

Response to Anonymous Referee #2

We would like to thank the reviewer for their detailed and comprehensive review. The comments and recommendations in this review have lead to a great improvement of this paper.

Referee comments and statements are in bold. Author replies are in italics. Updated figures are at the end (supplied for reference).

SUMMARY

This manuscript presents an evaluation of the Australian Community Climate and Earth System Simulator-Chemistry Climate Model in the Southern Hemisphere, focusing on the model's representation of Antarctic ozone depletion, stratospheric temperatures, the polar vortex, as well as past changes in surface winds. To this end, the model data are compared to both observations (ERA-I, Bodeker Scientific, ozonesondes) and multi-model datasets (CCMVal2, CMIP5). While such model-measurement comparisons are needed to improve the current generation of chemistry-climate models and to gain confidence in their performance, this study lacks new or significant results that would be of interest to the wider community. Also, the comparisons are not performed according to current best practices (see detailed comments below). I hence cannot recommend the manuscript for publication in the current form. I also suggest the authors to consider submitting this manuscript to the GMD journal instead of ACP, since it would fit that journal's scope and readership better.

We understand the concerns regarding the chosen journal. When discussing with coauthors where to submit this paper, we considered both ACP and GMD. We feel that the papers scope is closer to ACP, as it includes some nice results on how the model is incorrectly simulating Antarctic ozone depletion, and its links to ClO. Also, there has been very little new ACCESS-CCM model development since previously published papers in GMD (e.g. Morgenstern et al., 2009 (10.5194/gmd-2-43-2009); O'Conner et al., 2014 (10.5194/gmd-7-41-2014)), which also steered us towards ACP.

GENERAL COMMENTS

1) The current discussion of the results of the model is very descriptive and does not yield much insight on model performance or improvements over a previous (or parent) model. While it is good to see that the ACCESS-CCM does not completely lie outside of the CCMVal2 range, this was not to be expected anyhow since models with over- all rather weak performance metrics are included in the CCMVal2 database. A more insightful comparison could be achieved by highlighting/adding also the differences between the NIWA-UKCA model that was used in CCMVal2 and forms the basis of the new ACCESS-CCM. Commenting on

improvements in the performance in comparison to the NIWA-UKCSA model and reasons for possible improvements would be more valuable to the wider community than the current results.

Thank you for this comment. We think that including direct comparison to ACCESS-CCM predecessors is a great suggestion. However, NIWA-UKCSA did not contribute to CCMVal-2, but is actually directly comparable to our model. Therefore, we have made the comparison with our most direct predecessors in the time-series plots: UМУKCA-UCAM and UМУKCA-METO.

For relevant changes made to the text, please see comment reply corresponding to P19166.

2) The bigger problem I see with this manuscript however is that the authors try to interpret results from model-measurement comparisons that do not live up to the standards to current best practices, e.g. comparing the same time periods or accounting for potential sampling biases (as explained below in the case of CIO satellite measurements). Due to these deficiencies the conclusions of the paper (or explanations for processes/mechanisms behind the differences between the model and the observations) cannot be trusted either.

We have updated all climatology comparisons to only include 2005-2010 (the longest period available for all comparisons) averages to alleviate any concerns about not comparing same times.

Regarding the CIO, please see comment reply corresponding to P19175 L 25 to P19176 L4.

SPECIFIC COMMENTS(Note: page/line numbers indicated are from printer-friendly version of manuscript)

P19163 L12-14 Please specify why these model intercomparisons are expected to be of any help, e.g. ‘which focus on process-oriented evaluation of model performance’.

Thank you. We have included the suggested sentence

p. 19163, line 13. Appended: “...chemistry climate modelling projects, ...” with “...chemistry climate modelling projects, which focus on process-oriented evaluation of model performance, ...”

P19163 L20 change to ‘halting and reversing the’ since otherwise the full success of the Montreal Protocol is not acknowledged.

Thank you. This has been changed

p. 19163, line 20. Changed: “halting the build-up of halogens” to: “halting and reversing the build-up of halogens...”

P19163 L21-23 My understanding is that the community moved away from defining ozone recovery according to the three phases defined in the 2006 WMO ozone assessment, since the second phase does not explicitly account for the impacts climate change can have on ozone. E.g. increased stratospheric temperatures or increases in the Brewer Dobson circulation can alter ozone distributions and lead to an apparent ozone recovery, which may not be attributable to the decline in stratospheric chlorine and bromine. Also, knowledge has advanced since Dameris et al. (2014). A study by Shepherd et al. (Nature Geoscience, 2014, doi:10.1038/ngeo2155) now disentangles the effects of climate change and decreasing EESC and shows that ozone recovery is in fact already taking place.

Thank you for this comment. We agree, and have updated the following

p. 19163, line 21. Replaced the sentence: “This marks the first phase of ozone recovery, with the second phase being when ozone is consistently increasing.” With “Other recent studies have noted a detection in ozone recovery (e.g. Shepherd et al. (2015), deLaat et al. (2015)).”

P19164 L2-3 This impact does relate to the Earth system, in particular humans and ecosystems, but the way it is placed in the text implies that UV changes may affect climate. I suggest adding some clarification.

Thank you, we agree and have clarified with the following

p. 19164, line 2. Changed: “Another obvious surface impact is an increase in ultra violet (UV) radiation reaching the surface...” to “Another obvious surface impact, important for ecosystems, is an increase in ultra violet (UV) radiation reaching the surface...”

P19166 It is not clear whether the model underwent specific improvements since Morgenstern et al. (2009).

Morgenstern et al. (2009) describes the stratospheric chemistry only. With only very few minor changes made since this paper: Addition of VSLS tracers (CH₂Br₂, CHBr₃) which would add ~5pptv of bromine to the stratosphere, update to nitrogen advection settings.

The major chemistry related changes since UMUKCA model iterations are: FAST-JX is used instead of FAST-J2 and tropospheric chemistry including tropospheric isoprene.

A paragraph has been included to discuss differences with ACCESS-CCMs predecessors.

p. 19166, line 23: Included the paragraph: “The ACCESS-CCM model is a direct successor to the UMUKCA-UCAM and UMUKCA- METO CCMs that contributed to CCMVal-2, the second interaction of CCMVal. A number of

advancements to the model where made since. Regarding the stratospheric chemistry scheme. The UMUKCA models and ACCESS-CCM both follow Morgenstern et al. (2009), with only minor adjustments made to include the halogenated very short lived substances: CH₂Br₂, ChBr₃, update the advection of total nitrogen. Other more major changes to the chemistry in ACCESS-CCM are the introduction of FASTJX instead of FAST-J2 (Bian and Prather, 2002), the introduction of tropospheric chemistry, approximately doubling the number species and reactions included only in the stratospheric scheme (O'Connor et al., 2014), and the addition of isoprene for tropospheric chemistry. In addition, the UMUKCA models used HadGEM1 as the background climate model, with the major updates in HadGEM3 being to the convection, cloud and boundary layer schemes, among others, described in Hewitt et al. (2011)."

Also, we have included the UMUKCA-UCAM and UMUKCA-METO models in the time-series analysis plots.

P19167 L8 Why do you choose to follow RCP 6.0 after 2005? An explanation should be given, since this is a puzzling difference to the use of the first simulation.

The standard REF-C2 simulation as defined in the CCMI project specifies the use of RCP 6.0. This is different from the REF-C1 simulation, which uses RCP 8.5 after 2005, as this scenario matches closest with observations. Therefore, we have updated the following text to clarify

p. 19166, line 27. Added the sentence: "RCP 8.5 was chosen as this scenario best represents the observations between 2005–2010"

p. 19167, line 8. Added the sentence: "RCP 6.0 was chosen following the CCMI REF-C2 specifications (Eyring et al., 2013b)."

P19167 L17 If it is not prudent you should not do it. Please rephrase to something like 'a limitation to the comparison is given by the absence of winter-time observations'.

Thanks, you are correct. We have rephrased.

p. 19167, line 17. Changed: "It is important to note that it may not be prudent to directly compare Antarctic wintertime observations from this dataset to model data." To "It is important to note that a limitation of this comparison is the shortage of wintertime observations."

P19168/9 All descriptions of observations (3.1, 3.5, and 3.6) need statements about measurement quality and stability. It is not clear what you mean by having taken into account 'all data quality control considerations' in 3.6.

MLS measurements are supplied with different data screening parameters. It

is recommended that the user apply them to the data before being used scientifically. We have updated the following text

p. 19169, line 7. Changed the sentence: "...comparison of the model data with the MLS ClO measurements has taken into account all data quality control considerations." To: "...data quality control considerations, such as, precision, quality, status flag and convergence (see Livesey et al. 2011).

An updated description about the pros and cons of the Bodeker scientific database has been included. Please see reply to comment P19170 L7-15.

Quality control statements about ozonesondes have been introduced.

p. 19168, line 24. "The accuracy of ECC ozonesondes has been reported to range between 5–10% when following a standardized procedure (Smit et al. 2007).

P19169 L8-12 It is not clear to me why and how you account for the a priori of the measurement. Please improve this description.

We have changed the following sentence to avoid any confusion

p. 19169, line 8. Changed: "...consistently, this is done by adding the averaging kernel convolved model and a priori difference to the a priori (Livesey et al., 2011)." to "...consistently, this is done following Eq. 2 in Livesey et al. (2011), where the model data is modified to represent what MLS would observe. This is done by taking the difference between the model and a priori profiles, multiplying them with the averaging kernels, and adding the product to the a priori.

P19170 L7-15 The discussion of the figure seems limited. In order to be more valuable to the scientific community, it should also include a discussion of potential limitations of the Bodeker Scientific TCO database. For example, the comparisons by Hassler et al. (ACP, 2013, doi:10.5194/acp-13-5533-2013) of different TCO databases indicate that the Bodeker Scientific TCO database may be low-biased in the tropics and at high latitudes.

Thank you, we agree that a discussion of the potential limitations is a good idea. However, Hassler et al. (2013) compares vertically resolved ozone databases, and the Bodeker TCO database was not sourced from these vertical ozone databases. We have updated the description to make the known advantages and disadvantages of this dataset clear.

p. 19167, line 15. Updated the sentence: "This database is assimilated from satellite observations and spans the period from 1979–2012, where offsets between datasets have been accounted for using Dobson and Brewer ground-based observations." To "This database is assimilated from satellite observations and spans the period from 1979--2012, where dataset offsets

and drifts have been accounted for using Dobson and Brewer ground-based observations. This has the advantage of including long-term Dobson and Brewer measurement stability.”

p. 19167, line 16. Added in the sentence: “However, it is important to note that the version of the dataset used includes interpolation. Therefore, a limitation of this comparison is the shortage of wintertime observations. This...”

p. 19170, line 15. Added the sentence: “The differences between REF-C1 and observations at high southern latitudes during austral winter are likely less accurate due to the limited number of observations available at this time.”

P19173 L1 It is not clear how you determine the largest differences. Do you take the maximum difference anywhere along the profile even if it were to be in the troposphere? If so, I wouldn't see the value of having the table without indication of the altitude these numbers pertain to.

Thank you. We agree that the table is a little misleading. What we actually did was take the difference from the maximum values of each profile. Therefore to make this clearer, we have removed the table, and included the differences as a separate line in Figures 4 and 5 in the paper. Any numbers associated with the differences as described in the text have been updated accordingly. Please see Figures 3 and 4 at the bottom of the page for updated plots.

P19175 L 25 to P19176 L4 I don't agree that this is a fair comparison. ClO has a very strong diurnal cycle, especially at altitudes below 10 hPa and with night-time values that often come close to satellite instruments' detection limits. The differences can be expected to be larger in winter than in summer and vary with height, depending on the availability of sunlight. The following discussion (L5-23) of potential model shortcomings seems therefore too hypothetical. The comparison should be repeated for daytime values only (for both model and Aura-MLS) to allow for a fair model-measurement comparison. If you have already found that the results do not depend on taking into account the diurnal cycle, then you will need to show this in the paper or provide references that argue for the validity of the approach. PS: L14-16, Or maybe rather the inability of the authors to make a valid comparison?

Thank you for this comment. We understand the concern regarding ClO. We took measures to make sure that we provided a fair comparison, as described in the original text. For example, stating that this method only allows for a qualitative comparison of the vertical locations of ClO chemistry. However, we agree that a daytime comparison only would be much better, and would also further avoid readers mistaking the comparison for a quantitative one.

Therefore, we have re-produced the plot for only near-coincident times from MLS and model ClO, corresponding to daylight values. We have added in the following sentences in the text to state this.

p. 19175, line 21. Added the sentence: “Only 3 pm values from MLS are used in the average. The REF-C1 averages were produced using instantaneous 3 hourly output, with the closest coincident time to 3 pm used, corresponding to approximately 2 pm at Davis. 3 pm values were chosen as CIO has a strong diurnal cycle. This ensures the model averages represent the CIO observations.”

p. 19176, line 5–23. As we have changed the figure. We have updated the description paragraph.

P19176 L5-25 See previous comment. This may or may not be true.

Please see previous reply.

P19178 L7 I am not convinced that there is a direct link between the differences in the SAM trend and the ozone depletion in the two simulations. Strong ozone depletion after 1980 is seen in REF-C1, but there is no equivalent change in the SAM, instead the SAM stays more or less flat after 1980 in this simulation.

We agree that the large amount of year-to-year variability in the SAM plots will make it hard to directly attribute the trends in this analysis, and we thank the reviewer for taking the time to look into this in detail. There is a leveling off of the SAM index after 1990 in the REF-C1 simulation. However, this coincides with the leveling off of ozone depletion in the same simulation. Over 1980–1990, both the SAM and the total column ozone show large increasing and decreasing trends respectively. This link is what we would expect to see due to the ozone’s known influence on tropospheric circulation.

P19180 L20 You mentioned earlier that the heat-flux comparison showed discrepancies between model and ERA-interim. Maybe a too weak heatflux led to the too cold temperatures in the Antarctic middle stratosphere, which in turn may be the reason for a too strong ozone depletion, and not the other way around? I don’t think your evaluations allow for a conclusion of this chicken-and-egg problem.

Thank you, you are correct. There is not enough information to make this conclusion. Therefore, we have updated the sentence.

p. 19180, line 20. “This also induces a significant cold bias in the stratosphere during spring at the altitudes of ozone depletion in the model.” To “This is also accompanied by a significant cold bias in the stratosphere during spring at the altitudes of ozone depletion in the model.”

P19180 L23-25 See comments above on the validity of your model-measurement comparison, I hence don’t agree that you have attributed the differences in ozone to deficiencies in the representation of CIO in the model.

Please see reply to comment on regarding P19175 L 25 to P19176 L4.

P19181 L6-9 This seems to contradict your earlier statement (P19180 L12-14) that the ozone vertical profile at Melbourne shows very good agreement between ozonesondes and model during all seasons. The problem may be that the Bodeker Scientific TCO database indeed has a low bias as also indicated in Hassler et al. (2013)? In other words, there may be an inconsistency between the TCO and ozonesonde observations you use for your comparisons? Or did you mean inside the polar processing regions?

Thank you for picking up on this. Both Bodeker TCO and ozonesondes both show an ACCESS-CCM excess. Therefore, we do not think there is a disparity in the results here, but the statements need to be clarified.

p. 19180, line 12. Changed “Model-simulated seasonal averaged vertical profiles of ozone and temperature compared to Southern Hemisphere ozonesondes show very good agreement in ozone vertical distribution, concentration and seasonal variation for Melbourne.” To “Model-simulated seasonal averaged vertical profiles of ozone and temperature compared to Southern Hemisphere ozonesondes show very good agreement in ozone vertical distribution, concentration and seasonal variation for Melbourne, with only a small excess ozone bias in ACCESS-CCM.”

P19190 Figure 4 Why did you not make an apple-to-apple comparison using the common time period 2003-2010? I understand you do not want to include 2002 due to the ozone hole splitting event that year. It seems however dangerous to include years with different EESC loadings, the way it is done currently.

We chose the mismatched time periods to ensure a common climatology length of 10 years. We believe that a 10-year climatology would wash out any large resulting biases. However, we understand your concerns, and have therefore updated all climatology comparisons to use the 2005–2010 period (longest common period between all datasets).

P19194 Figure 6 Same as for previous figure, comparison should be made over the same time period so to avoid potential sampling biases resulting from trends in the ClO species.

Please see previous comment reply.

TECHNICAL COMMENTS

Title:

I expected a much broader evaluation of the performance of this model than what is offered. Please specify. E.g. suggest to fit in something like ‘Evaluation of Southern Hemisphere chemistry-climate processes in the ...’. Main problem really is the humongous name of the model under

evaluation, but I assume its name cannot be changed anymore.

Thank you. We understand your point regarding the broad title. Yes, unfortunately the name cannot be changed any more. We are considering changing the title to: "Evaluation of the ACCESS chemistry climate model for the Southern Hemisphere". However, unfortunately, this includes an acronym in the title.

P19163 L10&12 Needs some references (see WMO, 2014 and references therein)

Thank you, we have added appropriate references.

P19167 Section 3 Observational datasets You list here both observational and model datasets you are comparing to, so this title is not adequate. Suggest changing title or moving the model data used for comparisons into a new section.

Thank you for spotting this. We have update the section title to the following

p. 19167, line 9. Changed: "Observational datasets" to "Observational and model datasets"

P19168 L3 I don't think that you can evaluate the performance of your model with earlier model data, since these may be wrong too. You can at best compare them to each other to test improvements or consistency.

This is a good point. We have updated the following text to clarify this,

p. 19168, line 3. Substituted "to evaluate" for "to compare".

P19190 Figure 2 Why are the CCMVal2 data limited to 1965-2000, while the data should be available from 1960-2005?

You are correct: the data are available and used from 1960—2005. The lines in Fig. 2 have undergone a 10-year running mean. Therefore, the first 5 and last 5 years where removed from those lines.

P19172 L25 Please provide references that provide the theoretical backing for this approach.

We have done this under the assumption of normal statistics, and that a single ozonesonde sounding approximates a daily average. We understand that this may not give direct quantitative results. And have rephrased the sentence to clarify this. We have also updated the number used to 7.5. Which is a better representation.

p. 19172, line 25. Changed "The ozonesonde standard deviations are divided by sqrt(7.5) as we have presumed an average of one sounding per week.

With the assumption of normal statistics, this will approximate the standard deviation of a monthly average, consistent with the model data used.” To “The ozonesonde standard deviations are divided by $\sqrt{7.5}$ for visualisation purposes. We have presumed an average of one sounding per week, therefore, with the assumption of normal statistics, this will approximate the standard deviation of a monthly average, consistent with the model data used.

P19175 L7 ‘radiatively active gas’ → ‘radiatively active gases’ P19181 L4 ‘possible’ → ‘possibly’

Fixed, thank you.

P19182 L3 where/what is the CCMI web portal?

Thank you, we have added in the link to the web portal.

Updated Figures

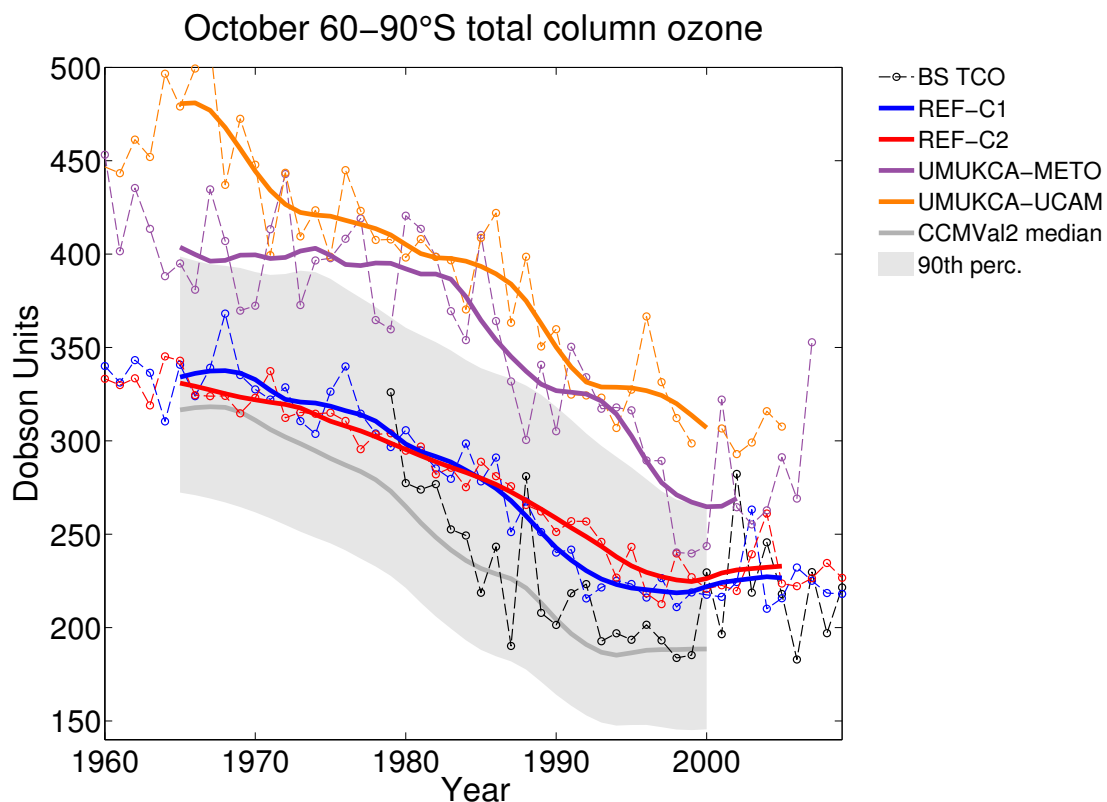


Figure 1. Update to figure 2.

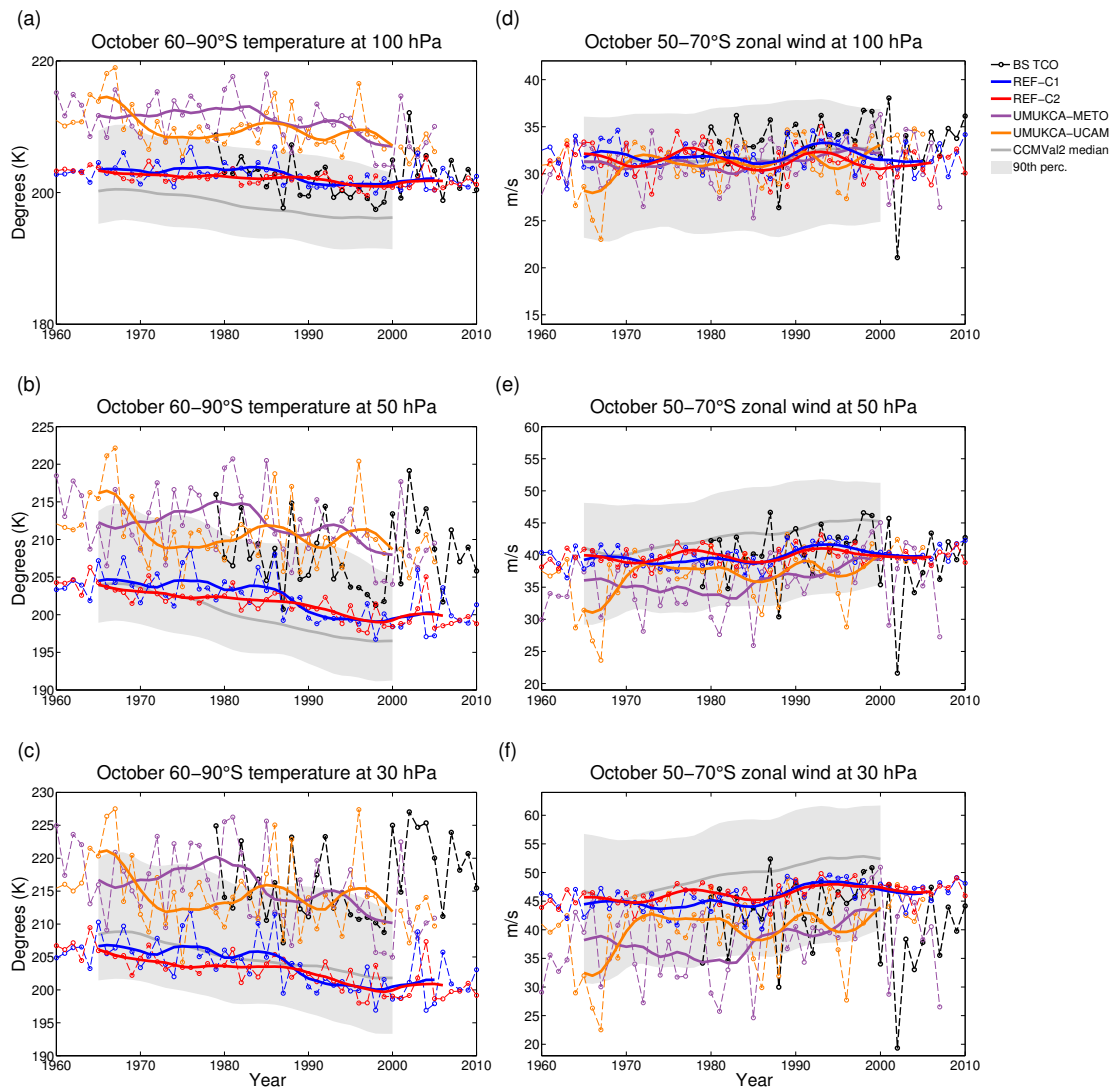


Figure 2. Update to figure 3.

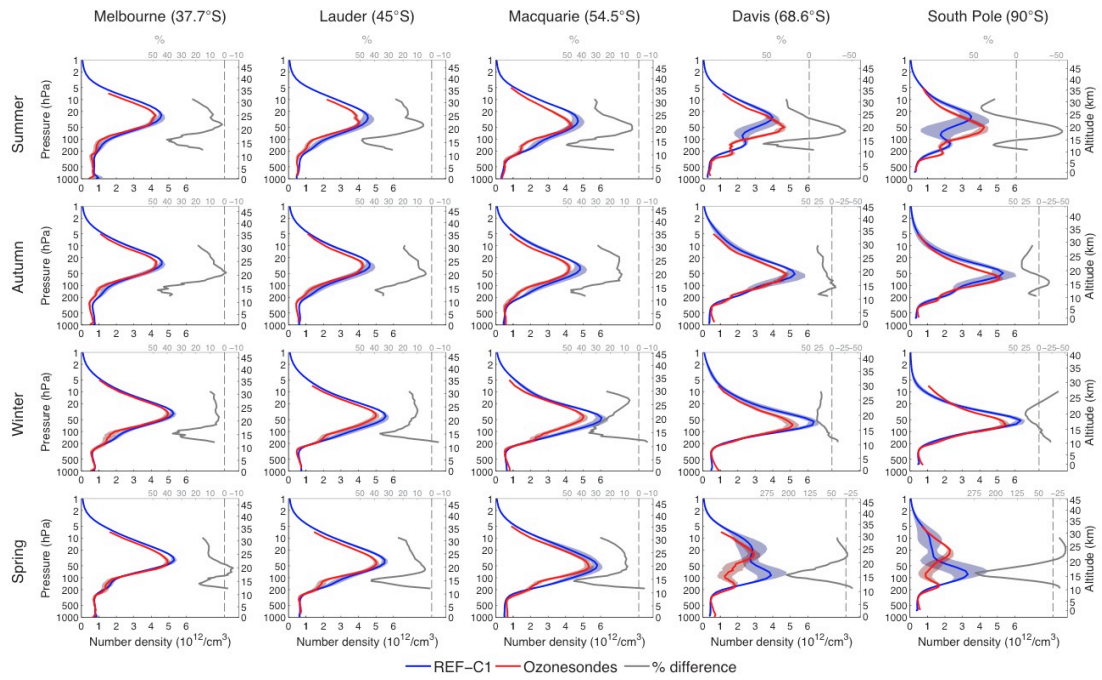


Figure 3. Update to Figure 4 in manuscript

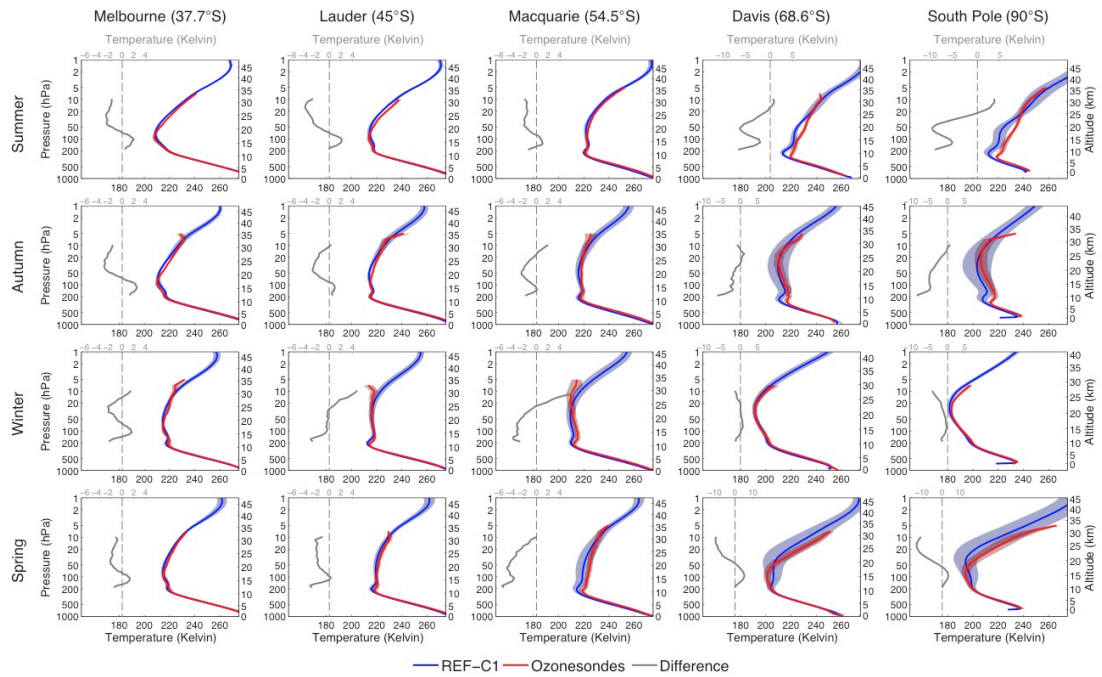


Figure 4. Update to Figure 5 in manuscript.

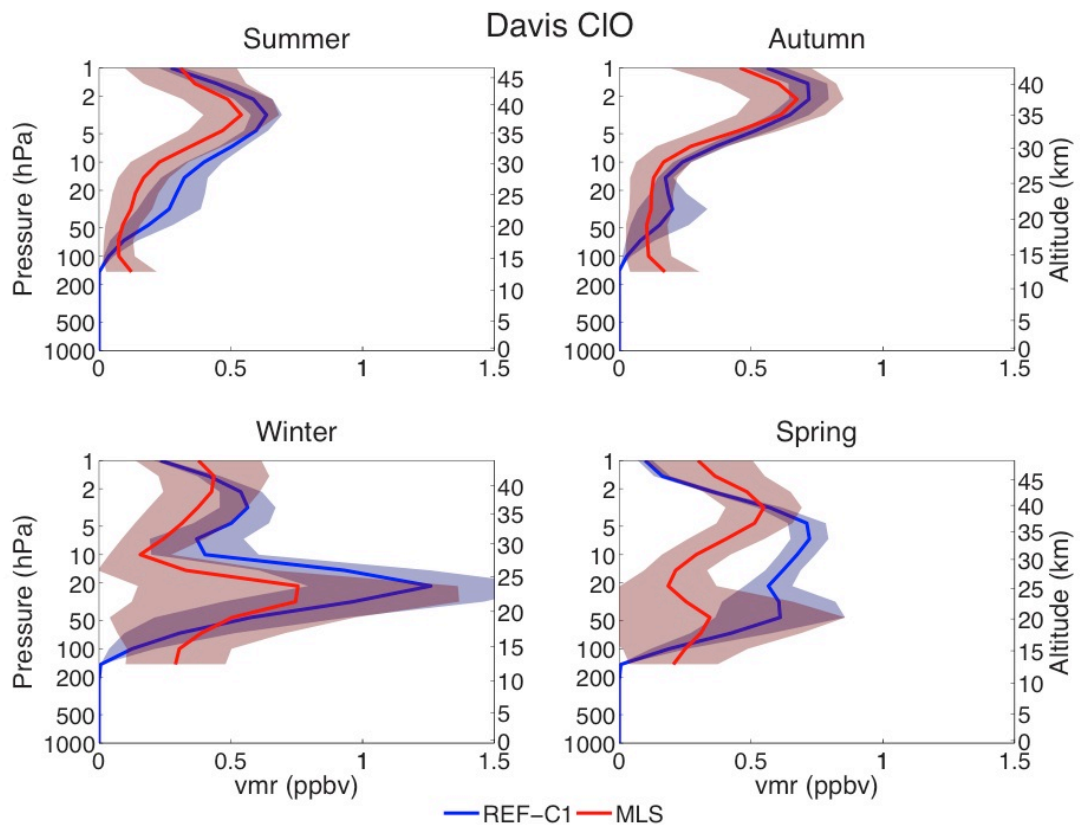


Figure 5. Update to Figure 5 in manuscript.