the surface as reference, hence including the transit time from the surface to the tropical tropopause. This distinction was not clearly made and led to inconsistent figures, as shown by several specific comments made by both reviewers.

Hence we decided to re-run our calculations and re-plot all figures (except for figure 8, see below) using as source region the tropical tropopause region (still defined as the 100 hPa isobar between latitudes 10°S and 10°N), and computing the AoA at each grid point as the time elapsed since the mixing ratio of the ideal tracer reached the same value in that source region. The last paragraph of section 2.1 has been re-written to fully explain the updated procedure for computing AoA.

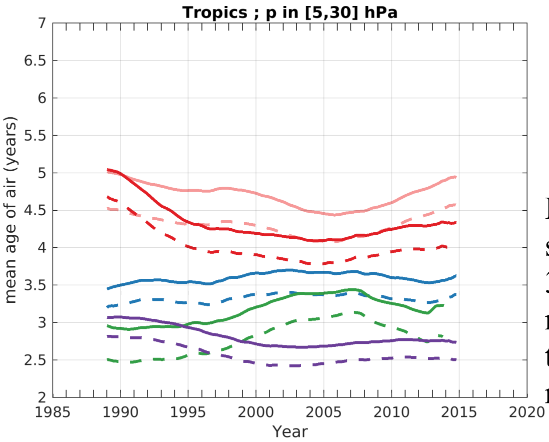
The figures did not change much from the discussion paper, indicating that this is a methodological issue which does not have a large impact on our findings. Besides figures 3 and 9 which are discussed below for specific comments, there is one other case where the figure changed sufficiently to warrant a minor update in the text: on figure 12 the positive AoA trends for ERA-I in 2002-2015 (top row, middle column) have become significant at all latitudes (in the discussion paper they were significant only in the polar latitudes). On figure 12 the signs and patterns of AoA trends did not change for any other reanalysis or period but the range of these trends increased by up to 50% (see min/max values above the plots); this led us to extend the scale of the color bar, from [-0.6,0.6] to [-0.9,0.9].

For figure 8 (and figure 8 only) we have kept the original calculations where the tracer was set to increase linearly throughout the surface, because this figure includes a comparison with observational values of AoA which used the surface as reference. We have moved to the discussion of figure 8 the description of this surface boundary condition and its propagation through the troposphere, because it is now irrelevant for all other figures. We have also added in this figure a plot showing tropical AoA computed both from the surface and from the tropical tropopause to show that the difference does not vary significantly with the simulated year (see next comment).

Replies to specific comments

- *(4, 20) "There is no other representation of convection": It is not clear that, in the Tropics, where deep convection can reach the 14-15 km level, this artificial diffusion can simulate vertical transport realistically. But perhaps this does not matter for assessments of AoA in the stratosphere if the base point for AoA calculations is taken to be at or near the tropical tropopause? Please comment (especially since it is not clear how the reference level for computing AoA is chosen). See also comment at (7, 24).*

As explained above, nearly all figures now use (100hPa, 10°S-10°N) as reference region hence the absence of deep convection in the CTM is not an issue for the updated figures. Since for figure 8 we keep using the surface as reference, we added tropical timeseries in the tropics which show both the surface-based evaluation (solid lines) and the tropopause-based evaluation (dashed lines):



Middle panel of revised Fig. 8. Mean AoA in the mid-stratosphere (5-30hPa) for the tropical latitudes (30°S–30°N). Solid lines show AoA using the surface as reference, dashed lines show AoA using the tropical tropopause as reference (i.e. as in all other figures of the revised manuscript).

This comparison between the two evaluations also allows a discussion on the impact of the omission of deep conviction in the model. The discussion of Fig. 8 in the revised manuscript includes the following paragraph:

> **These differences between the two calculations represent the transit times from the surface to the tropopause, are nearly independent of the simulated year and range between 3 months (with ERA-I or JRA-55) and 6 months (with MERRA). These values are close to the longest transit times reported in a recent intercomparison of global models (Krol et al., 2018), indicating a rather slow transport from the surface to the tropical tropopause which we attribute to the omission of deep convective transport in our model. Hence the model results in Fig. 8 may be slightly overestimated but these biases have no significant inter-annual variations and do not hinder the intercomparison between reanalyses.**

- *(6, 5) "at the wavelength number 47": "at wavenumber 47" might be better.*

Text corrected.

- *(6, 8) "Figure 1 compares the results": I do not believe you have stated how AoA is calculated. [...]*

Section 2.1 now describes explicitly the revised procedure to calculate AoA from the tropical tropopause region:

> **The age of air is defined as the spectrum of transit times from a source region to a given location, with the tropical tropopause usually defining the source region for studies of the stratosphere. In the case of ideal tracers which increase linearly in the source region and have no photochemical production or losses, the mean of this spectrum (denoted here AoA) is simply the time elapsed since the mixing ratio of this ideal tracer reached the same value in the source region (see e.g. Waugh and Hall, 2002). We follow here this classical approach, using for most simulations the 100 hPa isobar between latitudes 10°S and 10°N as source region.**

Section 3.2 describes explicitly the original procedure which has been kept only for figure 8:

> **For consistency the modeled AoA in this figure are evaluated as the time elapsed since the mixing ratio of an ideal tracer reached the same value at the surface, using as boundary condition a global constant increasing linearly with time at the surface.**

- *(7, 18) Figure 2: This figure, as well as Figure 3, would benefit from a color bar to indicate the values of the AoA isolines not explicitly labeled.*

A color bar has been added to figures 2 and 3.

- *(7, 24) Figure 3: I am confused by this comparison [...] So, where is the reference point in these simulations? If it is at the surface, then AoA will reflect the effects of transport not just in the stratosphere, but also in the troposphere, including the artificial diffusive transport between the surface and the middle troposphere. Unless I am misunderstanding what you have done here, it seems to me that, if AoA is intended to highlight transport in the stratosphere (e.g., Waugh and Hall, 2002, Sec. 3.1) then the choice of a base point in the troposphere confuses the issue, especially given the use of artificial diffusive transport in the lower troposphere.*

The reviewer was rightly confused and his interpretation is correct. We have followed this advice, choosing the tropical tropopause as reference point in the revised manuscript (see above). The relative differences between ERA-I and the four other reanalyses vanish at the reference point and the difference is not plotted at grid points where ERA-I AoA is smaller than 5 days.

- *(8, 5) "The intercomparison at 50 hPa": You should state explicitly in the text that in these comparisons AoA is "normalized" to be zero at the tropical tropopause (this is only stated in the caption of Figure 4). Otherwise, the reader will wonder, as I did, why the AoA shown in Figures 2-3 are different from the AoA in Figure 4. By the way, a problem with the "normalization" of AoA to zero at 100 hPa is that it gives the impression that AoA above that level is determined only by the stratospheric circulation, when in fact the AoA also contains the effect of transport in the troposphere.*

thanks to the direct calculation of AoA using the tropical tropopause as reference point, no "normalization" is performed any more for the figures 1 and 4 of the revised manuscript.

- *(8, 12) "overall, the spread . . . is larger than the 1-sigma. . .": One wonders how this result would change if AoA were computed with respect to a reference point at 100 hPa.*

In the revised manuscript the AoA are computed with respect to a reference point at 100 hPa. The differences in Figure 4 between the submitted and revised manuscripts are nearly indistinguishable. Hence the spread between the five simulations at 50 hPa is still larger than the 1-σ observational uncertainties in the tropics, and still nearly as large in the extratropics. We have not modified this sentence in the revised manuscript.

- *(8, 26) "The spread between the four reanalyses reaches a maximum of 0.2 years at 30 hPa": Are you referring here to the gradient comparison, Figure 4d? How is this "gradient" calculated? The figure legend refers to "MLNH-Tropics" and shows values in months, not per unit distance, so this is really a difference between the Tropics andmidlatitudes o f the NH. How are Tropics and NH midlatitudes defined?*

The words "(mean age) gradient profiles" or "latitudinal gradients (of mean age)" were meant with the same meaning as Neu et al. (2010) and Chipperfield et al. (2014) i.e. as the difference between AoA in NH midlatitudes and AoA in the Tropics. The vertical profiles on figure 4d simply show the differences between the corresponding profiles on figures 4c and 4b which are mean values for latitude bands 35°N-45°N and 10°S-10°N respectively (as stated in the figure of caption 4).

In the revised manuscript we have added the definition of the latitude bands in the discussion of figures 4b and 4c and we have added the following sentences in the discussion of figure 4d:

> **These "latitudinal gradients of AoA" were used in several CCM intercomparisons (Neu et al., 2010; Chipperfield et al., 2014). Figure 4d shows this diagnostic for the five reanalyses, i.e. the differences between the AoA profiles on Fig. 4c and Fig. 4b.**

We have replaced the words "latitudinal gradients" by "AoA differences" in the remainder of this discussion and in the caption of the figure.

- *(8, 30) "MERRA-2 yields an outlying vertical profile of AoA at northern midlatitudes": True with respect to the other reanalyses except for MERRA (Fig. 4c), and in fact, MERRA and MERRA-2 midlatitude profiles of AoA agree best with the observations. You keep referring to MERRA-2 as an "outlier", which carries negative connotations,but in fact being an outlier in this comparison is a good thing if one considers the data to be the "truth"..*

Thanks for pointing this out. We have corrected the discussion according to your comment:

MERRA and MERRA-2 yield larger AoA at northern midlatitudes than the three other reanalyses. In the case of MERRA-2 this results in a profile of AoA differences which are significantly larger than the profiles obtained with the four other reanalyses but agrees much better with the profile derived from the observations. Hence MERRA-2 apparently underestimates the tropical upwelling in the lowermost stratosphere (100-60~hPa), agrees better with the observations at 50~hPa than any other simulation, and joins the results of the four other reanalyses at higher levels.

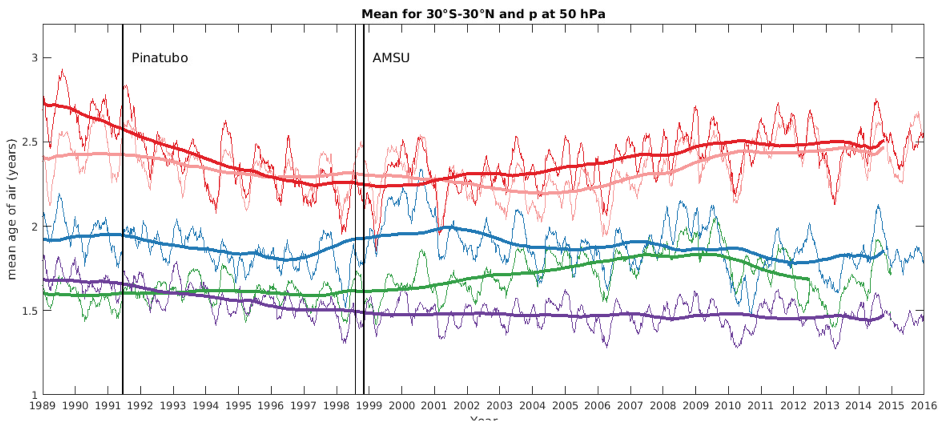
- *(9, 14) "MERRA-2 starts with much older values": This behavior does appear to be anomalous. Any idea what might be causing it?*

This issue is discussed in detail in sectin 5 (see paragraph starting with "MERRA-2"). We have inserted the following sentence in section 3.2:

The possible causes for this apparently anomalous behavior of MERRA-2 are discussed in section 5.

- *(9, 18) "The Pinatubo eruption does not appear to have any impact of the simulated AoA at 50 hPa": Insofar as one might expect that the largest impact of Pinatubo would be in the Tropics, it might be worthwhile to examine the AoA time series averaged over, say, 30N-30S.*

This was done and no impact was found for the tropical latitude band, as shown by the corresponding plot:



Note that any impact in the 30°S-30°N latitude band would have been seen in figure 7 which shows the 72°S-72°N band. The revised manuscript mentions the absence of volcanic impact in the tropical latitude band as well.

- *(10, 8) "observational trend is not significant": One would not expect any trend calculated from the smoothed, model time series shown on the right pane of Figure 8 to be significant either. By the way, you keep referring to "trends" in connection with the model results, but you have not calculated any trends...*

Not yet. As outlined in the introduction, modelled trends are evaluated extensively in section 4. The sentence now refers to this later finding:

> **ERA-I delivers a weakly positive trend over the period 1989-2015 and we will assess in section 4.3 that this trend in the model results is significant.**

And this finding has been added in the discussion of figure 12 (section 4.3):

> **The same plot** (*i.e. Fig. 12, top right)* **also shows that the positive trend which had been inferred visually for the northern mid-latitudes of the middle stratosphere (Fig. 8, left) is significant.**

- *...Note also that Garcia et al. (2011) have argued that, even using model output for an ideal AoA tracer, trends over periods as long as 30 years are often not significant when the ideal tracer is sampled like the available observations of stratospheric tracers....*

Great caution should indeed be exercised in comparisons of trends between model output and observational datasets which are sparsely and irregularly sampled. This is what we meant in the original manuscript with "..., but Engel et al. (2009, 2017) showed that the sign of this observational trend is not significant". Referring to Garcia et al. (2011) allows us to reinforce and clarify this call to caution. The revised manuscript states:

> **While the overall trend simulated with ERA-I is apparently in agreement with the balloon observations, this comparison should be considered with great caution because the sign of the AoA trend is not significant in the observations (Engel et al., 2009, 2017) and modelled trends over periods as long as 30 years are often not significant when the ideal tracer is sampled like the available observations of stratospheric tracers (Garcia et al., 2011).**

- *... Furthermore, trends derived from observation are also confounded by the fact that no real atmospheric tracer has a constant, linear growth rate.*

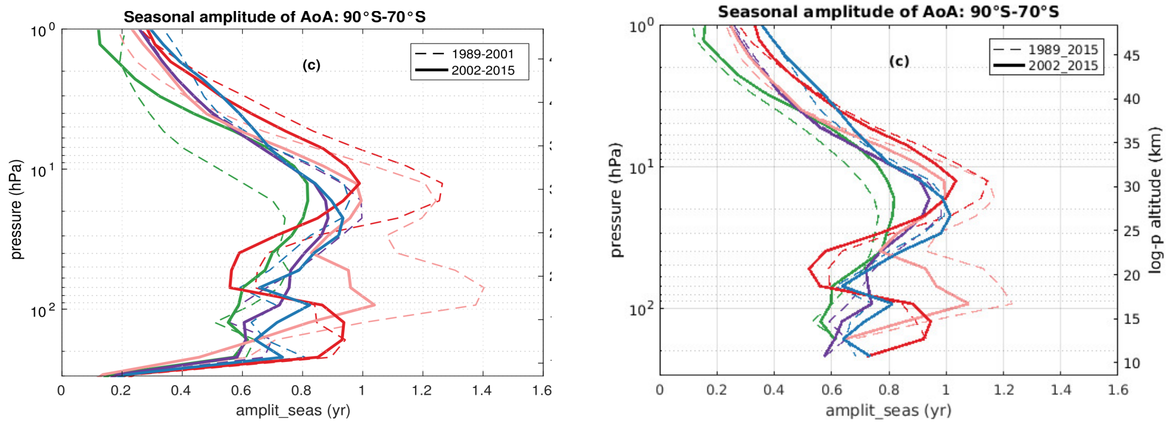
The non-linearity of the growth rate of CO_2 and SF_6 has of course been taken into account by Engel et al. (2009, 2017) for their derivation of AoA and also for their error analysis. The procedure is described in the supplementary information of Engel et al. (2009) and in the first paragraph of section 4 in Engel et al. (2017).

- *(11,16) "in the polar regions and midlatitudes": Why limit this to extratropical behavior only? It would be interesting to show the seasonal amplitude in the Tropics as well, say a ±30° average.*

Thanks for the suggestion. The seasonal amplitudes with ERA-I and MERRA-2 have a different vertical structure in the Tropics. We have added this plot to figure 9 in the revised manuscript.

- *(11, 24) "MERRA and MERRA-2 . . . different amplitudes depending on the period used or the analysis": Does this have to do with the development (1989-2001) and stabilization/decline (2002-2015) of the Antarctic ozone hole? To explore this issue, one would have to examine the actual seasonal climatology at high SH latitudes, not just the annual amplitude.*

The left plot below is extracted from fig. 9 in the ACPD manuscript while the right plot shows results with the new AoA calculation (i.e. using the tropical tropopause as reference):



This shows that the new AoA calculation delivers amplitudes of AoA seasonal variations which are much closer for periods 1989-2001 and 2002-2015. This is the case also for the other latitude bands, so for the revised manuscript we have removed from figure 9 the results for the early period 1989-2001 and we have dropped the corresponding part of the discussion.

- *(12, 9) "could not be expected from inspection of the native dynamic variables": I do not understand what you are trying to say here. Please elaborate. And note also that the discrepancies you mention ("up to 50% dependencies on the considered time period") are not even illustrated, so it is very difficult to even guess what the intent of your statement is.*

The dependencies of seasonal amplitudes on the considered time period reached 50% in fig. 9a and 9c of the ACPD manuscript but, as explained in the previous comment, they have disappeared from the revised manuscript thanks to the corrected AoA calculation. We removed the whole paragraph (i.e. last paragraph of section 4.2) from the revised manuscript because it is redundant with the second paragraph of section 5.

- *(12, 15) "unexpectedly increasing": Given the very short period covered by the SF_6 observations, it is not clear that one should "expect" any particular sign for the trends. Determination of AoA trends from observations of stratospheric tracers is fraught with many uncertainties; even in models where an ideal, linearly increasing artificial tracer is used, one has to rely on zonal-mean results over long periods to obtain trends that are clearly statistically significant. Arguably, examination of AoA trends determined from observations of stratospheric tracers is not the best tool for documenting changes in the BD circulation. See Garcia et al. (2011).*

We acknowledge that it is important to provide proper context about this topic. Section 4.3 now starts with a new paragraph explaining some caveats of interpreting AoA trends as changes in BDC. The next paragraph has been expanded and corrected to justify the interest of the section:

> **It is delicate to infer changes in the BDC from the examination of AoA trends which are derived from observations of stratospheric tracers over periods shorter than several decades. Even in models where an ideal, linearly increasing artificial tracer is used, one has to rely on zonal-mean results over long periods to obtain trends that are clearly statistically significant (Garcia et al., 2011). The statement that is often made that climate models simulate a decreasing AoA throughout the stratosphere only applies over long time periods and is not necessarily the case for the past 25 years, when most tracer measurements were taken (Garfinkel et al., 2017). For example, the analysis of a 1700 year simulation showed that it takes around 30 years for a modeled BDC trend to emerge from the noise of natural climate variability (assuming a 2%/decade trend in the BDC; Hardiman et al., 2017).**
>
> **While linear trends of AoA over shorter periods may represent transient changes due to climate variability, changes over such intermediate timescales (i.e. intermediate between the QBO and the multidecadal scales) are still relevant to the study of stratospheric dynamics.** Current research on AoA trends has largely focused on a dipole-like latitudinal structure for the period 2002-2012, which was first derived from satellite observation of SF_6 by the MIPAS instrument (Stiller et al., 2012). This structure of trends shows AoA decreasing in the Southern Hemisphere but ~~**unexpectedly**~~ increasing in the Northern Hemisphere **which was used to explain a recent increase of stratospheric HCl in the Northern Hemisphere (Mahieu et al., 2014) and interpreted as the** consequence of a southward shift of the subtropical transport barriers (Stiller et al., 2017) .

- *(13, 1) "unexpected growth": Again, there are no clear expectations about trends for short periods.*

We have removed the word "unexpected" from this sentence.

- *(13, 23) "the reversal is found for all five reanalyses": This is true, but the reversals are in the opposite sense in ERAi and CFS vs. JRA-55, MERRA and MERRA2, so it is hard to know what to make of this.*

Yes but we still want to highlight this striking feature in case some reader does find what to make of this. We have changed this sentence to:

> This reversal is found for all five reanalyses and in all regions of the stratosphere **but it is difficult to interpret because it goes in opposite directions in ERA-I and CFSR versus JRA-55, MERRA and MERRA-2.**

- *(13, 25) "unexpectedly large disagreements": I am not sure why you think the disagreements are "unexpected". While presumably all reanalyses use more or less the same observational data, the manner in which the data are assimilated and the physical parameterizations included in the different models (in particular, those for convection and mesoscale gravity waves) are different. Note that at (14, 8) you suggest that "the disagreements found here may lie in the differences between the underlying models"; I agree that this is the most plausible working hypothesis.*

See above the reply to the first general comment. The word "unexpectedly" had already been removed before publication in ACPD.

Grammar, typos, etc.

All errors have been corrected.

Added references

Garcia, R. R., Randel, W. J., and Kinnison, D. E.: On the Determination of Age of Air Trends from Atmospheric Trace Species, J. Atmos. Sci., 68, 139–154, doi:10.1175/2010JAS3527.1, 2011.

Garfinkel, C. I., Aquila, V., Waugh, D. W., and Oman, L. D.: Time-varying changes in the simulated structure of the Brewer–Dobson Circulation, Atmospheric Chemistry and Physics, 17, 1313–1327, doi:10.5194/acp-17-1313-2017, 2017.

Hardiman, S. C., Lin, P., Scaife, A. A., Dunstone, N. J., and Ren, H.-L.: The influence of dynamical variability on the observed Brewer-Dobson Circulation trend, Geophys. Res. Lett., doi:10.1002/2017GL072706, 2017.

Krol, M., de Bruine, M., Killaars, L., Ouwersloot, H., Pozzer, A., Yin, Y., Chevallier, F., Bousquet, P., Patra, P., Belikov, D., Maksyutov, S., Dhomse, S., Feng, W., and Chipperfield, M. P.: Age of air as a diagnostic for transport timescales in global models, Geoscientific Model Development, 11, 3109–3130, doi:10.5194/gmd-11-3109-2018, 2018.