Atmospheric Chemistry and Physics Discussions

Atmos. Chem. Phys. Discuss., referee comment RC3
https://doi.org/10.5194/acp-2021-206-RC3, 2021

Comment on acp-2021-206

Anonymous Referee #3

Referee comment on "The Brewer–Dobson circulation in CMIP6" by Marta Abalos et al., Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2021-206-RC3, 2021

General comment:

This paper investigates the stratospheric Brewer-Dobson circulation in CMIP6 models in terms of residual circulation and mean age of air. Both the climatology and trends, for both past and future, are inter-compared. The results show quite some differences between the models regarding the climatology, but the paper states that all models are within the uncertainty range from observations. The model trends show a clear acceleration of the shallow branch but a less robust pattern of the deep branch, which is likely related to differences in wave forcing among models.

Overall, I regard this paper a very valuable model inter-comparison which is likely of high interest in the community. The paper is well written and the results are clearly presented. However, I think at some places the paper could still be improved to enhance its relevance, clarity and also some discussion could be placed into better context. Please note that the list of detailed comments below is clearly related to my strong interest in the subject and not to criticism of the paper. I rate my comments in this regard as minor and specific but would encourage the authors to elaborate a bit further on these, and hopw this would further improve the paper. After addressing the comments, I would strongly recommend publication.

Minor comments:

1. Comparison to CMIP5:
I can imagine that a key interest of readers when considering this paper concerns changes from CMIP5 to CMIP6 models. Several of these are briefly reported in the manuscript (e.g., L87, L266, L193). However, the paper could benefit significantly from a clearer presentation and discussion of these changes. In this spirit, I would suggest to include results from the previous CMIP5 project in several figures (regarding both climatology and trends), similar to what is done in Fig. 5 for CCMI models, for ease of comparison. Also,

the various differences between CMIP5 and CMIP6 could be discussed together in a short subsection.

2. Comparison to observed trends:
This comment is related to the discussion of the comparison of mean age trends from models with those from balloon observations, mainly on p8/L150ff. I think this is a point of high interest to many readers. However, after reading the paper, at least to me, the question still remains: Do we really have a discrepancy between simulated and observed trends? The paper states at several places that there is "an inconsistency in BDC trends" (e.g., L4, L343), but also says that the recent results of Fritsch et al. (2020) and also potential sampling issues in the observational data (as argued by Garcia and Randel, 2011) could explain these differences. So my question is (related to a comment of another Reviewer): How large is the remaining trend difference if one accounts for method and sampling uncertainties? If the proper sampling of model data is not possible, at least the discussion of these aspects could be clarified. And if the conclusions are not clear, I would avoid too strong related statements in the abstract and conclusions section.

3. Abstract:
I find the abstract somewhat unspecific and coming short in stating the main results of the paper. E.g., it is said that "CMIP6 results confirm the well-known inconsistency in BDC trends", or "paper reflects the current knowledge and main uncertainties" but it remains unclear what the "well-known inconsistency" or "current knowledge" are. I would recommend to avoid such unspecific terms but clearly state the results of the CMIP6 investigation regarding BDC climatology, trends, and forcing (even giving some numbers, e.g., MMM trend values).

4. Definition of deep branch:
In this paper, 1.5hPa is chosen as a characteristic level for the deep branch. The authors briefly say on p5/L99 that this level is substantially higher than used in most other studies (actually all studies I'm aware of). I don't really understand the reasoning given on the same page. Why should the fact that "upwelling minimizes at 1.5hPa" be a good reason for choosing this level as characteristic for the deep branch? As the 1.5hPa choice here is very different from other studies, it would be good to further clarify this argumentation. Also, I would recommend to add a discussion of previous studies on separating the deep from shallow branch and their criteria. Related questions I have are: Why do Lin and Fu (2013, Fig. 3) use 30hPa for the deep branch and find a seasonal cycle compared to the semi-annual cycle found here? In my view, also the Birner and Boenisch (2011) results would be more consistent with this choise and finding. Could it be that the 1.5hPa surface used here is indeed located above the actual BDC deep branch and the semi-annual cycle found here is actually related to the secondary circulation associated with the semi-annual oscillation and not the BDC?

5. Relation to surface warming:
It is stated that there is a "close connection between the BDC shallow branch and surface temp." (e.g., P16, L275), but I'm unsure how Fig. 12 proves that connection. If I understand the figure correctly, it just illustrates mass flux and surface temperature trends from the different models as ratios. If this is true, I don't see why this suggests causality. Wouldn't it be better to plot BDC trends vs. temperature trends for all

models/simulations (e.g., scatter plots), to see whether those models with strongest surface warming also simulate strongest BDC trends?

Specific comments:

P2, L41: I would explicitly state that the Fritsch et al. results also concern the age trends, e.g. add "...observational AoA and trend estimates... "
P6, L116: Regarding the comparison to observed age (Fig. 3) I would also state that a clear difference is the generally too weak extratropics-tropics age gradient in the models. Only CNRM and MRI models simulate subtropical gradients somewhat similar to observations, but these models show a high-bias in tropical age. In addition, I'm wondering how the age gradients here are related to upwelling differences (using the relation derived by Linz et al., 2016). At first glance, it seems to me that e.g. the GFDL model simulates a rather weak gradient but has a rather slow upwelling at 70hPa (Fig. 2) - this would actually be opposite as expected from the simple Linz et al. relation. Any ideas/comments?
P6, L120: I find this sentence about the Linz et al. paper at the end of this paragraph quiet confusing. Has the extratropics-tropics age gradient been used here?
P7, L148: Maybe the sentence "reanalyses do not provide robust trend estimates" is too general. I understand that the result of Abalos et al. (2015) which is referred to here, is the fact that different w* estimates provide different trends for ERA-Interim. Aren't most estimates considered in that paper (8 out of 9) rather robust in their trends, at least in parts of the stratosphere? I would suggest to reformulate like "... for upwelling due to difficulties when calculating w* trends from reanalyses".
Figure 4: What is the reason for the strange trend pattern in MMM 1998-2014 trends (Fig. 4f) in the tropics above about 10hPa? (A similar pattern occurs also in Fig. 7e).
P10, L183: Why is there larger inter-annual variability in the deep branch than in the shallo branch? It would also be helpful to say what variability is represented in the considered simulations (here for the UKESM model), and how reliable this is. Can the results based on UKESM simulations be generalized to other models, as the text here suggests?
P11, L195: "A similar but more modest behavior is found in the CMIP6 MMM...". Actually, I can see this widening only for HadGem and WACCM.
Figure 7: Why is the upwelling region for the GISS model so broad? Is this realistic? Maybe add some comment on this.
P15, L255: Any idea why the GFDL model has that low contribution from resolved waves? (Similar for WACCM at 1.5hPa, L263).
Figure 13: It would be good to add a measure of significance in Fig. 13 (and Fig. 12), e.g. the standard deviation from the regression as error bars.

Technical comments:

P2, L20: "Lagrangian"
P2, L26: "net strength in stratospheric tracer transport"
P3, L73: "significance in"
P11, L204: "substances"
P17, L293: delete "a" after "about"