Response to reviewers – Stratospheric ozone change and related climate impacts over 1850–2100 as modelled by the ACCMIP ensemble

We are grateful for the feedback of the three reviewers. We hope their comments and concerns are addressed below. Our responses (i.e. changes and information) follow each reviewer comment in **blue**.

Clarification in the discussion of the results, after the reviewers' comments, has been a major improvement to the revised manuscript. Furthermore, greater caution has been taken when no evidence was available (i.e. limited diagnostics in ACCMIP models), which highlights the need for a more comprehensive output in future multi-model intercomparison activities (MIPs).

Responses to reviewer #1

(a) Main concerns:

Comment: My one significant concern is the loose way in which the authors have broadly applied the term 'GHG concentrations', referring to the degree of radiative forcing or global warming, to explain certain differences in the response of ozone. The term 'GHG concentrations' appears in numerous places through Section 3.2 and is used as a general term to distinguish between the changes projected in RCP2.6 and RCP8.5, the two RCP scenarios investigated here. Yet it is critical when assessing the response of ozone to keep in mind that RCP8.5 is not just a scenario with a large increase in tropospheric radiative forcing by 2100, but it is also the only RCP scenario with a large future increase in methane. The problem is first apparent in the Abstract, Page 25176, Lines 21-24 (and discussed further in the body of the article on Page 25193, lines 16-22) with the following:

'Future TCO changes in the tropics are mainly determined by the upper stratospheric ozone sensitivity to GHG concentrations, due to a large compensation between tropospheric and lower stratospheric column ozone changes in the two RCP scenarios.'

The finding of the ’large compensation between tropospheric and lower stratospheric column ozone changes’ is largely a result of the particular scenarios that have been investigated. The RCP2.6 scenario has weaker global warming and, one assumes, a weaker increase in tropical upwelling associated with the acceleration of the Brewer-Dobson circulation (BDC) along with a correspondingly weaker decrease in lower stratospheric ozone. The RCP8.5 scenario would have a much larger climate change signal which results in a much larger decrease in lower stratospheric ozone due to the increase in tropical upwelling, as can be seen in Figure 5(i). But the compensating increase in tropospheric ozone in RCP8.5 is most certainly largely due to the increase in methane specified for RCP8.5 at 2100 and less the result of an increase in GHG concentrations in general. If one were to investigate RCP6, would one find that future TCO changes are determined by the upper stratospheric sensitivity to GHG concentrations due to a cancelling of changes between the lower stratosphere and the troposphere? As written, the statement is correct – for the two particular RCP scenarios analyzed here – but more care must be taken that the proper caveats are applied on more general statements.

Response: We did not mean to imply general GHG concentration and we understand that in the RCP8.5 emission scenario methane plays a key role in the tropospheric column ozone, as it approximately doubles its concentration by 2100 relative to present-day (2000) (Lamarque et al., 2013). Also, stratospheric influx (i.e. stratospheric-tropospheric exchange) is the other key driver of the increase in tropospheric ozone in the RCP8.5 (Young et al., 2013). Furthermore, we are aware that the ‘large compensation between tropospheric and lower stratospheric column ozone changes’ may be a result of different drivers and processes acting in the RCPs emission scenarios investigated here. Therefore, we agree that proper caveats need to be applied on general statements. The sentence (Page 25193, lines 16-22) have now been rewritten explaining the caveats pointed out above:

"Future TCO changes in the tropics are mainly determined by the upper stratospheric ozone sensitivity to GHG concentrations, due to a large compensation between tropospheric and lower stratospheric column ozone changes in the **RCP2.6 and RCP8.5 emission scenarios. Notice that tropospheric column ozone in the RCP8.5 2100 time slice is largely the result of future increase in methane (Lamarque et al., 2013)**".

We think, however, that there is no need to address the 'origin' of the tropospheric column ozone change in the RCP8.5 in the abstract (Page 25176, Lines 21-24), as here we do not address the likely 'causes' of the finding.

Comment: I have a related concern about the discussion of mid-latitude ozone changes discussed in Lines 25-29 of Page 25193, where the response of ozone in the lower stratosphere is '...positively correlated to GHG concentrations'. By only analyzing RCP2.6 and RCP8.5 it is not possible to separate the effects of a general climate change-associated response (here, an increase in transport of ozone from the tropics to mid-latitudes due to an acceleration of the BDC) from the photochemical effects associated with the increased methane particular to RCP8.5. Randeniya et al. (2002), Fleming et al. (2011) and Reader et al. (2013) have all shown the importance of methane to ozone in the extra-tropical lower stratosphere and it is likely that part of the lower stratospheric response of ozone is also due to the increase in methane and not purely a result of climate change. I would urge the authors to be careful about generalizing a response to 'GHG concentrations' when only analysing two RCP scenarios.

Response: Previous studies have been shown that changes in methane burden can affect ozone production as a result of its oxidation in the extratropical lowermost stratosphere (i.e. methane burden positively correlated with ozone production). Therefore, the sentence (Page 25193, Lines 25-29) has now been rewritten citing this work:

"In **contrast to the tropics**, the mid-latitudes lower stratospheric ozone is positively correlated to GHG concentrations (Fig. 5b and d) **mainly** due to the influx of relatively "rich" ozone air from lower latitudes (e.g. WMO, 2011) **from a strengthened BDC**. **Additionally, the increase in methane emissions in the RCP8.5 scenario results in chemically-driven increases in ozone in this region (e.g. Randeniya et al., 2002; Reader et al., 2013)**."

(b) Minor comments:

Page 25178, Lines 19-22. In the discussion of the effects of CO2-cooling on upper stratospheric ozone you should not ignore the straight-forward effects of temperature on oxygen chemistry. A large fraction of the response of upper stratospheric ozone to cooling has been attributed to the temperature dependence of O+O2+M -> O3 +M.

Response: Indeed, there is a temperature dependence of oxygen chemistry (i.e. cooler temperatures lead to both, faster production and slower destruction). The sentence has been rewritten:

"… slowing gas-phase ozone loss processes (e.g. reduced NOx abundances; reduced HOx-catalysed ozone loss**; and enhanced net oxygen chemistry**) resulting in ozone increases, particularly in the middle-upper stratosphere and high latitudes (e.g. Haigh and Pyle, 1982; Randeniya et al., 2002; Rosenfield et al., 2002; Jonsson et al., 2004)."

Page 25179, Line 13. Plural 'columns' in 'Recent past stratospheric columns ozone...'

Response: Fixed. Thanks.

Page 25181, Line 2. The authors introduce the idea that the SAM trend is not solely the result of ozone depletion by stating 'as it opposes the effect of increasing GHG concentrations.' without introducing the idea and referencing work that suggests the SAM is affected by both ozone trends and GHGs. A complete discussion of SAM trends does require mentioning the effect of GHGs.

Response: This suggestion will help a reader have a better overview of the topic. A sentence has been included to introduce the effect of increasing GHGs concentration on the SAM:

"… **Furthermore, some modelling studies have projected a poleward shift (i.e. positive change) in the SAM due to future increases in GHGs (e.g. Fyfe et al., 1999; Marshall et al., 2004)** …"

Page 25184, Lines 17-20. It is stated that two models (CESM-CAM superfast and MIROC-CHEM) submitted time-slice simulations yet introduced ODSs into these simulations as an emission. Since in 1980 the atmospheric abundance of ODSs was very far from being in steady-state with emissions, evidenced by how rapidly the tropospheric concentration of these species was increasing, how did these models ensure that the atmospheric concentration of ODSs is realistic for 1980 conditions? Since the authors of the present manuscript are not responsible for how these simulations were setup, perhaps a fairer question is to ask if the halogen loading in these simulations is realistic for 1980 conditions.

Response: The ODSs concentrations were not saved in ACCMIP models. However, we have calculated and looked for significant trends in stratospheric ozone (i.e. globally and >65S between 100-10 hPa) for those years that form part of the Hist 1980 time-slice in CESM-CAM-superfast and MIROC-CHEM. From this analysis, we find no significant trends in these models and time slice, even though the ODSs were specified as emissions. However, this is a good point and we have now included a sentence to note this issue:

"Note that no significant trends are found for stratospheric ozone in those years that form part of the Hist 1980 time slice for the latter models, even though ODSs were specified as emissions (i.e. any trends in ODS concentration in the stratosphere due to transport timescales do not significantly affect ozone concentrations)."

Page 25187, Lines 20-26. I think there should be discussion of the complication of calculating 1980-2000 trends from timeslice experiments for 1980 and 2000 conditions. My motivation here is that in 1980 the tropospheric concentrations of many ODSs is rapidly increasing. Given the 3 to 5 year lag for transport into the stratosphere, running for 10 years with constant 1980 tropospheric concentrations will produce stratospheric halogen loading that is more like that found around 1985, assuming a few years additional simulation were discarded to allow the model to properly spin up. I imagine the effect on trends is not large, but there should be discussion of this consideration.

Response: We are aware that there are particular issues with time slice simulations, including the concerns noted by the reviewer. As noted in our previous response, we agree that this is an important issue to highlight, yet we also note that there are no significant trends in ozone depletion through the time slices of the models that specified emissions. Clearly it would have been useful for these ACCMIP models to have archived their ODS concentrations. We hope that these lessons learned about model diagnostics will inform future MIPs (CCMI is already saving far more diagnostics that ACCMIP and CCMVal2, in order to address questions related to both the troposphere and stratosphere).

Page 25188, Lines 23-24. Here it is stated that Tier 1.4 of the BDBP ozone database is based on a regression model to the original observations. Are all terms of the regression conserved when deriving the trends that are used for the comparisons shown in Figure 2?

Response: The authors do not alter the BDBP ozone database (Bodeker et al., 2013). Therefore, the trends calculated in the manuscript conserve all terms. According to the authors of the Binary Database of Profiles (BDBP), Tier 0 refers to the monthly means inferred from ozone measurements. Tier 1.4 is calculated by fitting a regression model (i.e. including a number of functions such as ENSO, QBO, solar cycle and so forth) to the former tier. Note that regression coefficients for each of the above terms are fully consistent along the time-series. Therefore, all terms of the regression are consistent and conserved along the time-series of the BDBP data set and in the trends calculated in this study.

Page 25188, Line 28. It might help the reader transition from the previous introduction of Figure 2 to the reference to Figure 1b by stating that you are discussing total column ozone in the tropics.

Response: We agree that this is not clear. A sentence has been included stating that we are discussing total column ozone in the tropics:

"In the tropics (Fig. 1b), **TCO in all data sets agrees fairly well with observations**."

Page 25189, Lines 10-13. The statement 'ACCMIP models fail to represent observed ozone depletion occurring in the lower and middle stratosphere region, which may be linked to a poor representation of the HOx and NOx catalytic loss cycles (e.g. Lary, 1997; Nedoluha et al., 2015)' seems to be a significant bit of speculation. One should certainly be suspicious of the HOx chemistry, as it is dominant in the lower stratosphere, but can one rule out problems with trends in transport? And the reference to Nedoluha et al. (2015) seems out of place as they discuss trends in ozone around 10 hPa, where the models are not doing too badly. Further, the Nedoluha et al. findings of NOx effects on ozone trends was explained as being due to trends, secular or transient it is not clear, in N2O transport where N2O is the source gas for reactive nitrogen.

Response: We agree that transport may have contributed and that it cannot be ruled out. Furthermore, Nedoluha et al. (2015) is misplaced and has been removed. The sentence has now been rewritten:

"ACCMIP models fail to represent observed ozone depletion occurring in the lower and middle stratosphere region, which may be linked to a poor representation of the HOx catalytic loss cycle **and upwelling in this region** (e.g. Lary, 1997; Randel et al., 2007)."

Page 25190, Lines 4-5. The use of the term 'low biased' is not as clear as it should be in 'the ACCMIP multi-model mean is low biased compared to the BDBP data (Fig 2e).' The models generally underestimate the large negative trends in the BDBP data and the use of 'low bias' could mean that the models are more negative.

Response: The terms 'low biased' and 'high biased' alone were not clear. This has been clarified in the revised manuscript for both terms. For example, the sentence above commented has been rewritten to clarify the term 'low biased':

"…, the ACCMIP multi-model mean is **underestimating larger negative trends** compared to the BDBP data (Fig. 2e)."

Page 25190, Lines 22-25. Here the authors state 'ACCMIP models show fairly good agreement with BDBP Tier 1.4 decadal trends at various altitude regions, except around 70–30 hPa, likely linked to NOx ozone loss chemistry associated to stronger temperature trends than observed (see Sect. 5).' Assuming that much of the reactive nitrogen is sequestered in PSCs or has been removed by sedimentation of PSCs during at least the early part of SON shown in Figure 1f, the authors should explain more completely their thinking behind how NOx ozone loss chemistry can explain the underestimated ozone loss.

Response: Within uncertainty, the ACCMIP multi-model mean shows more negative temperature trends than observations, but weaker ozone depletion than the BDBP data (Tier 0 observation compilation and Tier 1.4 regression model) – true for most of the ACCMIP CHEM and NOCHEM models. However, the reviewer is correct to point out that NOx chemistry changes are unlikely to be the main driver of the discrepancy given that we expect the nitrogen to have been removed from the stratosphere through PSC chemistry at this time. Given the lack of available diagnostics, it is difficult to isolate the mechanisms behind the discrepancy in the ACCMIP simulations, so we leave this open for future study (e.g. CCMI, where will be more output saved). We also note that this temperature/ozone trend relationship has been reported before in our rewritten text:

"ACCMIP models show fairly good agreement with BDBP Tier 1.4 decadal trends at various altitude regions, except around 70–30 hPa, **which is also the region where the modelled temperature trends are more negative than observed** (see Sect. 5). **This is consistent with previous analysis which suggested that models potentially simulate too strong negative trend for a given ozone depletion (e.g. Young et al., 2011) and this discrepancy warrants further investigation in future model intercomparison studies, where there is more model output available.**"

Page 25198, Line15. I'll admit to always feeling on thin ice when discussing statistics, so if I am mistaken please accept my apologies. Why did the authors choose to use a paired sample Student's t-test when testing the significance of changes in the SAM index across the different experiments? My understanding is that a paired sample requires matched pairs within each of the populations being compared. Here, since the individual years in each of the experiments are completely independent isn't the independent samples t-test the appropriate one?

Response: The authors agree that the independent Student's t-test is more appropriate for this analysis. We have recalculated it and the results have not changed. The sentence has now been rewritten:

“By using a **two independent samples** Student’s t test, we find that SAM index changes between Hist 2000 and 2100 relative to Hist 1850, are significant for the RCP2.6 at the 5 % level, although is not significant for the RCP8.5.”

Page 25199, Lines 11-13. In discussing the large spread of model projected changes for RCP8.5 at 2100 the authors state that the spread is ’...likely linked to sensitivity of ozone to future GHG emissions uncertainty (i.e. various direct and indirect processes affecting ozone amounts in the troposphere and the stratosphere).’ It is a really minor point, but GHG emissions uncertainty is usually referenced in discussing the spread across scenarios. Here, all models used the same RCP8.5 specified emissions and the spread in model responses arises as models respond differently to these large forcings. The text in the parentheses, I think is a good description of the cause but it is not the same as future GHG emissions uncertainty.

Response: Authors agree that we did not phrase this very clearly. We meant ’...likely linked to **uncertainties due to** sensitivity of ozone to future GHG emissions (i.e. various direct and indirect processes affecting ozone amounts in the troposphere and the stratosphere).’ The sentence has now been re-written as above stated.

Page 25201, Lines 26-28. The effect of prescribing ozone or having interactive ozone and the role of zonal asymmetry was discussed in general terms on pages 25199-25200, but here a direct link between the different dynamical responses of the CHEM and NOCHEM models is attributed to the specification of zonally symmetric ozone in the NOCHEM models. It has been discussed earlier in the article that the SPARC/IGAC CMIP5 ozone database underestimated Antarctic ozone depletion and that the NOCHEM ACCMIP models show less ozone depletion than the CHEM models. Given the many different factors that may have affected the comparison of the CHEM and NOCHEM models, can the authors conclude that the use of zonally symmetric ozone is the cause of the differences they find?

Response: We meant that zonally symmetric ozone may be one of the possible causes in the NOCHEM models to underestimate or overestimate SH climate response. It is also known that the SPARC/IGAC ozone dataset is at the low end compared to observations in this region (e.g. Young et al., 2014). The discussion on pages 25199-25200 has been revisited. Furthermore, the sentence has now been rewritten:

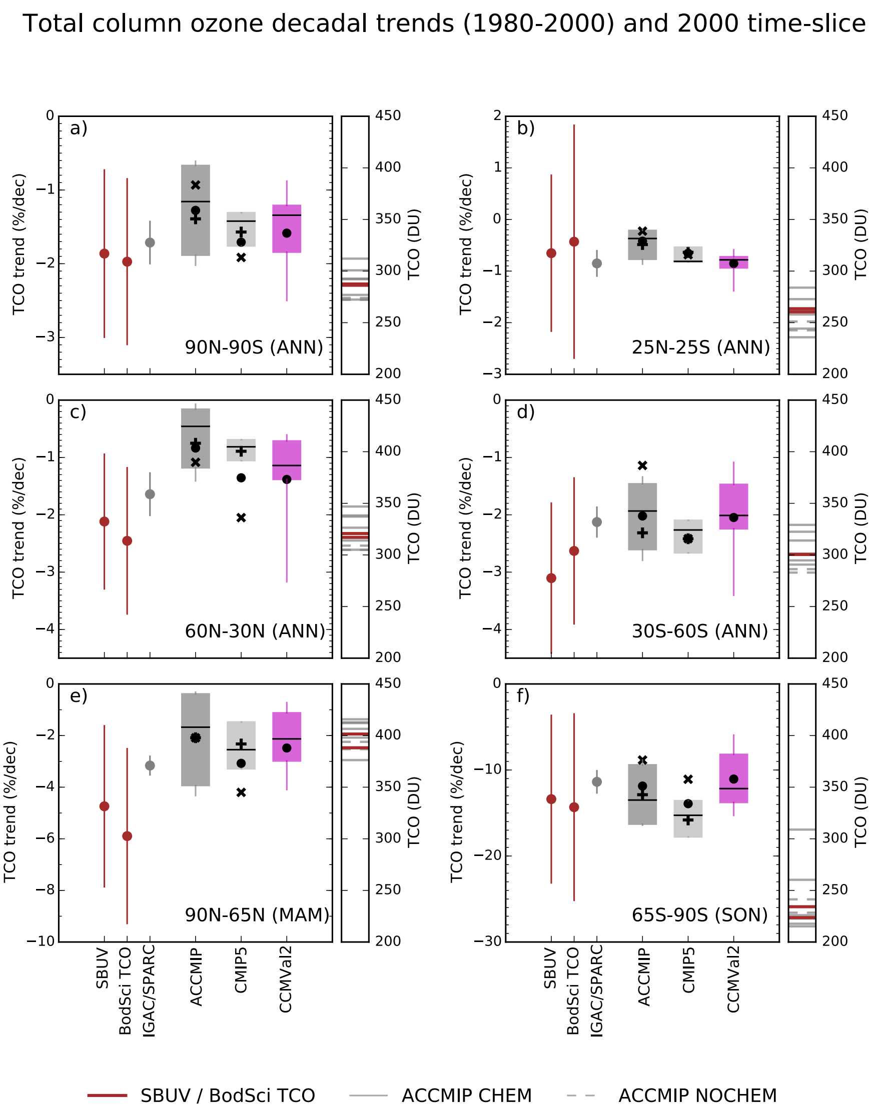
"This highlights the importance of the ozone database used to drive models on the climate response. For example, Young et al. (2014) found 20-100 % larger tropospheric climate responses in this region and season with a climate model driven by the BDBP data set compared to the SPARC/IGAC data set used in NOCHEM models here."

Responses to reviewer #2

(a) Summary:

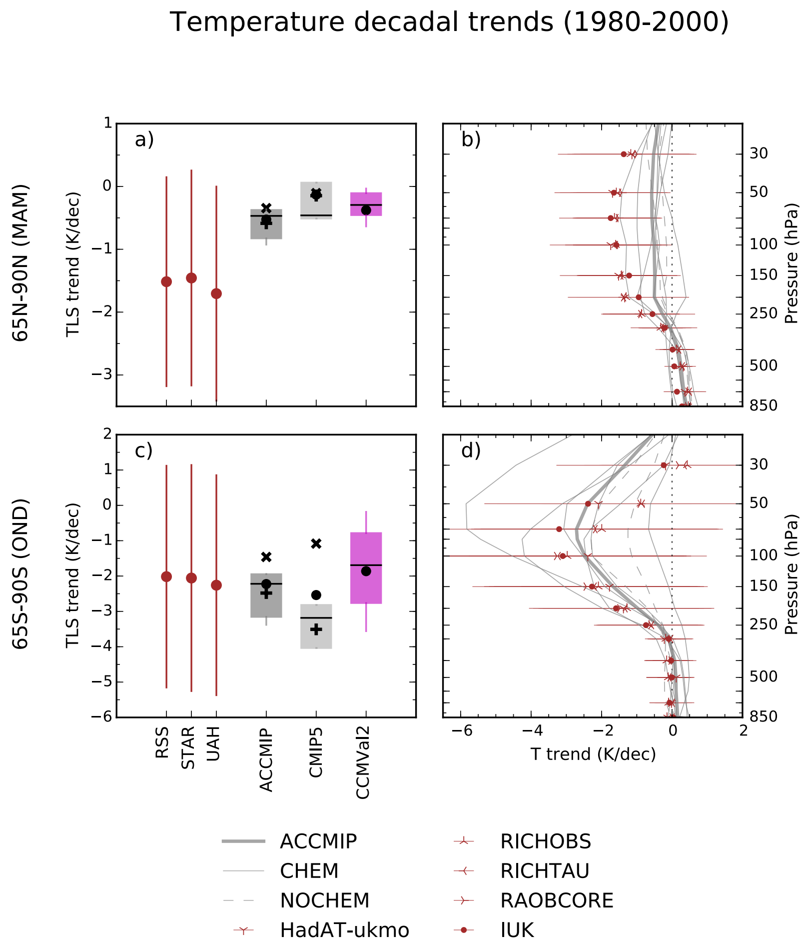
The paper's main strength is the large amount of model information ingested into the study. The authors use data from three recent multi-model intercomparison projects. The analyses themselves are straightforward. I don't have any major issues with the diagnostics except that an analysis of total column ozone should exclude models with prescribed stratospheric ozone. Also the analysis of stratospheric temperature trends should distinguish between model with and without stratospheric ozone chemistry. The latter group presumably would show less inter-model variability than the former.

Response: In both analyses, total column ozone and temperature trends between 1980 and 2000, CHEM and NOCHEM models (for ACCMIP and CMIP5) are distinguished by 'plus' and 'cross' symbols, respectively. This allows a number of comparisons between both, different type of models and multi-model intercomparison projects or MIPs (e.g. some of them addressed in the manuscript). We have chosen to include all the models in each of the multi-model intercomparison activities participating in this study (i.e. ACCMIP, CMIP5 and CCMVal2) in the boxplots to give a clearer overall picture of the multi-model ensembles. Nevertheless, we provide below these analyses including only CHEM models (in the stratosphere).



Total column ozone decadal trends (1980-2000) and 2000 time-slice
a)
90N-90S (ANN)
b)
25N-25S (ANN)
c)
60N-30N (ANN)
d)
30S-60S (ANN)
e)
90N-65N (MAM)
f)
65S-90S (SON)
TCO trend (%/dec)
TCO (DU)
SBUV
BodSci TCO
IGAC/SPARC
ACCMIP
CMIP5
CCMVal2
SBUV / BodSci TCO
ACCMIP CHEM
ACCMIP NOCHEM

1



I'm also not sure about the paper's final conclusion that analyses of the last decade have comprehensively demonstrated that there are benefits in interactively coupling ozone and climate. These points amount to a minor revision. The paper provides the type of summary information which is likely to be of use in the 6th Assessment Report of IPCC, the Tropospheric Ozone Assessment (TOAR) and the 2018 WMO Ozone Assessment. Thus the paper needs to progress to ACP after my detailed concerns, given below, are addressed. The language is generally adequate; in a few places, there are minor grammatical or stylistic issues that further proof-reading by a native speaker would help address.

Response: To support the paper's final conclusion on the benefits of including processes interactively in models (i.e. fully resolved stratosphere), we include a couple of references of key papers on this regard. The sentence has now been rewritten:

“These results and work over the last decade have shown that changes in stratospheric ozone are tightly coupled to the climate **(e.g. SPARC-CCMVal, 2010; Nowack et al., 2015)**, **supporting the idea** of including these processes interactively in models.”

Minor concerns have been addressed (see below) and the language has been checked in the revised manuscript.

(b) Minor comments:

P25183L19f. Are you sure HadGEM2 uses a look-up table approach which differs from UM-CAM? The models are of the same heritage. In the troposphere, HadGEM2 uses the same look-up table approach as UM-CAM (i.e. ozone column does not enter the calculation). In the stratosphere, possibly a different type of look-up table was used which depends on ozone column, pressure, temperature, and solar zenith angle. But this does not directly influence tropospheric rates. Please confirm with the HadGEM2 PIs that this is correct.

Response: Many thanks for the comment. HadGEM2 uses a look-up table similar to UM-CAM (i.e. ozone column does not enter the calculation). The paragraph has now been re-written:

“…The simplest scheme is for **HadGEM2 and** UM-CAM, where the photolysis rates are derived from a look-up table as a function of time, latitude and altitude only, and using a climatological cloud and ozone fields (i.e. the rates are the same for all simulations)…”

P25188L11. This is a strange formulation. The IGAC/SPARC dataset is based on observations?

Response: This sentence was not clear. We meant that the SPARC/IGAC dataset (i.e. used in NOCHEM models) shows better agreement with observations (i.e. SBUV and BodSci TCO) than CHEM models outside extratropical SH regions (i.e. although within observational uncertainty, it has been shown that SPARC/IGAC has a “conservative” ozone depletion in the SH). The sentence has now been re-written:

“However, outside extratropical SH regions, IGAC/SPARC **ozone data set (i.e. used to drive the majority of ACCMIP and CMIP5 NOCHEM models)** tends to show better agreement with observations than CHEM models.”

P25189L11. How many ACCMIP models actually had comprehensive stratospheric chemistry? My impression was that most used prescribed or simplified ozone in the stratosphere.

Response: This study includes 8 out of 16 ACCMIP models. These are 6 CHEM models and 2 NOCHEM models (as shown in **P25221 Table 1**). Although, CESM-CAM-superfast is included in the CHEM group, it has simplified stratospheric chemistry.

P25189L16. You give a trend as "-1.64–2.45±1.2%/dec”. The notation is strange. I suggest to replace this with a central estimate followed by the lower and upper (2.5% and 97.5% confidence) bounds, or for symmetric bounds the central estimate ± its uncertainty range at the 95% confidence interval. This also applies to the notations used in various other places in the text.

Response: The authors agree that the notation is not clear and these have been replaced by the central estimate and its uncertainty range at the 95% confidence interval:

P25189L16. “… -2.29 ± 1.2 % dec-1”

P25190L1. “… -5.32 ± 3.3 % dec-1”

P25190L17. “… -13.86 ± 10.4 % dec-1”

P25189L17. It’s no surprise that the NOCHEM models produce better agreement with obs than the CHEM models – aren’t they constrained with observations? Is this a fair comparison? The unsuspecting reader might conclude that adding interactive chemistry to a model is counterproductive. . .

Response: We would have thought so, however, globally (90N-90S ANN) and in regions with strong ozone depletion (i.e. 30S-60S ANN and >65S SON), CHEM models show better agreement with observations than NOCHEM models. It may be interesting to know where prescribing ozone fields (i.e. with the IGAC/SPARC ozone data set) would be appropriate. However, this breaks the connection between ozone and climate, which could lead to inconsistencies.

P25190L2. This is more than could be said for CCMVal2. I suspect that this is again because quite a few ACCMIP models prescribe ozone, hence this is as expected. You should exclude from the TCO analyses models that use prescribed stratospheric ozone.

Response: The authors agree that it may not be fair comparing CCMVal2 multi-model mean with ACCMIP and CMIP5 multi-model means (i.e. as these include some models with prescribed ozone fields). Therefore, in the revised manuscript we have changed this and comparisons between these three data sets are now done for ACCMIP CHEM, CMIP5 CHEM and CCMVal2 multi-model means.

P25190L17. The notation for the trend is strange, see above.

Response: The notation has been change to: "… -13.86 ± 10.4 % dec-1" (see above).

P25191L6. ". . . tropospheric ozone columns" (word order)

Response: Fixed thanks.

P25191L25. "the magnitude depending on region"

Response: Fixed thanks.

P25195L12. Morgenstern et al. (2010) did not consider the NAM. You mean Morgenstern et al. (2010), Anthropogenic forcing of the Northern Annular Mode in CCMVal-2 models, JGR, 115, D00M03, doi:10.1029/2009JD013347.

Response: The reference has been amended.

P25195L18f. Slightly strange sentence structure.

Response: The sentence was not completely clear. It has now been rewritten for further clarification:

"… observational estimates based on Microwave Sounding Unit (MSU) **retrievals by** the Remote Sensing Systems (RSS – version 3.3)…"

P25196L6. "simulate" (plural)

Response: Fixed. Thanks.

P25197L12. The word "temperature" is missing.

Response: The word "temperature" has now been added:

"… The slight **temperature** increase"

P25198L6. The MMM is within the uncertainty estimates of the observations, so at the 95% confidence interval there is no disagreement.

Response: We agree that at the 95% confidence interval there is no disagreement, however, we just want to point out that the ACCMIP multi-model mean is somewhat weaker (i.e. less positive trend) than the observation estimates.

P25198L17. ". . .but are not significant for the RCP8.5. . ."

Response: Fixed. Thanks.

P25200L18. Replace ",which" for "and"

Response: Fixed. Thanks.

P25200L25ff. As alluded above, this analysis only makes sense if you restrict it to models that interactively calculate stratospheric ozone.

Response: As commented above, CHEM and NOCHEM models (for ACCMIP and CMIP5) are distinguished by 'plus' and 'cross' symbols, respectively. Boxplots include all models participating here for each MIP to give a clearer overall picture of the multi-model ensembles. However, for a fair comparison between MIPs, the CCMVal2 multi-model mean is only discussed against ACCMIP and CMIP5 CHEM models.

P25202L9. I don't think you have established "additivity" or linearity here. How about "While in the recent past both ozone depletion and increasing GHGs have favoured a strengthening of the SAM during summer, under projected ozone recovery they will drive the SAM into opposite directions" or so.

Response: The sentence has now been re-written (thanks):

"**While in the recent past both ozone depletion and increasing GHGs have favoured a strengthening of the SAM during summer, under projected ozone recovery they will drive the SAM into opposite directions.**"

P25202L20. I don't think the leading effect is due to the representation of ozone (interactive or prescribed). My impression is that the main problem is that if ozone is prescribed, it can be inconsistent with the applied GHG and ODS forcing, which can skew the climate change signal due to changing GHGs. Your paper does not address the pure question of exactly what the differences are just due to the representation of ozone. Son et al., JGR, 2010, assess whether interactive (CCMVal-2) models behave differently from CMIP3 models, for some climate indices. The answer is, no, for these indices and models.

Response: The authors agree with the point made above. The sentence has now been re-written:

"We have demonstrated both its key role in the present and future SH climate **and the importance of how it is represented in climate models.**"

Responses to reviewer #3

(a) General comments:

This study would be helpful to better understand the uncertainty of SPARC ozone data and the projected changes of the stratospheric and tropospheric ozone in a warm climate. Inter-model comparison, i.e., ACCMIP, CCMVal-2, and CMIP5, could be also applicable to other modeling project such as ongoing CCMI project. However, the present study is missing the detailed explanations. In many places, the authors argued that such differences or discrepancies are "likely" caused by photolysis and stratospheric circulations without presenting any supporting evidences.

Response: The authors agree that more caution should be taken when no supporting evidence is provided. We would like to probe the 'causes' and drivers of the different findings presented in the manuscript. However, either because of limited output available (i.e. diagnostics) or because of the lack of sensitivity simulations for ACCMIP, this has not been possible in many cases. The above reviewer comment points to the need for more detailed diagnostics (and experiments) and it is hoped that CCMI will produce more in that direction. Nevertheless, we have rephrased the text in a number of places to acknowledge the reviewer concern.

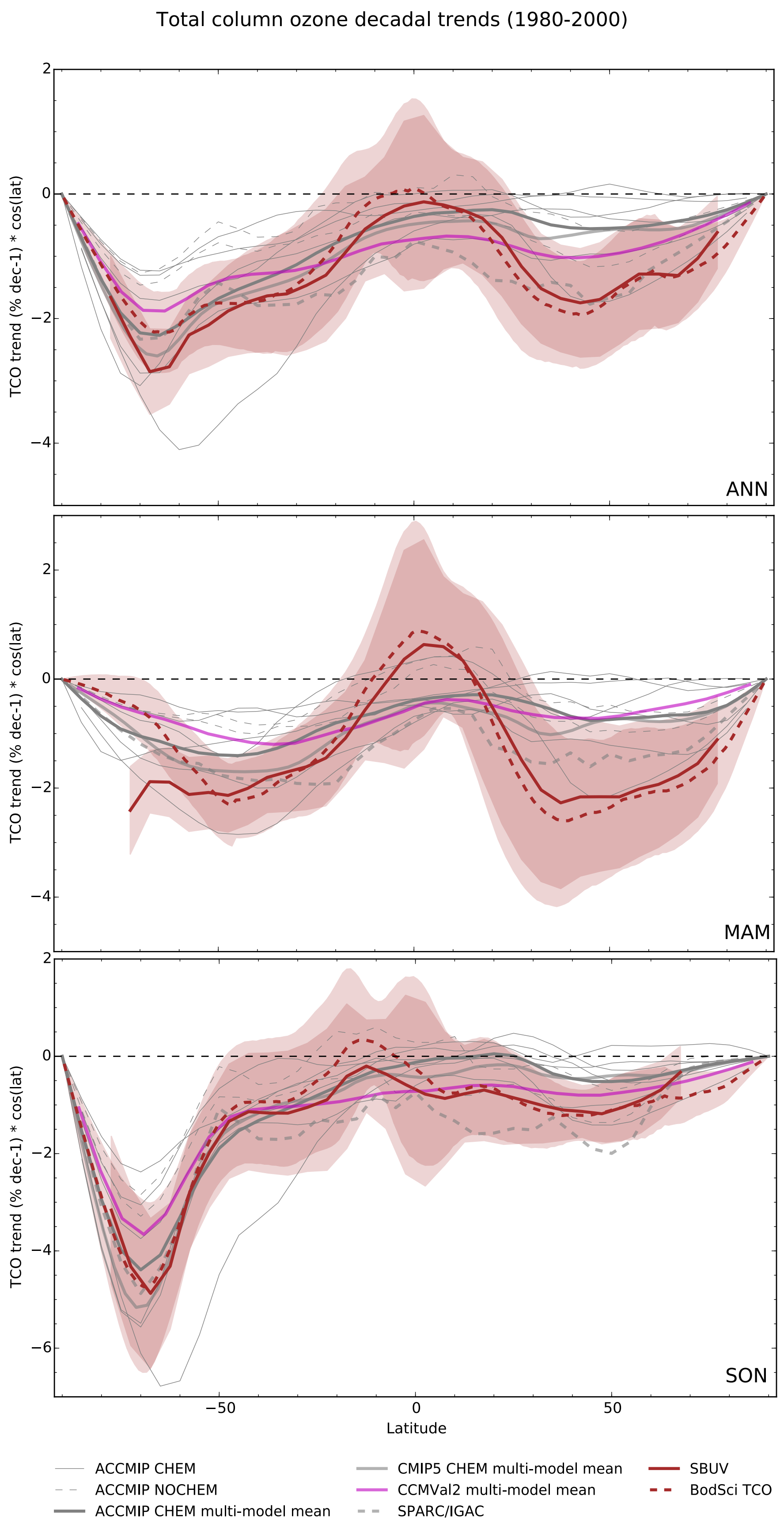
I understand that the main purpose of this study is to evaluate the ACCMIP simulations. However, this paper could become more exciting paper if additional analyses and figures that can support their arguments are presented. For example, intensification of the BDC and its differences among the models are repeatedly stated. But, no figures are shown for the equatorial upwelling or BDC. Since the computation of w* is not difficult, it could be evaluated at least for the ACCMIP models. The results could be compared to the tropical upwelling in the CCMVal-2 and CMIP5 high-top models which are presented in Butchart et al. (2010JCLI) and Charlton-Perez et al. (2013JGR).

Response: Indeed, calculating the BDC would have helped better understand the analyses made in this study. This is a really good suggestion, but the output specified for the ACCMIP simulations does not allow this calculation. We discuss below some extra analyses made for the total column ozone trends between 1980 and 2000 as a function of latitude and its absolute values for year 2000 (i.e. similar to Fig. 1 in the manuscript).

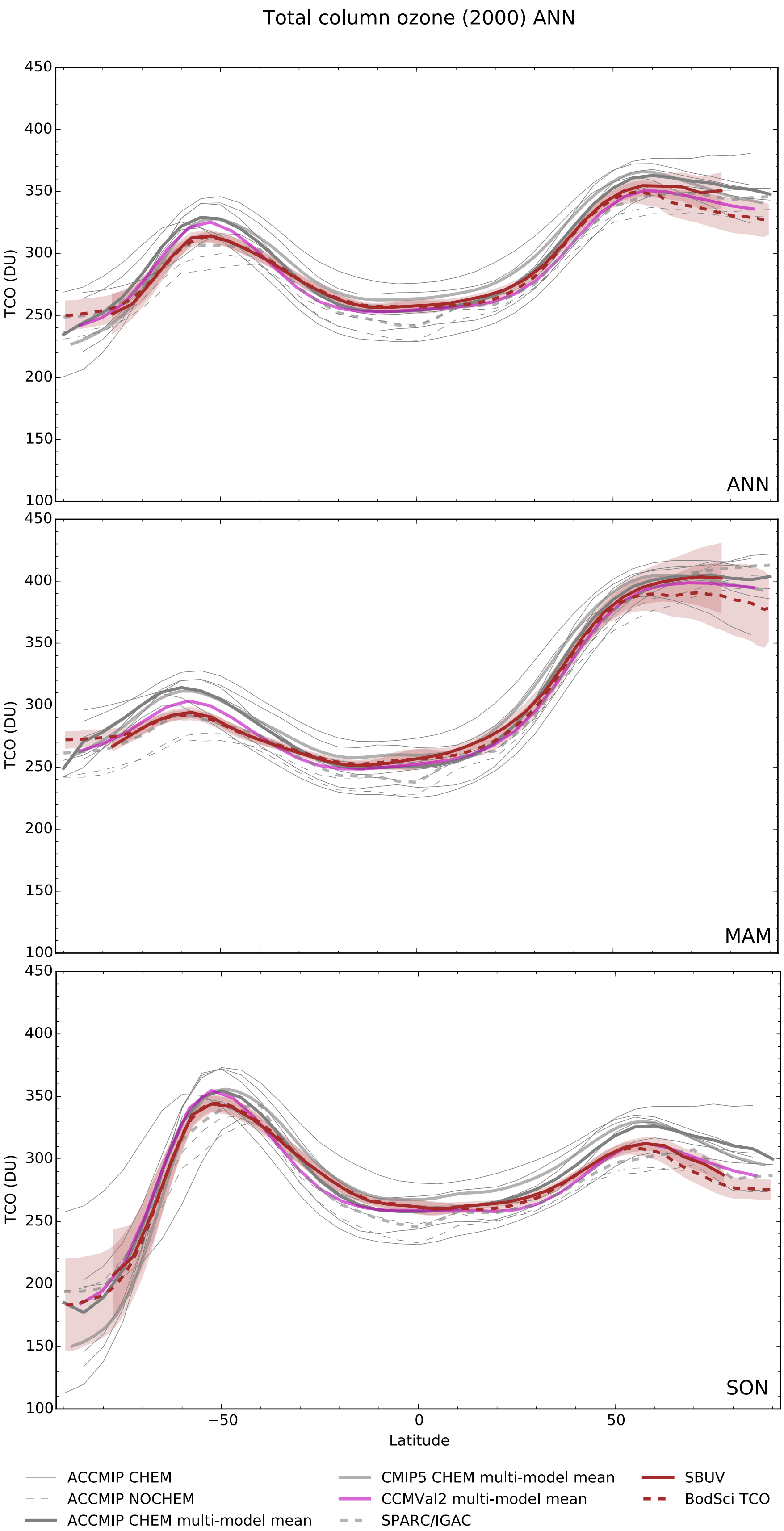
(b) Specific comments:

1. Evaluation as a function of latitude in all figures, latitudinally-averaged quantities are presented. But, I think the latitudinal profile of annual-mean TCO or the latitude-month plot of monthly-mean TCO is much more useful. Such figure would be especially important to evaluate the extent of the polar vortex and its trend. I suggest authors to evaluate the climatological TCO (1980-2000) as a function of latitude (instead of the one presented in a small box in Fig. 1). Likewise, authors can present long-term trends of TCO, stratospheric O3 and tropospheric O3 as a function of latitude and season.

Response: We are grateful for the above suggestion and present some extra analyses. The first figure shows total column ozone trends between 1980 and 2000 as a function of latitude (i.e. similar as in Fig.1 in the manuscript). The second figure presents the absolute values of total column ozone for year 2000 as a function of latitude (i.e. again similar to Fig.1 small boxes). After a careful consideration, we believe that Fig.1 in the manuscript summarises fairly well the information provided by these two extra figures (below).



Total column ozone decadal trends (1980-2000)
TCO trend (% dec-1) * cos(lat)
ANN
MAM
SON
Latitude
ACCMIP CHEM
ACCMIP NOCHEM
ACCMIP CHEM multi-model mean
CMIP5 CHEM multi-model mean
CCMVal2 multi-model mean
SPARC/IGAC
SBUV
BodSci TCO



Total column ozone (2000) ANN
TCO (DU)
ANN
MAM
SON
Latitude
ACCMIP CHEM
ACCMIP NOCHEM
ACCMIP CHEM multi-model mean
CMIP5 CHEM multi-model mean
CCMVal2 multi-model mean
SPARC/IGAC
SBUV
BodSci TCO

1

2

2. BDC To explain the biases of the tropical and NH midlatitude O3 concentration, authors mentioned the importance of the BDC. Such influence could be simply illustrated by a scatter plot of tropical O3 and NH midlatitude O3 for all ACCMIP models. For example, if the modeled BDC is stronger than observation, negative relationship between the two would be stronger. Based on Fig. 5, I suspect that 50 hPa in the tropics (decreased O3 by the intensified upwelling) and 150 hPa in the extratropics (increased O3 by the enhanced downwelling) would be reasonable choice for the scatter plot. This scatter plot would also reveal the relationship between the mean biases and trends of tropical O3 and those of extratropical O3.

Response: Inferring the BDC is a nice idea, but we do not feel that the results would be conclusive given that we are dealing with time slice simulations. A comparison of out-of-phase ozone relationships (or even temperature) might indicate something about differing strengths of the BDC between models, but is not an absolute measure we could compare to observations with confidence. Again, we await the more detailed diagnostics that will become available in future intercomparisons. Nevertheless, we have rewritten our text slightly removing the reference to the BDC strength in this region:

"… though simulates weaker ozone depletion in the lower stratosphere **compared to observations (although not the Tier 1.4 regression model), which may be** associated with **the weaker than observed ozone depletion over the Arctic compared to observations (see below)**."

3. Interactive and zonally asymmetric stratospheric ozone. It is argued that "eliminating zonal asymmetry may lead to a poor representation of stratospheric and tropospheric climate trends in the SH". This point is repeated raised in the manuscript. However, no evidence is presented. In fact, Gerber et al. (2013BAMS) documented that, based on the inter-comparison between the CMIP5 (prescribed ozone) and CCMVal-2 models (interactive ozone), the response of the SH circulation is NOT dramatically sensitive to the interactive or zonally asymmetric ozone. Such sensitivity might be true in a single model (e.g., Waugh et al. 2009b). However, its impact is likely within the uncertainty in the multi-model framework (Gerber et al.). This issue should be clearly re-investigated in the revised manuscript.

Response: As commented above, in this case we are not able to investigate further the 'causes' of the differences between climate models with prescribed ozone and those with fully resolved chemistry. Here, we are just noting that eliminating zonal asymmetry may result in a poorer representation of the climate response both, in the stratosphere and the troposphere. However, there may be a number of reasons explaining the differences between CHEM and NOCHEM models (i.e. the ozone database used to drive NOCHEM models and/or how it was implemented). Therefore, we have revisited the manuscript on this regard, due to this and some comments of the other reviewers (i.e. **Page 25201, Lines 26-28; Page 25202, Line 20**).

4. Others - Fig. 1b versus Fig. 2b: Fig. 1b shows comparable TCO trends to observations. However, a large difference is found in Fig. 2b especially in the tropical UTLS. Are they consistent? BTW, it would be helpful if zero line is included in Fig. 2.

Response: Fig. 1b and Fig. 2b are consistent. Notice where ACCMIP models are low biased (i.e. small negative trends; UTLS) compared to observational estimates, the contribution to the total column ozone is relatively small. A **'zero' line has been included** now in both Fig. 2 and Fig. 6, thanks.

- Fig. 4: In the introduction, definition of tropopause used in the study is extensively discussed. However, in Fig. 4, tropopause is simply set to 150 hPa. Is there any reason to choose 150 hPa? This pressure level is certainly the upper troposphere rather than the lower stratosphere.

Response: The notation used in this analysis was not clear. The tropopause definition used in this paper applies for both the upper and lower stratospheric columns ozone. The later refers to the stratospheric column ozone below 15 hPa. The notation has been changed:

"Figure 4. As Fig. 3, but for the upper stratosphere (10–1 hPa), lower stratosphere (**>15 hPa**) and tropospheric columns ozone (DU) in the tropics."

- Too many references: I am not sure whether that many references, over 6 pages, are necessary for the present paper.

Response: We have removed some of the less relevant references as requested, particularly from the introduction section.

- Typos: This paper is technically well written. But there are still several typos. I believe authors can easily correct them when revising the manuscript.

Response: The new manuscript has been checked for typos after all comments and revisions.

We thank again the three reviewers for their comments.

References:

Bodeker, G. E., Hassler, B., Young, P. J., and Portmann, R. W.: A vertically resolved, global, gap-free ozone database for assessing or constraining global climate model simulations, Earth Syst. Sci. Data, 5, 31-43, 10.5194/essd-5-31-2013, 2013.

Fyfe, J. C., Boer, G. J., and Flato, G. M.: The Arctic and Antarctic oscillations and their projected changes under global warming, Geophys. Res. Lett., 26, 1601-1604, 10.1029/1999GL900317, 1999.

Haigh, J. D., and Pyle, J. A.: Ozone perturbation experiments in a two-dimensional circulation model, Quart. J. Roy. Meteor. Soc., 108, 551-574, 10.1002/qj.49710845705, 1982.

Jonsson, A. I., de Grandpré, J., Fomichev, V. I., McConnell, J. C., and Beagley, S. R.: Doubled CO2-induced cooling in the middle atmosphere: Photochemical analysis of the ozone radiative feedback, J. Geophys. Res., 109, D24103, 10.1029/2004JD005093, 2004.

Lamarque, J. F., Shindell, D. T., Josse, B., Young, P. J., Cionni, I., Eyring, V., Bergmann, D., Cameron-Smith, P., Collins, W. J., Doherty, R., Dalsoren, S., Faluvegi, G., Folberth, G., Ghan, S. J., Horowitz, L. W., Lee, Y. H., MacKenzie, I. A., Nagashima, T., Naik, V., Plummer, D., Righi, M., Rumbold, S. T., Schulz, M., Skeie, R. B., Stevenson, D. S., Strode, S., Sudo, K., Szopa, S., Voulgarakis, A., and Zeng, G.: The Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP): overview and description of models, simulations and climate diagnostics, Geosci. Model Dev., 6, 179-206, 10.5194/gmd-6-179-2013, 2013.

Lary, D. J.: Catalytic destruction of stratospheric ozone, J. Geophys. Res., 102, 21515-21526, 1997.

Marshall, G. J., Stott, P. A., Turner, J., Connolley, W. M., King, J. C., and Lachlan-Cope, T. A.: Causes of exceptional atmospheric circulation changes in the Southern Hemisphere, Geophys. Res. Lett., 31, L14205, 10.1029/2004GL019952, 2004.

Nowack, P. J., Luke Abraham, N., Maycock, A. C., Braesicke, P., Gregory, J. M., Joshi, M. M., Osprey, A., and Pyle, J. A.: A large ozone-circulation feedback and its implications for global warming assessments, Nature Clim. Change, 5, 41-45, doi:10.1038/nclimate2451, 2015.

Randel, W. J., Park, M., Wu, F., and Livesey, N.: A Large Annual Cycle in Ozone above the Tropical Tropopause Linked to the Brewer–Dobson Circulation, J. Atmos. Sci., 64, 4479-4488, 10.1175/2007JAS2409.1, 2007.

Randeniya, L. K., Vohralik, P. F., and Plumb, I. C.: Stratospheric ozone depletion at northern mid latitudes in the 21st century: The importance of future concentrations of greenhouse gases nitrous oxide and methane, Geophys. Res. Lett., 29, 10-11-10-14, 10.1029/2001GL014295, 2002.

Reader, M. C., Plummer, D. A., Scinocca, J. F., and Shepherd, T. G.: Contributions to twentieth century total column ozone change from halocarbons, tropospheric ozone precursors, and climate change, Geophys. Res. Lett., 40, 6276-6281, 10.1002/2013GL057776, 2013.

Rosenfield, J. E., Douglass, A. R., and Considine, D. B.: The impact of increasing carbon dioxide on ozone recovery, J. Geophys. Res., 107, ACH 7-1-ACH 7-9, 10.1029/2001JD000824, 2002.

SPARC CCMVal: SPARC CCMVal Report on the Evaluation of Chemistry-Climate Models, edited by: Eyring, V., Shepherd, T. G., and Waugh, D. W., SPARC Report No. 5, WCRP- 132, WMO/TD-No. 1526, 2010.

Young, P. J., Rosenlof, K. H., Solomon, S., Sherwood, S. C., Fu, Q., and Lamarque, J.-F.: Changes in Stratospheric Temperatures and Their Implications for Changes in the Brewer-Dobson Circulation, 1979-2005, J. Clim., 25, 1759-1772, 10.1175/2011jcli4048.1, 2011.

Young, P. J., Archibald, A. T., Bowman, K. W., Lamarque, J. F., Naik, V., Stevenson, D. S., Tilmes, S., Voulgarakis, A., Wild, O., Bergmann, D., Cameron-Smith, P., Cionni, I., Collins, W. J., Dalsoren, S. B., Doherty, R. M., Eyring, V., Faluvegi, G., Horowitz, L. W., Josse, B., Lee, Y. H., MacKenzie, I. A., Nagashima, T., Plummer, D. A., Righi, M., Rumbold, S. T., Skeie, R. B., Shindell, D. T., Strode, S. A., Sudo, K., Szopa, S., and Zeng, G.: Pre-industrial to end 21st century projections of tropospheric ozone from the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP), Atmos. Chem. Phys., 13, 2063-2090, 10.5194/acp-13-2063-2013, 2013.

Young, P. J., Davis, S. M., Hassler, B., Solomon, S., and Rosenlof, K. H.: Modeling the climate impact of Southern Hemisphere ozone depletion: The importance of the ozone data set, Geophys. Res. Lett., 41(24), 9033-9039, doi:10.1002/2014GL061738, 2014.