Author response to referee comments on Chrysanthou et al. "The effect of atmospheric nudging on the stratospheric residual circulation in chemistry-climate models" submitted to ACPD

Reply to Anonymous Referee #2

In this paper, the authors compare the free running and specified dynamics simulations of the CCMI models. They find that although the nudging involved in the SD does make the models short term variability line up, the residual circulation in the nudged runs matches neither the original models nor the reanalyses to which the models are nudged.

This paper is important. It is interesting and provides a cautionary tale against using these SD runs to examine transport of long-lived trace gases, which one would expect would be significantly impacted by these inconsistencies. It is well reasoned and well written. I recommend publication following some minor revisions. An expanded introduction would be extremely valuable to the reader, and I have included recommendations for that below. The primary scientific points that would add significantly to the paper are:

1. Discuss the choice of metric for the BDC in more detail. (Line 191)
2. Reconsider the results of the multiple linear regression in light of the importance of the relative phasing of the annual cycle and the QBO.
3. Consider the differences between the models and the reanalyses they are being nudged towards, especially vertical profiles, to try to better understand why the results are what they are.

Thank you for your positive comments and suggestions for improving the paper. We reply to the specific points raised below in red.

Specific comments:

1) Recent work (Abalos et al. 2015, Linz et al. 2019) has raised the issue that the differences due to different calculations of the residual circulation are as significant as (or more than) the differences between the reanalyses themselves (c.f. Seviour et al. 2012, also). Therefore, I would appreciate more discussion of the choice made here to use the TUMF and downward control (which is used to distinguish the different wave forcing components). How do these choices impact the results, if at all? A thermodynamic calculation might be very different, since the temperature is nudged, which will introduce spurious diabatic terms. How is TUMF different from w* averaged between turnaround latitudes?

Thank you for this suggestion. We agree that some further discussion of the chosen measures of residual circulation is important to place the results into a broader context. As indicated by the reviewer, the choice of diagnostics is likely to be a particular issue for interpreting the nudged models and reanalyses due to the additional tendencies imposed in the model equations. Part of our choice is pragmatic, based on the availability of output from the CCMI model dataset, which provides direct estimates of residual circulation diagnostics following Andrews et al. (1987) and resolved and parameterized wave forcing components. As TUMF is an integrated measure of w* between TA latitudes, the results regarding interannual variability and long-term trends are similar to if an average w* measure is used. Arguably, the TUMF is a first-order measure in order to evaluate whether the mass circulation is affected by nudging and it has been widely used in previous multi-model comparison and reanalysis studies and so we include it here (e.g. Butchart et al., 2006, Butchart et al., 2010, Butchart et al., 2014 and references therein; Rosenlof, 1995; Seviour et al., 2012; SPARC CCMVal, 2010). The important point is that we use consistent diagnostics in the free running and nudged experiments and the reanalyses, so that we are always comparing the same measures across the different simulations. Unfortunately, as heating rates are not provided by all models in the CCMI output, it would not be possible to compute a consistent (i.e. with the appropriate model radiation scheme) estimate of the residual circulation using the thermodynamic equation. It is possible that for the nudged simulations calculation of the residual circulation based on the thermodynamic equation might result in different behaviour to the direct TEM formulation used here. We now suggest this as an interesting topic for future research in the conclusions. The downward control principle (DCP; Haynes et al., 1991) calculations provide a framework to assess the extent to which the wave forcing and directly estimated residual circulation become decoupled as a consequence of the additional tendencies imposed by the nudging.

Based on the above we have expanded/edited the relevant parts of the text to the following :

End of section 2.2.1, (just before the introduction of DCP) :

"The TUMF has been used widely as a measure of the strength of the BDC (Rosenlof, 1995; Butchart et al., 2006, Butchart et al., 2010, Butchart et al., 2014 and references therein; Seviour et al., 2012), so it's use here enables a direct comparison with earlier studies. Arguably, the strength of the TUMF is a first-order metric to evaluate changes to the mass circulation as a consequence of nudging. As mentioned above, by calculating the annual means of TUMF accounting for the seasonal cycle of the turnaround latitudes we capture the correct evolution of the intraseasonal (not shown) and interannual variability in the TUMF."

End of section 2.2.2, (after introduction of DCP) :

"Both the direct and DCP methods rely on the applicability of the quasi-geostrophic approximation to interpret the results. We note that in addition to the direct and DCP approaches used here, the residual circulation can also be estimated using the thermodynamic

equation. Unfortunately, heating rates were not available from all CCMI model simulations to perform this calculation. Studies have shown that the estimates from the different methods for evaluating the residual circulation can differ (Abalos et al., 2015; Linz et al., 2019; Seviour et al., 2012), particularly in reanalyses where the normal conservation laws are not required to be met. Similar issues are likely to beset the nudged model simulations owing to the additional tendencies included in the model equations. The differences between the calculation methods for the residual circulation can be as large as, or larger than, the differences between reanalysis datasets for the same diagnostic (Abalos et al., 2015; Linz et al., 2019), and may further depend on choices around averaging between fixed latitudes or the turnaround latitudes (Linz et al., 2019), so it is important to bear this in mind in interpretation of the results presented here. Nevertheless, we compute the diagnostics for the residual circulation in a self-consistent manner in the models and reanalyses to enable comparison with earlier multi-model studies (e.g. Butchart et al., 2006, Butchart et al., 2010, SPARC, 2010)."

2) Although the method used here to do the multiple linear regression is somewhat standard, it does not actually remove all of the variability that is related to these signals. In particular, the relative phasing of the QBO and the annual cycle will cause an enhanced wintertime upwelling in some years, so the result that the leftover variability is highly correlated could be related to this. Randel et al. 1999 have a nice treatment. I would consider using the method in that paper (described on p. 458, see also Randel and Wu 1996) to determine if the relative timing of the annual cycle and QBO is the cause of the covariation of the nudged time series. This phenomenon is also mentioned in the cited Baldwin et al. 2001 QBO review.

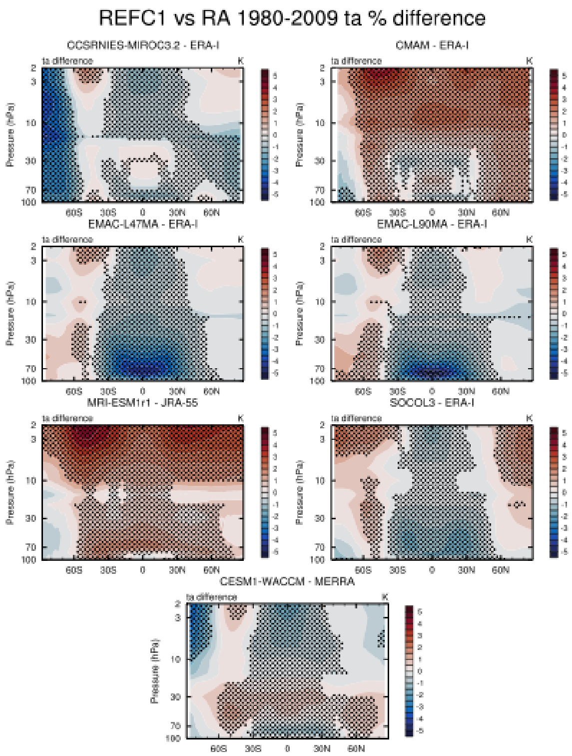
Thank you for raising this interesting point. The MLR model is applied to the annual mean TUMF timeseries and hence the QBO terms are resolved at an annual timescale. Therefore, the method presented by Randel et al. (1999) to account for interannual variations in the relative phasing of the QBO with the annual cycle is not readily applicable in our analysis. We further note that all of the models except CMAM and MRI-ESMr1 also nudge the phase of the QBO in the REF-C1 simulations (Morgenstern et al., 2017). Hence, if this is a main factor for explaining the MLR residuals, we expect that such an effect would also be seen in the REF-C1 simulations. However, the temporal correlation of the residuals is much weaker in the REF-C1 experiments for those models that nudge the QBO, suggesting this is not a major factor for explaining the correlated residuals in the REFC1-SD experiments.

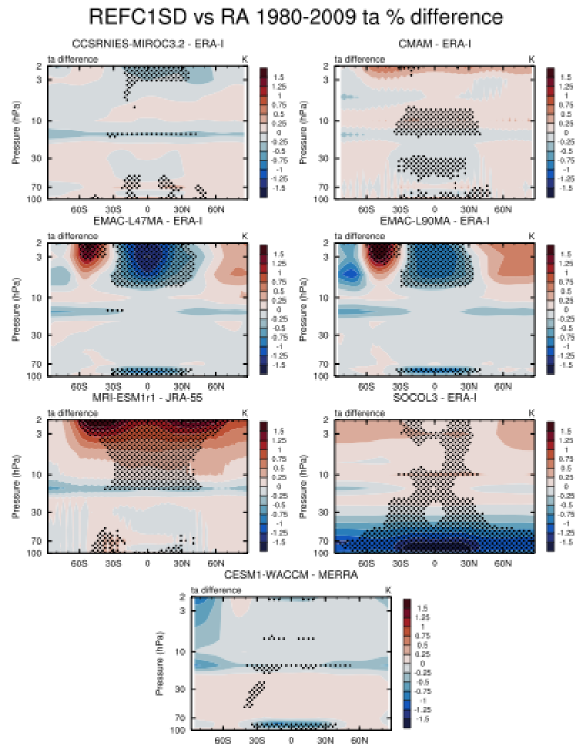
3) The authors find that the reanalyses differ from the nudged runs which in turn differ from the free running models. I would like to see some presentation of the difference between the mean states of the reanalyses and the mean states of the models without nudging. The result that the models tend to have stronger upwelling in their SD runs could be caused by a systematic low bias in the model temperature in the tropics compared to the reanalysis temperature.The repeated

nudging with a primarily positive value would then cause an apparent increase in upwelling, without changing the underlying model physics that are the cause of any systematic mean state bias. To understand why the nudging does what it does, I think the difference between the free running models and the reanalyses needs to be explored in much more detail–perhaps for just one model. A continuous spurious forcing because of a mismatch in the mean state (and likely also in the seasonal cycle) should impact the mean while leaving the interannual variability alone (line 612).

Thank you for this suggestion. We present below the percent differences in zonal mean temperature in the stratosphere between the REF-C1 simulations and the respective reanalysis they are nudged towards (stippling denotes statistical significance at 95% based on a two-tailed Student's t-test); the equivalent differences for the REFC1-SD simulations are also shown in the right panels (note the reduced colour scale).

Most models show a cold bias relative to the reanalyses in the tropical tropopause layer, except for MRI-ESMr1 which shows a warm bias. The two versions of EMAC show the largest relative cold bias in the TTL. Much of this bias is alleviated in the nudged simulations, as expected, but EMAC (l47/90), CESM1-WACCM and SOCOL3 still show a (smaller) cold bias in the TTL. Larger differences remain above the level where nudging ceased to be imposed (e.g. 10 hPa in EMAC and 40 hPa in MRI-ESM1r1). The opposite sign of the TTL temperature bias in the free-running MRI-ESMr1 model provides a useful test bed for the point raised by the reviewer. The REFC1-SD MRI-ESMr1 simulation still shows stronger upwelling near the equator compared to the free-running model (Figure 3c) despite having a warm bias; this result is similar to the effect of nudging in the other models that show a cold TTL bias. Hence from this cursory investigation, we do not identify a systematic relationship between the intrinsic model bias and the effect of the nudging on the residual circulation. To get a better handle on this, a detailed analysis of the sensitivities within one model to relaxation timescales and nudging parameters such as the altitude range etc. would be needed in order to robustly establish how nudging affects the mean residual circulation.





Minor comments

I find the second paragraph of the introduction to be lacking. Although Abalos et al. (2015) did conclude that there was likely a 2-5%/decade trend in the circulation, they found substantial differences with different calculations for the circulation strength. Ploeger et al. (2019) have a new paper out on the inconsistency of the BDC trends for the different reanalyses, looking at the age spectra. Age of Air is mentioned somewhat abruptly and with no explanation. The results are also misrepresented. Engel et al. 2009 did not find a statistically significant increase in the mean age of air in the midlatitudes. They found no statistically significant change. There are limitations both on the availability of measurements and on the interpretation of age of air as a direct proxy for the residual circulation, and much more recent papers discuss this, rather than Waugh and Hall 2002. There are new satellite based measurements–ACE-FTS (e.g. Ray et al. 2016), MIPAS (Haenel et al. 2015), and MLS (Linz et al. 2017)–and new measurements from Air Core (Engel et al. 2017). The interpretation of the data is also difficult because of the separation of the residual circulation from the mixing (e.g. Garny et al. 2014, Ploeger et al 2015, as mentioned in the first paragraph–perhaps just move this comment to this paragraph?). The only attempt to calculate the residual circulation from tracer data did not calculate the residual circulation directly, but instead the diabatic circulation (Linz et al. 2017), and they did not calculate trends.

Thank you for this comment. We have expanded this part of the introduction to give a more balanced and detailed synopsis of the literature:

"Past studies have shown substantial spread across models in the mean strength of the residual circulation (e.g. Butchart et al., 2010). Nevertheless, climate and chemistry-climate models (CCMs) consistently simulate a long-term strengthening of the residual circulation with an increase of ~2% decade^{-1} (e.g. Butchart et al., 2010; Hardiman et al., 2014), though there are differences across models in the relative contribution to trends from resolved and parameterized wave forcing. Reanalysis datasets also suggest a strengthening of the residual circulation over the past several decades of the order 2-5% decade^{-1} (Abalos et al., 2015; Miyazaki et al., 2016) apart from one (ERA-Interim) which shows a weakening of the deep branch of the BDC (Seviour et al., 2012; Abalos et al., 2015). However, reanalyses are subject to multiple caveats, particularly in their suitability for trend studies, and there are substantial differences in residual circulation trends calculated from the same reanalysis using different methods (Abalos et al., 2015). Given the limitations of reanalyses, evaluating the fidelity of model estimates of residual circulation variability and trends is challenging since there are no direct measurements of the residual circulation.

The only direct estimates of the stratospheric mass circulation come from tracer measurements, which can be used to calculate stratospheric age-of-air (AoA) (Kida, 1983; Schmidt and Khedim 1991; Waugh and Hall, 2002). AoA represents the combined effects of advection by the residual circulation and turbulent mixing processes, and as such cannot be directly related to the residual circulation. While progress has been made in separating the relative effects of these two factors for AoA calculated from models (Garny et al., 2014; Dietmuller et al., 2018, Eichinger et al., 2019, Šácha et al., 2019), using Lagrangian models driven by reanalysis fields (Ploeger et al., 2015a, Ploeger et al., 2015b, Ploeger and Birner, 2016, Ploeger et al., 2019) and comparing these effects in both CCMs and Lagrangian models (Dietmuller et al., 2017), this is more difficult to achieve in observations. Engel et al. (2009) and Stiller et al. (2012) used balloon-borne measurements of stratospheric trace gases and found a statistically non-significant increase in AoA in the middle stratosphere at northern mid-latitudes, which has been corroborated by more recent study using balloon-borne measurements at two midlatitude sites in the Northern Hemisphere (Engel et al., 2017). It has been hypothesised based on analyses of more recent satellite datasets, which have greater spatial and temporal coverage, that the subtropical AoA trends can be explained by a weakening of the mixing barriers at the edge of the tropical pipe (Neu and Plumb, 1999) that is masking the effects of an increase in tropical upwelling on AoA (Stiller et al., 2012; Haenel et al., 2015). In contrast with the AoA trends derived from observations, CCMs forced with observed sea-surface temperatures (SSTs), greenhouse gases and ozone-depleting substances show a decrease in AoA throughout the stratosphere (World Meteorological Organization, 2018; Li et al., 2018; Morgenstern et al., *2018; Abalos et al., 2019; Polvani et al., 2019). Theoretical approaches based on the tropical leaky pipe model (Neu and Plumb, 1999) have shown promise for bridging the information on the stratospheric circulation derived from observations with outputs from GCMs/CCMs (Ray et al., 2016).*

As explained above, it is difficult to infer information on the residual circulation from AoA without carefully accounting for the effects of mixing. More recent theoretical developments offer a means of calculating the diabatic circulation using stratospheric tracers (Linz et al., 2017), which is a promising avenue as this is more closely related to the residual circulation than AoA. The results of Linz et al. (2017) showed consistent estimates of the diabatic circulation in the lower stratosphere, but large uncertainties of up to a factor of two in the mean strength in the upper stratosphere. Targeted measurement strategies to better characterise long-term changes in the observed stratospheric meridional circulation have been proposed (Moore et al., 2014; Ray et al., 2016).''

A more detailed explanation of what nudging is, how it is implemented, why it is used, and the motivation for this work would help the introduction. The current paragraph in the introduction (line 85) is good, and I think that including more specific examples would be very helpful. Similarly, in the conclusion, commentary on what these nudged runs could be useful for would be appreciated in addition to the entirely appropriate cautionary words.

Thank you for this comment. We have expanded this paragraph to include your suggestions above. The relevant part of the manuscript is now (nb: black italic below denote unchanged parts of the manuscript) :

''In an attempt to obtain a closer comparison with observed stratospheric trace species, some studies have used model simulations with meteorological fields nudged or relaxed towards a reanalysis dataset (Jeuken et al., 1996); these include many studies of ozone variability and trends (e.g. van Aalst et al., 2003; Solomon et al., 2016; Hardiman et al., 2017b; Ball et al., 2018), comparisons between models and satellite-based multi-species observational records (Froidevaux et al. 2019) as well as the chemical and climatic effects of volcanic eruptions (Löffler et al., 2016; Solomon et al., 2016; Schmidt et al., 2018). Nudged simulations have also been used to study mechanisms for dynamical coupling between the stratosphere and troposphere (Hitchcock and Simpson, 2014). Nudging involves adding additional tendencies to the model equations by constraining the simulated meteorological fields to reanalysis fields (Kunz et al., 2011). Nudged variables can include horizontal winds (or divergence and vorticity), temperature, surface pressure, latent and sensible heat fluxes. However, vertical winds, which are small residual from horizontal divergence, are not nudged and the underlying model physics can yield quite different results from the reanalysis that they are nudged towards (Telford et al., 2008; Hardiman et al., 2017). A recent study by Orbe et al. (2018) analyzed tropospheric tracers

in nudged CCM simulations and found large differences in the distributions of the tracers, which could be partly traced to differences in the model convection schemes. They urged users to adopt a cautious approach when interpreting tracers in nudged simulations given their dependence on not only large-scale flow but also sub-grid parameterizations. In contrast, critical evaluation of the stratospheric residual circulation in models relaxed to reanalysis fields has been rather limited to date."

As for the conclusions the now expanded relevant part of the manuscript (second from last paragraph of the manuscript):

"Despite the limitations described here, some success has been reported in studies that used nudged simulations to investigate specific meteorological events, such as Sudden Stratospheric Warmings, and in particular for exploring processes beyond the top of the nudging region in the Mesosphere-Lower Thermosphere (e.g., Tweedy et al., 2013; Chandran and Collins, 2014; Pedatella et al., 2014). In order to reduce discrepancies between nudged and free-running simulations, various nudging techniques have been investigated. The role of gravity waves in the error growth that the nudging introduces over time has been highlighted for a single model (Smith et al., 2017). Constraining just the horizontal winds without the temperature was found to be a good strategy when investigating the aerosol indirect effects without affecting significantly the mean state (Zhang et al., 2014)...

...A dedicated study of the sensitivities within one model to relaxation timescales, nudging parameters, nudging height range, and model resolution would be needed to offer a detailed explanation for these differences."

Do any of the "free running" models have a nudged QBO, and if so, which ones?

All of the free-running models presented in the study apart from CMAM and MRI-ESMr1 impose a QBO "nudging" in their simulations as mentioned in the last paragraph of section 3.4 (Table S8 of supplement of Morgenstern et al., 2017).

Since the focus of this paper is really on the difference between SD and free running models, it might make sense to omit the models that do not include both. I respect the authors choice to have those comparisons in the supplementary information. However, I think I would find the plots more manageable without the additional lines.

Thank you for this recommendation. In response to this comment and also the comments of reviewer 1 we have removed the models which did not perform the SD simulation from the main body of the paper (GEOSCCM, NIWA-UKCA and ULAQ-CCM) and the main text now focuses on the seven models for which both REF-C1 and REFC1-SD simulations are available. For completeness a subset of diagnostics are shown for those models in the supplementary material.

Line 106: aren't they necessarily inconsistent?

We have corrected this to: "*will add forcings that are inconsistent with the model state*."

Lines 100-129: excellent discussion.

Thank you for the positive comment!

Line 153: Do you mean Table 2?

Thank you for spotting this, we have now corrected this.

Line 155 etc.: More details here would be good. I don't understand what is meant by T (with wave-0), and I assume that "temperature" is also T (for CMAM), but if that's true then they all nudge temperature so maybe mention that? Including the sentence "Nudging timescales range from 6 - 50 hours." (or some equivalent) would also emphasize how different these treatments are.

We have added some more details here as per your recommendation. In Table 2, temperature is referred to consistently as T. The wave-0 was intended to refer to the additional nudging of the global mean temperature in EMAC. We now state this explicitly. We have also changed L152-153 in the text to: "*The REF-C1SD simulations nudge temperature and other meteorological fields such as horizontal winds, vorticity and divergence and some surface fields (Table 2), while the chemical fields are left to evolve freely. The nudging timescales range from 6 - 50 hours and the height range over which nudging is applied varies (Table 2).*"

Line 159-160: Table 1?

Thank you for spotting this, we have now corrected this.

Line 168: "primary variable"–not sure what is meant by this

We meant prognostic. We have now corrected this.

Line 442: I have trouble seeing this from the plot. Could you provide the numbers? Is the difference in the peak-to-peak amplitude significant?

Yes, due to the congestion of the multiple lines is not that easy to see but the difference is not significant. This sentence now includes the numbers of the peak-to-peak amplitude

"Comparing the MMM annual cycle of the REF-C1 runs with the MMM REF-C1SD of Figure 7b reveals that on average the REFC1-SD models show a slightly larger peak-to-peak amplitude (0.22 mm s^{-1} vs. 0.19 mm s^{-1}) in the annual cycle."

Line 476. . .: How do these results relate to Abalos et al. 2014, who found such a strong dependence of the upwelling on the location of the EPFD?

Thank you for this interesting point. We agree it would be interesting to examine how the patterns of EPFD vary between the free running and nudged simulations and how this relates to fluctuations in upwelling. This work is planned for a follow-up study.

Line 488-90: I found this sentence confusing.

Thank you for spotting this unclear sentence. We have removed this sentence as it was related to an earlier sentence on lines 483-485, which has been edited to:

"This result indicates that although nudging does not constrain the mean residual circulation, it does constrain the interannual variability and produces very similar contributions to variability across models from both resolved and parameterized wave forcing.

Line 498: remove "due to"

Done.

Section 3.6: Nice.

Thank you for the positive comment!

Line 606.: This paragraph is very long and contains a lot of important information. Turn into two or even three shorter paragraphs?

Thank you for this comment, we have now turned this paragraph into three shorter ones to improve the clarity.

Figures: Consider stippling the insignificant part of the plot. I read this on a number of different devices, and on some it looked fine, but on one of them the plots were impossible to see because the colors weren't visible behind the stippling. This wasn't an issue for the last two figures for some reason.

It may take a while for these plots to be loaded properly in a pdf viewer, as they are quite large files. We changed the subtleties of the stippling patterns as seen in the figures and hopefully they are more clear now.

It could be useful to have models that are nudged to the same reanalysis plotted in the same color family to aid interpretation of the complicated line plots.

We have now colour coded the model results nudged towards ERA-I in red(-ish) colours. We hope that without the additional lines of the extra models, the results are more clear.

References:

Linz et al. 2019: https://www.atmos-chemphys.net/19/5069/2019/acp-19-5069-2019.html
Randel et al. 1999 https://acomstaff.acom.ucar.edu/randel/JAS%201999.pdf
Randel and Wu 1996
https://journals.ametsoc.org/doi/10.1175/1520-0469%281996%29053%3C2546%3AIOTOQI%3E2.0.CO%3B2
Ploeger et al. 2019 https://www.atmos-chem-phys.net/19/6085/2019/
Ray et al. 2016 https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2015JD024447
Haenel et al. 2015 https://www.atmos-chem-phys.net/15/13161/2015/acp-15-13161-2015.html
Linz et al. 2017 https://www.nature.com/articles/ngeo3013
Abalos et al. 2014 https://journals.ametsoc.org/doi/10.1175/JAS-D-13-0366.1

Abalos, M., Legras, B., Ploeger, F. and Randel, W. J.: Evaluating the advective Brewer-Dobson circulation in three reanalyses for the period 1979-2012, J. Geophys. Res. Atmos., 120(15), 7534–7554, doi:10.1002/2015JD023182, 2015.

Abalos, M., Polvani, L., Calvo, N., Kinnison, D., Ploeger, F., Randel, W. and Solomon, S.: New Insights on the Impact of Ozone-Depleting Substances on the Brewer-Dobson Circulation, J. Geophys. Res. Atmos., 124(5), 2435–2451, doi:10.1029/2018JD029301, 2019.

Andrews, D. G., Holton, J. R. and Leovy, C. B.: Middle Atmosphere Dynamics, International Geophysical Series, Vol. 40, 1987.

Butchart, N.: The Brewer-Dobson circulation, Rev. Geophys., 52(2), 157–184, doi:10.1002/2013RG000448, 2014.

Butchart, N., Scaife, A. A., Bourqui, M., Grandpré, J., Hare, S. H. E., Kettleborough, J., Langematz, U., Manzini, E., Sassi, F., Shibata, K., Shindell, D. and Sigmond, M.: Simulations of anthropogenic change in the strength of the Brewer-Dobson circulation, Clim. Dyn., 27(7–8), 727–741, doi:10.1007/s00382-006-0162-4, 2006.

Butchart, N., Cionni, I., Eyring, V., Shepherd, T. G., Waugh, D. W., Akiyoshi, H., Austin, J., Brühl, C., Chipperfield, M. P., Cordero, E., Dameris, M., Deckert, R., Dhomse, S., Frith, S. M., Garcia, R. R., Gettelman, A., Giorgetta, M. A., Kinnison, D. E., Li, F., Mancini, E., McLandress, C., Pawson, S., Pitari, G., Plummer, D. A., Rozanov, E., Sassi, F., Scinocca, J. F., Shibata, K., Steil, B. and Tian, W.: Chemistry–Climate Model Simulations of Twenty-First Century Stratospheric Climate and Circulation Changes, J. Clim., 23(20), 5349–5374, doi:10.1175/2010JCLI3404.1, 2010.

Chandran, A. and Collins, R. L.: Stratospheric sudden warming effects on winds and temperature in the middle atmosphere at middle and low latitudes: a study using WACCM, Ann. Geophys., 32(7), 859–874, doi:10.5194/angeo-32-859-2014, 2014.

Dietmüller, S., Garny, H., Plöger, F., Jöckel, P. and Cai, D.: Effects of mixing on resolved and unresolved scales on stratospheric age of air, Atmos. Chem. Phys., 17(12), 7703–7719, doi:10.5194/acp-17-7703-2017, 2017.

Dietmüller, S., Eichinger, R., Garny, H., Birner, T., Boenisch, H., Pitari, G., Mancini, E., Visioni, D., Stenke, A., Revell, L., Rozanov, E., Plummer, D. A., Scinocca, J., Jöckel, P., Oman, L., Deushi, M., Kiyotaka, S., Kinnison, D. E., Garcia, R., Morgenstern, O., Zeng, G., Stone, K. A. and Schofield, R.: Quantifying the effect of mixing on the mean age of air in CCMVal-2 and CCMI-1 models, Atmos. Chem. Phys., 18(9), 6699–6720, doi:10.5194/acp-18-6699-2018, 2018.

Engel, A., Möbius, T., Bönisch, H., Schmidt, U., Heinz, R., Levin, I., Atlas, E., Aoki, S., Nakazawa, T., Sugawara, S., Moore, F., Hurst, D., Elkins, J., Schauffler, S., Andrews, A. and Boering, K.: Age of stratospheric air unchanged within uncertainties over the past 30 years, Nat. Geosci., 2(1), 28–31, doi:10.1038/ngeo388, 2009.

Froidevaux, L., Kinnison, D. E., Wang, R., Anderson, J. and Fuller, R. A.: Evaluation of CESM1 (WACCM) free-running and specified dynamics atmospheric composition simulations using global multispecies satellite data records, Atmos. Chem. Phys., 19(7), 4783–4821, doi:10.5194/acp-19-4783-2019, 2019.

Garny, H., Birner, T., Bönisch, H. and Bunzel, F.: The effects of mixing on age of air, J. Geophys. Res., 119(12), 7015–7034, doi:10.1002/2013JD021417, 2014.

Haynes, P. H., McIntyre, M. E., Shepherd, T. G., Marks, C. J. and Shine, K. P.: On the "Downward Control" of Extratropical Diabatic Circulations by Eddy-Induced Mean Zonal Forces, J. Atmos. Sci., 48(4), 651–678, doi:10.1175/1520-0469(1991)048<0651:OTCOED>2.0.CO;2, 1991.

Hitchcock, P. and Simpson, I. R.: The Downward Influence of Stratospheric Sudden Warmings*, J. Atmos. Sci., 71(10), 3856–3876, doi:10.1175/jas-d-14-0012.1, 2014.

Kunz, A., Pan, L. L., Konopka, P., Kinnison, D. E. and Tilmes, S.: Chemical and dynamical discontinuity at the extratropical tropopause based on START08 and WACCM analyses, J. Geophys. Res. Atmos., 116(24), 1–15, doi:10.1029/2011JD016686, 2011.

Li, F., Newman, P., Pawson, S. and Perlwitz, J.: Effects of Greenhouse Gas Increase and Stratospheric Ozone Depletion on Stratospheric Mean Age of Air in 1960-2010, J. Geophys. Res. Atmos., (2010), 2098–2110, doi:10.1002/2017JD027562, 2018.

Miyazaki, K., Iwasaki, T., Kawatani, Y., Kobayashi, C., Sugawara, S. and Hegglin, M. I.: Inter-comparison of stratospheric mean-meridional circulation and eddy mixing among six reanalysis data sets, Atmos. Chem. Phys., 16(10), 6131–6152, doi:10.5194/acp-16-6131-2016, 2016.

Moore, F. L., Ray, E. A., Rosenlof, K. H., Elkins, J. W., Tans, P., Karion, A. and Sweeney, C.: A Cost-Effective Trace Gas Measurement Program for Long-Term Monitoring of the Stratospheric Circulation, Bull. Am. Meteorol. Soc., 95(1), 147–155, doi:10.1175/BAMS-D-12-00153.1, 2014.

Morgenstern, O., Hegglin, M. I., Rozanov, E., O'Connor, F. M., Abraham, N. L., Akiyoshi, H., Archibald, A. T., Bekki, S., Butchart, N., Chipperfield, M. P., Deushi, M., Dhomse, S. S., Garcia, R. R., Hardiman, S. C., Horowitz, L. W., Jöckel, P., Josse, B., Kinnison, D., Lin, M., Mancini, E., Manyin, M. E., Marchand, M., Marécal, V., Michou, M., Oman, L. D., Pitari, G., Plummer, D. A., Revell, L. E., Saint-Martin, D., Schofield, R., Stenke, A., Stone, K., Sudo, K., Tanaka, T. Y., Tilmes, S., Yamashita, Y., Yoshida, K. and Zeng, G.: Review of the global models used within phase 1 of the Chemistry–Climate Model Initiative (CCMI), Geosci. Model Dev., 10(2), 639–671, doi:10.5194/gmd-10-639-2017, 2017.

Morgenstern, O., Stone, K. A., Schofield, R., Akiyoshi, H., Yamashita, Y., Kinnison, D. E., Garcia, R. R., Sudo, K., Plummer, D. A., Scinocca, J., Oman, L. D., Manyin, M. E., Zeng, G., Rozanov, E., Stenke, A., Revell, L. E., Pitari, G., Mancini, E., Di Genova, G., Visioni, D., Dhomse, S. S. and Chipperfield, M. P.: Ozone sensitivity to varying greenhouse gases and ozone-depleting substances in CCMI-1 simulations, Atmos. Chem. Phys., 18(2), 1091–1114, doi:10.5194/acp-18-1091-2018, 2018.

Neu, J. L. and Plumb, R. A.: Age of air in a "leaky pipe" model of stratospheric transport, J. Geophys. Res. Atmos., 104(D16), 19243–19255, doi:10.1029/1999JD900251, 1999.

Pedatella, N. M., Fuller-Rowell, T., Wang, H., Jin, H., Miyoshi, Y., Fujiwara, H., Shinagawa, H., Liu, H.-L., Sassi, F., Schmidt, H., Matthias, V. and Goncharenko, L.: The neutral dynamics during the 2009 sudden stratosphere warming simulated by different whole atmosphere models, J. Geophys. Res. Sp. Phys., 119(2), 1306–1324, doi:10.1002/2013JA019421, 2014.

Ploeger, F. and Birner, T.: Seasonal and inter-annual variability of lower stratospheric age of air spectra, Atmos. Chem. Phys., 16(15), 10195–10213, doi:10.5194/acp-16-10195-2016, 2016.

Ploeger, F., Abalos, M., Birner, T., Konopka, P., Legras, B., Müller, R. and Riese, M.: Quantifying the effects of mixing and residual circulation on trends of stratospheric mean age of air, Geophys. Res. Lett., 42(6), 2047–2054, doi:10.1002/2014GL062927, 2015a.

Ploeger, F., Riese, M., Haenel, F., Konopka, P., M??ller, R. and Stiller, G.: Variability of stratospheric mean age of air and of the local effects of residual circulation and eddy mixing, J. Geophys. Res. Atmos., 120(2), 716–733, doi:10.1002/2014JD022468, 2015b.

Ploeger, F., Legras, B., Charlesworth, E., Yan, X., Diallo, M., Konopka, P., Birner, T., Tao, M., Engel, A. and Riese, M.: How robust are stratospheric age of air trends from different reanalyses?, Atmos. Chem. Phys. Discuss., 1–32, doi:10.5194/acp-2018-1281, 2019.

Polvani, L. M., Wang, L., Abalos, M., Butchart, N., Chipperfield, M. P., Dameris, M., Deushi, M., Dhomse, S. S., Jöckel, P., Kinnison, D., Michou, M., Morgenstern, O., Oman, L. D., Plummer, D. A. and Stone, K. A.: Large impacts, past and future, of ozone‐depleting substances on Brewer‐Dobson circulation trends: A multi‐model assessment, J. Geophys. Res. Atmos., 2018JD029516, doi:10.1029/2018JD029516, 2019.

Rosenlof, K. H.: Seasonal cycle of the residual mean meridional circulation in the stratosphere, J. Geophys. Res., 100(D3), 5173, doi:10.1029/94JD03122, 1995.

Šácha, P., Eichinger, R., Garny, H., Pišoft, P., Dietmüller, S., de la Torre, L., Plummer, D. A., Jöckel, P., Morgenstern, O., Zeng, G., Butchart, N. and Añel, J. A.: Extratropical Age of Air trends and causative factors in climate projection simulations, Atmos. Chem. Phys. Discuss., (2018), 1–37, doi:10.5194/acp-2018-1310, 2019.

Seviour, W. J. M., Butchart, N. and Hardiman, S. C.: The Brewer-Dobson circulation inferred from ERA-Interim, Q. J. R. Meteorol. Soc., 138(665), 878–888, doi:10.1002/qj.966, 2012.

Smith, A. K., Pedatella, N. M., Marsh, D. R. and Matsuo, T.: On the Dynamical Control of the Mesosphere–Lower Thermosphere by the Lower and Middle Atmosphere, J. Atmos. Sci., 74(3), 933–947, doi:10.1175/JAS-D-16-0226.1, 2017.

SPARC: SPARC CCMVal Report on the Evaluation of Chemistry-Climate Models. V. Eyring, T. Shepherd and D. Waugh (Eds.), SPARC Rep. No. 5, WCRP-30/2010,WMO/TD-No.40, available at www.sparc-climate.org/publications/sp [online] Available from: http://www.sparc-climate.org/publications/sparc-reports/sparc-report-no5/, 2010.

Stiller, G. P., Von Clarmann, T., Haenel, F., Funke, B., Glatthor, N., Grabowski, U., Kellmann, S., Kiefer, M., Linden, A., Lossow, S. and López-Puertas, M.: Observed temporal evolution of global mean age of stratospheric air for the 2002 to 2010 period, Atmos. Chem. Phys., 12(7), 3311–3331, doi:10.5194/acp-12-3311-2012, 2012.

Tweedy, O. V., Limpasuvan, V., Orsolini, Y. J., Smith, A. K., Garcia, R. R., Kinnison, D., Randall, C. E., Kvissel, O. K., Stordal, F., Harvey, V. L. and Chandran, A.: Nighttime secondary ozone layer during major stratospheric sudden warmings in specified-dynamics WACCM, J. Geophys. Res. Atmos., 118(15), 8346–8358, doi:10.1002/jgrd.50651, 2013.

World Meteorological Organization (WMO): Scientific Assessment of Ozone Depletion: 2018, Global Ozone Research and Monitoring Project – Report No. 58, Geneva, Switzerland, 2018., 2018.

Zhang, K., Wan, H., Liu, X., Ghan, S. J., Kooperman, G. J., Ma, P. L., Rasch, P. J., Neubauer, D. and Lohmann, U.: Technical note: On the use of nudging for aerosol-climate model intercomparison studies, Atmos. Chem. Phys., 14(16), 8631–8645, doi:10.5194/acp-14-8631-2014, 2014.