

General comments:

The authors present a combination of Age-of-Air and Transformed Eulerian Mean (TEM) analyses based on CMIP6 simulations and supporting water vapor observations by the Aura Microwave Limb Sounder (MLS) to investigate the effect of increased greenhouse gases on the Brewer-Dobson Circulation (BDC) and compare more recent developments of the BDC to historical trends, as well as propose a vision of the future based on the 1ptCO₂ (annual increase of atmospheric CO₂ by 1%) scenario.

In its current form, the manuscript has several significant issues, which would demand additional attention to varying degrees.

Regarding its structural and linguistic form, the manuscript suffers from a few spelling mistakes and bad punctuation, which do make understanding the potentially novel ideas more difficult. Such issues should have been addressed prior to submission more rigorously. Parts of the manuscript, in particular the section „Discussion & Conclusions“, and in part the section „Results“, appear to unnecessarily repeat certain passages, without providing new information. From their discussion, the significance to repeat them in this form does not become clear to me. The usage of „codes“ used to describe certain simulation configurations or variables should be revised, as these are likely opaque to anyone not directly familiar with the simulation results. For example, „utendepfd“ does not help a first-time reader to understand what it is referring to, similarly to „i1p1f1“, which the authors have provided the means to interpret as „initialization method 1, physics 1 (whatever that means) and forcings 1“, but do not make clear what any of these configurations actually mean. Figures in the manuscript suffer from tiny and often redundant labels, which cannot be deciphered without greatly zooming in. The size of these figures also varies greatly, which is not critical, but could be improved upon.

Regarding its content, the authors present their analysis as a form to utilize many forms of analyses in a unified framework as a proposed framework for future analyses, but ultimately present a collection of several, established analyses, whose synergy mostly lies on the qualitative side. While I do not criticize each analysis itself, the initial framing of the work might have been a bit unfortunate. The analyses themselves (naturally) contain the same results, since most models are indeed consistent with themselves across different properties. For example, the distribution of Age-of-Air does depend on the transport scheme used in a model, and therefore any quantity derived from such a scheme cannot directly contradict the scheme itself, to some extent. The argument for this particular approach is therefore relatively weak. Regarding the novelty of the results, I have to question if there are insights gained from this study that were unknown to me previously. An acceleration of the BDC under increased greenhouse gas forcing has been hypothesized by Rind et al. 1990. Streamline functions have been used by to investigate the trends of the BDC e.g. Okamoto et al. 2011, which do attribute the acceleration to increased wave activity. The analysis of using water vapor anomalies to quantify the BDC is loosely based on Flury et al. 2013, which the authors acknowledge. Indeed, while the authors do reference previous work adequately in the respective „Results“ subsections, the novelty of their results does come into question.

I found that some results presented in this manuscript, mostly through the various figures, are not necessarily well supported. Overall, unfortunately I cannot recommend the manuscript in its current form for publication. Parts of the analyses should be revised and it should be made sure that they all support the overarching claim of the manuscript. The style of the manuscript has a lot of potential for improvement, and maybe this work could be rephrased as a comprehensive study on BDC dynamics under greenhouse gas forcings rather than a new analysis framework, if the combination of these different analysis methods yields some benefit.

Minor comments:

Overall spelling and punctuation must be improved. If requested, I can provide a list of the errors I found, but they are numerous and trivial to fix. Figures need to be adjusted/redone to ensure

visibility. The linguistic style of the manuscript could be adjusted to more closely resemble standard English. Equations should be numbered, unless trivial (such as the conversion from volume mixing rate to part per million). The lag correlation coefficient should be expressed here for a discrete time series, the summation limits need to be corrected.

Abbreviations should be revised to allow a more intuitive understanding of the abbreviation without needing to look up the meaning from a few pages ago. Citations sometimes are missing brackets or unnecessarily have them. Citations need to actually support the statements they are mentioned with. ACP conventions should be upheld. A summary at the end of the manuscript would be very useful here.

Major comments:

- What are the weights used for your low-pass filter? If I'm reading the manuscript correctly, these weights are never mentioned. They are an essential part of your methodology. Every choice of parameters will have an influence on the results.

- Line 121: „...necessitating ensemble averaging to obtain reliable conclusion...“

I take it that you average mean Age-of-Air and vertical wind speeds. As mentioned in section „Limitations and Future Perspectives“ as a third point, generally you cannot trivially average mean Age-of-Air distributions for the reasons mentioned in this section. There are effects such as multi-model mean dilution, biased in absolute values of Age-of-Air and such. I am willing to assume that this has been considered in this analysis, but it must be addressed in the manuscript as well.

- State what „normalized“ and „standardized“ mean

The term „normalization“ is generally used quite broadly, but in my field it refers to scaling to the interval (0, 1). Standardization usually refers to scaling to zero-mean, unit variance. If you do this – motivate why you do it. You must differentiate between scaling between multiple models and within multiple models. For the later, the resulting scaled cannot be trivially compared to each other. As of now, this raises a lot of unnecessary, technical questions.

- Usage of Pearson correlation coefficients in table 5

The Pearson correlation coefficient requires the samples to be statistically independent. Applying a smoothing filter does violate this assumption, which could explain these very low values. This metric is not suitable here.

- Core statement of the manuscript

What are the actual benefits of using many analysis methods in tandem? I am missing some form of synergy. As presented right now, the manuscript presents the findings to simply agree with each other, implying the opposite claim that a single analysis method would have been sufficient. However, some results appear to actually not agree with each other, actually contradicting the initial statement. Ideally, these apparent contradictions would be resolved through the interplay of the different methods, which they do not appear to do so. This needs to be revised.