Review of the manuscript

The stratospheric Brewer–Dobson circulation inferred from age ofair in the ERA5 reanalysis

submitted to ACP by F. Ploeger et al.

S. Chabrillat, BIRA-IASB, February 2021

General Comments

The representation of the BDC and its time variations in reanalyses is an important topic for modellers of atmospheric composition in the middle atmosphere, and ERA-5 is set to become as broadly used as ERA-Interim has been for the twelve last years. Hence this study is of high interest and very timely for ACP. The modelling tools are very well tailored for these aims, and the modelling experiments are well designed. The manuscript is well structured and well written. All my comments deal with the description of the methods and observations, and the interpretation of the results. Hence I recommend publication of the paper after a minor revision (i.e. no additional calculations should be necessary).

In my opinion the manuscript comes short in the comparisons with trace gas observations (section 5). These suggest that the diabatic Age of Air (AoA) is overestimated in ERA-5, at least in the lower stratosphere at most latitudes (for some unspecified time period; Fig. 10a) and also in the middle-latitudes during the first half of the 1990's (Fig. 10b). It is repeatedly stated that the ERA-5 results are "*at the upper margin of the observational uncertainty range*" but these observational uncertainties are not clearly described in Figures 10 and 11, preventing a serious assessment of the significance of the apparent high bias using ERA-5. The AoA overestimation is confirmed by correlations between AoA and CFC-11, using as reference a campaign of aircraft observations (Fig. 11) but here also there is no indication of the observational uncertainties. These comparisons are briefly discussed in section 6 (lines 355-360) but mainly to tone down their importance and to avoid overinterpretation. The authors seem so unsure about these comparisons that they completely overlooked section 5 in the abstract.

It is well understood that these comparisons are not certain and that the same team is preparing a deeper investigation. Yet the results already shown in this manuscript are clearly of high interest for all readers and should be treated with due care. I recommend to better explain the observational uncertainties (see specific comments SC25 and SC26 below) and spend more effort to assess the significance of the overestimations of AoA by ERA-5 (see specific comment SC29 below). The abstract should highlight the results of section 5, and of course mention the discussion about their significance.

Major comments

MC1. The introduction and the discussion draw an often-stated parallel between the decreasing AoA found by some reanalyses in some hemispheres during some periods (maximum 30 years) and the decreasing AoA robustly found by most climate models over ~100 years due to increasing greenhouse gases. The validity of this parallel is doubtful, as correctly noted in the discussion (lines 392-393): "*The 30-40 years time series is presumably still too short for computing long-term trends in reanalysis, where stratospheric variability might be larger compared to climate models.*"
Fortunately this parallel is not necessary because AoA variations in the "actual world" over timescales of 1 to 3 decades are worth studying on their own. On this topic see also SC24 and SC31 below.

MC2. Section 2.1 explains two changes in methodology with respect to the previous age of air studies by this team: the mean age is now computed from a separate clock-tracer (rather than as the first moment of the AoA spectra); and the vertical velocity is not corrected any more for missing balance in the cross-isentropic mass flux. While the impacts of these changes are briefly described in that section and mentioned in section 4, they are not shown in the paper. In view of the long and successful series of AoA studies relying on ClaMS, it is important to show the impact of these changes on mean AoA. The impact of the first change especially could easily be shown on some existing figures, e.g. adding dashed black lines on Fig. 4 and 5 to show mean AoA from clock-tracer calculations.
It would be even more interesting to add a figure showing the Age spectrum for a given latitude and potential temperature. Ploeger et al. (2019) did precisely this in their Fig. 12a, showing clearly the importance of the spectrum tail in MERRA-2. In view of the apparent similarities between ERA-5 and MERRA-2 (when compared with ERA-I), this could be highly relevant here as well. Of course the mean AoA from the spectrum would also be compared with the mean AoA from the clock tracer.

MC3. Section 2 lacks a (short) description of the linear regression approach used for section 4: when de-seasonalized time series are regressed, do the linear regressions use "climate proxies" such as QBO, ENSO, volcanic forcings? If yes, what are the proxies actually used? The regression method also provides standard deviations that are used to assess the significance of the regressed trends (fig. 6). While this is standard, it remains important to explain how these standard deviations are ovtained (in a few words and with a reference).

MC4. Section 2.2 states (lines 139-140): "*However, we maintain the full vertical resolution. Hence, the ERA5 data to drive the ClaMS model in this study has 137 hybrid ECMWF model levels,* ***to compare with the*** *ERA–Interim 60 levels.*" How is it possible to maintain vertical resolution while going from sigma-pressure to sigma-theta? What is the actual vertical resolution in both simulations? The reader must get a sense that the ERA5 simulation makes full use of its increased vertical resolution. This is important as it certainly plays a role in the "diabatic heating gap" that is found at 350K (Fig. 1e) and repeatedly mentioned in the manuscript.

MC5. According to Diallo et al. (2020), the vertical coordinate that is used for derivation of the residual vertical velocity $\overline{w}^*$ is log-pressure height, i.e. a constant-pressure grid assuming a scale height of 7 km. Yet Fig. 2 uses "Altitude" for the Y-axis. This is confusing and prevents matching the levels of Fig. 2 with the iso-pressure levels explicitly drawn of Fig. 1, 3, 6, 8. If the actual Y-axis for Fig. 2 is log-pressure height, I recommend to re-draw Fig. 2 with pressures on the (logarithmic) Y-axes.

MC6. Fig. 8 compares the RCTT in ERA-5 and ERA-I, and the accompanying text states that "*Differences between mean age and RCTT are related to mixing effects*". Yet the corresponding paragraph discusses only the changes in RCTT and does not attempt to disentangle the contributions of mixing and residual circulation to AoA and its time variations. In practice it is not possible to visually compare AoA (Fig. 3) and RCTT (Fig.8) because AoA is shown sparately for DJF and JJA while RCTT is shown as an annual mean (also the color maps differ).

Specific Comments

SC1. Line 10: "…*changes in the assimilation system*" : reanalyses avoid any change in the assimilation system (i.e. the model and assimilation software); they suffer instead from changes in the observing system that is assimilated into the reanalysis. This formulation should be corrected throughout the manuscript.

SC2. Lines 16 and 20 : "*The Brewer-Dobson circulation (BDC) is the global transport circulation* ***in the stratosphere***" and "*The BDC is characterized by upwelling motion in the tropics, poleward transport* ***in the stratosphere*** *and downwelling above middle and high latitudes*." While these characterizations are very common, they overlook the important branch of the BDC that extends into the mesosphere where it goes all the way from the summer pole to the winter pole (e.g. Fig.1 in Bönisch et al., 2011).

SC3. Lines 40-43: "*On the one hand, climate models ~~show~~* ***predict*** *a robust strengthening and acceleration of the BDC with climate change (e.g., Butchart et al., 2010), manifesting in an increase in tropical upwelling and a decrease in global mean age of air. On the other hand…*" On what timescale does the climate model prediction appear? See MC1: the opposition between long-term AoA changes in climate models and multi-decadal AoA changes in reanalyses is misleading and not necessary.

SC4. Line 53: "…*while inter-annual BDC variability seems to be well represented in reanalyses…*". This statement should be more specific and have specific references because it is really not obvious to me, e.g. Chabrillat et al. (2018, Fig.10) found that the amplitude of the QBO impact on AoA can vary by a factor of 2 depending on the input reanalysis.

SC5. First paragraph of page 3: consider inserting here a reference to Fujiwara et al. (2017) as it is a general introduction of the reanalyses and their diversity, with a comprehensive description of the observing systems that they assimilate.

SC6. Lines 91-94: this paragraph should be expanded to introduce more smoothly diabatic transport and its advantages in the stratosphere (with references). It is espeecially important to define θ before it is used.

SC7. Line 131: please provide a web link or reference to the ECMWF technical documentation for IFS CY41R2. "*The horizontal resolution is about 30km (T639)*": explain (in a few words) the spherical harmonics number.

SC8. Line 133: Simmons et al. (2020) describe the need for the ERA5.1 update and its differences with ERA5.0. It is really necessary to cite that reference here.

SC9. Line 138: what is the spectral truncature (wavenumber) corresponding to this 1°x1° resolution? This is important to compare with the T639 number provided above.

SC10. Lines 141-146: For clarity, it is necessary to first explain in a few words what is the "temperature tendency". Is the total diabatic heating rate Q_{tot} the same quantity as the 5th temperature tendency provided by ERA, i.e. the "*mean temperature tendency due to* [both? all?] *parameterizations*"? Is Q_{tot} nothing else than the sum of the temperature tendencies due to short-wave and long-wave parameterizations in all-sky conditions? Please clarify.

SC11. $\mathrm{d}\theta/\mathrm{d}t$ is variously described in the text as the "cross-isentropic vertical velocity" (e.g. lines 145 and 163) or as the "total diabatic heating rate" (e.g. caption of Fig. 1). In view of Eq (3), I think that the second formulation is not rigorous. In any case the same formulation should be used consistently throughout the manuscript.

SC12. Line 166: "*Cross-isentropic tropical upwelling maximizes in boreal winter*". But the equinox seasons are not shown or discussed anywhere, and it is problematic to write about tropical upwelling "in" boreal winter. Consider instead: "*Cross-isentropic tropical upwelling* ***is larger in DJF than in JJA.***" This comment applies throughout the manuscript.

SC13. Line 170: "*…at lowest TTL levels around the level of zero radiative heating* ***in ERA-5*** *(about 350K)…*" This is important since there is no such level in ERA-I. But maybe I did not understand correctly (see SC27 below) – in such case the text should be clarified.

SC14. Line 175: for clarity I suggest to insert here a preview of the Discussion, inserting few words e.g. "*This much weaker upwelling in the TTL and tropical lower stratosphere causes stronger restrictions on large-scale advective upward transport in ERA5* ***and appears to correct an overestimation in ERA-I (see section 6).***"

SC15. Lines 178-179: "*The upward velocities in ERA5 in that region are more consistent with the residual circulation vertical velocity (see next paragraph and Fig. 2).*" This transition does not work – "see next paragraph" should be avoided. I suggest to move that sentence to the end of the next paragraph.

SC16. Lines 188 and 189: upwelling and downwelling are "stronger" but not "strongest" and tropical upwelling is stronger in DJF not in boreal winter (see SC12).

SC17. Lines 196-197: "*up to 30-40%, see Fig. 1*": normalized differences can not be seen on Fig.1 "*Above about 20 km…*": what is the corresponding theta? See MC5. A proper comparison would actually require a figure that overlays the two vertical profiles of tropical $\mathrm{d}\theta/\mathrm{d}t$ as function of θ with the two vertical profiles of tropical $\overline{w}^*$ as a function of p.

SC18. Figure 3c and 3d are difficult to read due to lack of contrast between the blue colors that are also obscured by the black contour lines. Consider removing these black color lines (pressure levels and zonal wind contours are well sufficient) and/or changing the color map.

SC19. Line 223: the differences are not so "*clear*" since closer inspection is required. In the tropics the differences are definitely not clearly visible and would require actual lineplots of the spectra (which could be worth adding; see MC2).

SC20. Line 226: the words "*against exchange with middle latitudes*" are superfluous

SC21. Line 237: "*…faster transport of young air towards the pole in ERA–Interim than in ERA5.*" This is not true everywhere: the northernmost latitudes in Fig. 5c show older modal age in ERA-I than in ERA-5.

SC22. Lines 244-252: the negative age trend with ERA-5 is significant in (nearly) the whole stratosphere for the period 1989-2018. Similarly, the trends found with ERA-I for 1989-2018 are either negative either positive but significant nearly everywhere. These significances should be highlighted.
Line 252: "*In the lowest tropical ~~and sub-tropical~~ stratosphere ERA5 shows **significantly** increasing age, ~~although non-significant changes in some regions,~~ compared to decreasing age **or insignificant trends** in ERA–Interim in this region.*"

SC23. Lines 265-271: I find this paragraph quite confusing. Here is a suggestion for a clarified version (assuming I understood correctly):
"*The clearest difference to ERA–Interim regarding structural age spectrum changes emerges at middle and high latitudes at upper levels (here 600K, Fig. 7b and d). ERA5, on the one hand, shows a shift of the spectrum to younger ages~~, although not as clear as in the tropics and in the SH.~~ **;** ERA–Interim, on the other hand, shows a decrease in the fraction of air younger than about 4 years and an increase of the fraction of older air. This increas**ing** fraction of air older than about 4 years in ERA–Interim indicates **that ERA-5 has** a strengthening in the deep branch of the residual circulation. The different spectrum changes in ERA5 and ERA–Interim cause the opposite mean age changes in the two reanalyses (Fig. 6).~~, and are related to different trends in the deep BDC branch (see next paragraph).~~*"
The end of this paragraph aims to introduce the follwing one: this transition should be re-worded (see also MC6).

SC24. Lines 299-304: please re-write for clarity and elaborate on the interpration. Here is a suggestion:
"*…mean age appears to increase before about 1991 **in both hemispheres** and after about the year 2000 ~~(except in the SH)~~ **in the NH** and decreases in between. These steplike changes are evident in both reanalyses**. In the beginning of the 1990's they** ~~and~~ could be related to **the Pinatubo eruption (i.e.** true atmospheric variability**) while in the end of the 1990's they could be related** to changes in the **observing system assimilated by the** reanalys**e**s. In particular…*"
Chabrillat et al. (2018) discussed the obscuring impact of the observing system change that happened less than a decade after Pinatubo (their section 3.2) so this reference may be used again here.

SC25. Figure 10a: It seems to me that these AoA use as source region the tropical 100 hPa levels (rather than the surface in all your other plots). Please check and (if correct) state this fact in the figure caption. What time period covered in the ClaMS simulations is used for this figure? Does it match the observational values taken from Waugh (2009)? What is the exact uncertainty range indicated by the error bars, i.e. do they show 1-σ (66% confidence level) or 2-σ (95% confidence level) on each side of the observation? Better explanations about this figure will allow improving the corresponding discussion and go a long way in addressing the general comment.

SC26. Section 5 needs to be substantively improved (see General Comments). Here are some questions and superficial suggestions:

Lines 310-311: "*...ERA5 mean age values are just at* ***outside*** *the upper margin of the uncertainty range of the in-situ SF 6 -based observations.* ***This apparent overestimation of mean age in ERA5 may be even larger in reality because*** *Leedham Elvidge et al. (2018) have* ***recently*** *shown that…"*

Lines 315-316: "*..ERA5 age is at* ***or above*** *the upper margin of the observational uncertainty range before about 1997, while ERA–Interim is ~~at~~ in the lower ~~margin~~* ***part of the range*** *(Fig. 10b).*"

Line 317: "*...the more gradual increase in ERA–Interim mean age...*" : over what period?

Line 320: "*...we note the trend values from a simple linear regression...*": over what period?

Line 328 and caption of Fig.11: "…(*for a detailed measurement data description, see Laube et al., 2020).*" What tracer was used to derive this observational AoA? This should be stated in the text, along with an estimate of the associated uncertainty and also the uncertainty of the CFC-11 observations. It would be even better to show these uncertainties graphically in Fig. 11 (if possible).

SC27. Line 347: "…*the minimum in tropical upwelling around the level of zero radiative heating (around 350K) is much lower for ERA5 than for ERA–Interim...*". I do not understand this formulation: is there a "level of zero radiative heating" in ERA-I as well? Do you mean something else than the minimum value itself? This is related to SC13 above.

SC28. Lines 353-354: it would be good to remind the reader about the corresponding pressure ranges i.e. "*...while in the tropical lower stratosphere* ***(50-150hPa)*** *the representation of tropical upwelling seems to beimprovedin ERA5, it is unclear whether the very weak total diabatic heating rates in the lower TTL* ***(150-300hPa)*** *are realistic.*"

SC29. Lines 355-360: this paragraph should be revised (see general comment). More specifically:

Line 355: "*...age of air in ERA5 is* <u>*slightly*</u> *high-biased...*" – but according to Fig. 10a this bias is significant and still ~1 year older, i.e. as large as the difference between MERRA-2 and ERA-I in Chabrillat et al. (2018) or Ploeger et al. (2019). Hence "slightly" seems rather subjective to me…

Line 358: "*It should be noted that these differences are small...*" – with respect to what?

Line 359: "…*when taking all uncertainties into account.*" This is a quite vague formulation, especially considering that the uncertainties were not sufficiently described (see SC25 and SC26).

SC30. Line 362: "*...which was argued by Stiller et al. (2017) to agree qualitatively with the structural circulation change observed by MIPAS.*" These MIPAS observations of structural circulation changes were first reported by Stiller et al. (2012). The agreement between AoA trends in observations and in ERA-Interim was first reported by Mahieu et al. (2014) using observations of HCl.

SC31. Line 369: "*The globally negative age of air trend in ERA5 over 1989–2018 agrees with results from climate model simulations, showing an accelerating BDC and decreasing age* <u>*over multi-decadal time scales*</u> *in response to increasing greenhouse gas concentrations.*"
This agreement may very well be for different reasons than increasing GHG concentrations because in ERA-5 NH the decreasing age is <u>entirely</u> due to the decrease in the 1990's (~1993-2003) i.e. a timescale much shorter than in the CCM simulations (and probably related to Pinatubo eruption). See also MC1 and SC24 above.

SC32. Line 395: "*…also before 1979 (as available from an extended ERA5 data product to be released soon).*" We can expect this backward extension to bring its own artifacts as the BDC will be much less constrained by satellite observations, with results more akin to those by climate models. In such a case the differences between CCM and ERA5 would most probably be related to shortcomings in the underlying model of ERA5 (i.e. IFS is an NWP model and not yet a climate model).

SC33. Line 404: "*…raises the question whether the remaining steplike change could be related to an incomplete bias correction in ERA5.*"
Or maybe to a bias correction that started too late? A cursory look at the ECMWF report on ERA5.1 (Simmons et al., 2020) would help in this part of the discussion.

SC34. Line 421: "*...and in the TTL ...*" According to your own discussion, it is not clear that the upwelling changes in the TTL are a correction.

SC35. Line 427: "*...largely related to a steplike change around the year 2000.*"
See above: your figures rather show two steplinke changes, one around 1993 and the other one around 2000.

SC36. Line 435: what is the availability of theobservational data shown in Fig. 10 and 11?

Typos, wording etc.

The points below are only suggestions to improve the wording and also to highlight a few typos:

- Line 4: "*ERA5-based results are compared to those* ***using*** *the preceding ERA–Interim reanalysis.*"
- Line 34: Waugh and Hall (2002) is a review paper on this diagnostic, hence the "e.g." is not necessary
- Line 101: "*…a pulse* ***period*** *of 2 months*".
- Line 153: "*For instance, the forecast data* ***starting*** *at 6UTC and* ***on the*** *5th hour forecast step…*". Same for line 161.
- Line 157: "***Here*** *the hourly ERA5 data is* ***not*** *used to drive model transport but data sub-sampled in time* ***is used instead, hence*** *the temperature tendencies…*"
- Line 178: "*where ERA5 shows upward* ***velocities*** *whereas* ***they are downward in*** *ERA–Interim.*"
- Line 205: "*…apparent in both reanalys****e****s.*"
- Line 207: "*...criti****c****ally…*"
- Line 220: abbreviation NH not yet defined
- Line 231: "*clear ~~detailed~~ differences occur in the* ***details of the*** *spectrum shape*"

- Line 246: "**However** above about 500 K **and in particular in the NH**, the signs of the trends…"
- Line 286: "*...shows increasing RCTTs,* ~~*clearest*~~ ***with strongest positive trends*** *in the NH.*"
- Line 290: "*…(compare* ***with*** *Fig.6).*"
- Line 293: "*…the* ~~*variability*~~ ***periodicity*** *is similar at all locations…*" or maybe "*the* ***timing of the*** *variability is similar at alo locations*"
- Line 309: "***The latitudinal distribution in*** *ERA-Interim* ***(Fig. 10a)*** *agrees well with…*"
- Line 372: "*... this* ***accelaration of the*** *residual circulation* ~~*acceleration*~~*…*"

Additional bibliographical references

Bönisch, H. et al. :On the structural changes in the Brewer-Dobson circulation after 2000 *Atmos. Chem. Phys.,* **2011***, 11*, 3937-3948

Fujiwara, M. et al.: Introduction to the SPARC Reanalysis Intercomparison Project (S-RIP) and overview of the reanalysis systems. *Atmos. Chem. Phys.,* **2017***, 17*, 1417-1452

Mahieu, E. et al.: Recent Northern Hemisphere stratospheric HCl increase due to atmospheric circulation changes. *Nature,* **2014***, 515*, 104-107

Simmons, A. et al.: Global stratospheric temperature bias and other stratospheric aspects of ERA5 and ERA5.1. *ECMWF Tech. Mem.*, **2020**, doi: 10.21957/rcxqfmg0