Response to interactive comment on "Assessing stratospheric transport in the CMAM30 simulations using ACE-FTS measurements" by Felicia Kolonjari et al. from Referee #3

RC = Reviewer Comment
AR = Author Response

RC: In this paper, the authors have used numerous techniques to compare a nudged CMAM model run with long-lived tracer observations from ACE-FTS. Their sampling technique is excellent and allows for a like-to-like comparison, as much as possible when comparing model output and data. The use of the TLP model as well as the tracer-tracer JPDFs are appropriate. I think this paper is appropriate for publication in ACP, subject to addressing a significant number of concerns about the dynamical interpretation and discussion of the results. I therefore waver between minor and major revisions. Only one of my comments requires additional calculations.
Generally, the following need to be addressed: 1) Improved discussion of BDC 2) Improved discussion of pathways for mixing across the tropopause and if feasible, 3) More specific discussion of the implications of these results for model development

AR: We thank Referee #3 for their detailed and constructive comments on our manuscript. We have addressed the general and the specific comments below.

RC P3L12: Hardiman et al. (2017) found a time of emergence for a trend of 30 years and showed that any trend less than 12 years could be the wrong way due to dynamical variability. The results of Mahieu and pretty much all of our observational records are too short.

AR: Thank you for identifying this recent publication. The authors were not aware of it. We have added the following sentence to the end of the paragraph.

New sentence: "Recently, Hardiman et al. (2017) determined that a period of 30 years is required for a trend to be identified from noise due to natural variability. They also found that dynamic variability can obscure a trend in the BDC if it is based on less than 12 years of data."

RC P3L13: I don't think there are fundamental questions about the mechanisms driving the stratospheric circulation. The mechanisms driving changes to the circulation are less clear, but the fact that data don't show the trend predicted to emerge from models over a much longer timescale is not surprising in light of the results of Hardiman et al.

AR: Agreed. Fluid dynamics provides a robust explanation for the mechanism driving the stratospheric circulation. We had meant to refer to the ability of models to quantitatively simulate the contribution of different processes to the BDC and the stratospheric age of air. Perhaps these are not fundamental questions. We have rephrased the sentence. This was also noted in our response to Reviewer #2.

Original sentence: "Fundamental questions remain as to the mechanisms driving the stratospheric circulation because there has been evidence of changes in the BDC that has not been projected by models (e.g., Butchart, 2014; Mahieu et al., 2014)."

New sentence: "Discrepancies between observations and model projects may be due to the short time scales of observation systems (Butchart, 2014; Hardiman et al., 2017)."

RC P3L16: Is it true that understanding how the structure of the BDC will change depends greatly on the ability to simulate its current behavior? The models examined by Butchart et al. 2011 have pretty different mean upward mass flux at 70 hPa, but the community still interprets their agreement on the strengthening of the BDC as robust. Getting the mean and present day right are important, but not necessarily for the trends.

AR: While all models agree on the sign of the change in the BDC, both the absolute and relative rates of change in the BDC show considerable variation across models. In addition to this, while we have confidence in a future acceleration, the authors do not think there is a consensus on the magnitude of the change or how it would be distributed between the shallow and deeper branches of the BDC. It should also be noted that our confidence in a future acceleration is not purely because of the model consensus, but also significantly because of physical processes that are directly tied to changes in the large-scale temperature structure of the atmosphere for which we do have great confidence. We do think that increasing confidence in our future projections of change would be strengthened by having more confidence in our capacity to model the present-day state of the BDC.

RC P3L16: "Typically"–please provide some citations demonstrating how typical.

AR: We have added four examples to demonstrate this.

Original sentence: "This is typically assessed by investigating how capable a model is at simulating tracer concentrations; in particular …"

New sentence: "This is typically assessed by investigating how capable a model is at simulating tracer concentrations (e.g. Jin et al., 2005; Allen et al., 2009; Park et al., 2013; Pendlebury et al., 2015); in particular …"

RC P4L11: You haven't definced CMAM30 before.

AR: CMAM30 is defined on P3L23.

RC P4L13: I would appreciate some discussion either here or in 2.2.2 of the potential problems with CMAM30. In particular, the model is nudged to reanalysis. As far as I am aware, neither the nudging process nor the reanalysis itself conserves mass, energy etc. Are there any studies that show how that influences tracers? Are your tracers transported conservatively and do their budgets close? I'm not *necessarily suggesting you calculate the tracer budgets, though such analysis might be interesting, but please address these concerns to whatever extent you can.*

AR: We are not aware of any studies that have investigated the effects of nudging on tracer conservation. While the process of nudging the dynamical fields most definitely violates conservation of energy, we believe that tracer advection is globally conservative as it is in the free-running model. The nudging is applied by relaxation with a 24-hour time constant and will be 'felt' by the model dynamical fields as an additional tendency, similar to that produced from any other parameterization of unresolved physical processes. Therefore we do not believe nudging itself introduces new problems for the conservation of advected tracers. The tracers are advected using spectral advection, which, while not positive definite does conserve global mass. The correction of negative concentrations by 'hole filling' will violate mass conservation and any addition of mass to correct for negatives when spectral fields are transformed to grid-point space is tracked and corrected for. The tracers analysed here, CFC-11, CFC-12 and N_2O, are long-lived with 'smooth' distributions that are well represented in spectral space resulting in very little problem with the generation of negatives. No nudging of surface pressure was used and we rely on the standard running correction of the global average surface pressure, the first spectral coefficient of the surface pressure, towards a pre-defined constant to ensure that the background mass of the atmosphere does not exhibit a trend at all.

In Section 2.2.2, P7L2 we have added the following text:
"As noted, tracers in CMAM30 evolve freely subject to advection by the resolved circulation and vertical redistribution by physical parameterizations. Advection of tracers is calculated using spectral advection, which is inherently mass conservative though not necessarily positive definite. The generation of negative concentrations upon transformation from spectral to physical space is corrected through 'hole filling' with any artificially added mass to remove negatives tracked and corrected for in the global average. The tracers analysed here are long lived and smoothly varying, resulting in spatial distributions that are well represented in spectral space and produce minimal problems with the generation of negative concentrations. No nudging of surface pressure is performed and the global average surface pressure is continually corrected back to a predefined constant value in the CMAM30 simulation, in the exact same manner as is done in free-running simulations. While mass conservation in the CMAM30 simulation has not been analysed specifically, no significant differences with free-running simulations have been seen for quantities such as the evolution of total stratospheric chlorine."

RC P4L17: "morphologies" of CFC11, CFC12 and N2O.

AR: Thanks for catching this.

Original sentence: "Section 4 examines the measured and simulated zonally-averaged morphologies."

New sentence: "Section 4 examines the measured and simulated zonally-averaged morphologies of CFC-11, CFC-12, and N_2O."

RC P6L19: the model isn't being constrained to follow observations. It's being constrained to the reanalysis, which is a model-data product that is our best guess at a representation of reality.

AR: You are correct. We have clarified this in the text.

Original sentence: "The ability to constrain the dynamical fields to follow the observations more closely enables direct model-measurement comparisons of chemical tracers in the model by eliminating the internal variability in the simulated circulation."

New sentence: "The ability to constrain the dynamical fields to follow the reanalysis (the best approximation of reality) enables direct model-measurement comparisons of chemical tracers …"

RC Section 2.2.4: How does CMAM BDC compare to ERA-I BDC? Mean tropical w at a few levels would be sufficient. The speculation in this section is not necessary when you can do direct comparisons. Additionally, this section would benefit from considering the extratropical vs. tropical age difference (e.g. Neu and Plumb 1999, Linz et al. 2016) rather than just talking around the relationship between the age and the circulation. For example, the near-zero differences in 2a between 50S and 50N do not mean that the lower branch of the BDC is the same in the two simulations because the polar age on the same level is older in the free running model. This discussion would be aided by the conversion to isentropic coordinates. This section is one place to address general comment 1) above. E.g. "filtered out close to the tropopause" could be explained in terms of the physical mechanisms of wave propagation (Charney Drazin).*

AR: To the authors knowledge, the CMAM BDC has not been compared to the ERA-Interim BDC. While the analysis suggested here may not be well beyond the scope of the paper, the authors feel that including such an analysis would add a lot of text and perhaps additional figures to an already extensive manuscript. As explained in response to a Reviewer #1, the purpose of including a comparison to the free running simulation was only to provide an important caveat to the results for the CMAM30 nudged simulation. While the use of a nudged simulation allows for a time- and space-matched comparison to the ACE observations, we wanted to make the reader aware that the residual circulation in CMAM30, along with age of air and the distribution of long-lived tracers, is different than that which we find in the freely running version of the model. An analysis of the cause of these differences would constitute a separate study. We have modified sections of the paper where we speculate on possible causes of the differences to make it clear that the reasons for the differences are unclear at the present time and are outside of the scope of the paper.

RC P12L11: "CHAM" → "CMAM"

AR: Fixed. Thanks!

RC P12L30: "Air in the polar vortex … representative of older air …" The terminology "representative of" is confusing to me. Isn't air in the polar vortex composed of a larger fraction of older air transported from upper levels?

AR: We agree that the use of "representative of" is a bit confusing.

Original sentence: "Air in the polar vortex is typically representative of older air brought down from higher altitudes."

New sentence: "Air in the polar vortex is typically composed of older air brought down from higher altitudes."

RC P12L34-35: If the variability of the vortex edge is responsible, why is there so much difference in the middle of the vortex, and why is the Southern hemisphere, where the vortex variability is much weaker pretty comparable to the Northern hemisphere?

AR: This is a valid point, the differences observed are not just due to the variability of the vortex edge but also reflect that ACE-FTS samples the large scale downwelling of air within the vortex differently from the zonal mean average. The text has been modified to reflect this detail.

Original sentence: "The differences seen in Fig. 4a occur because comparing the full output to measurements does not account for the variability of the vortex edge in both longitude and latitude."

New sentence: "The differences seen in Fig. 4a occur because comparing the full output to measurement-like samples of the output does not account for the variability of the vortex edge in both longitude and latitude, nor does it account for the differences in spatially- and temporally-sampled large-scale downwelling of air within the vortex compared to a zonal mean that includes the model simulation at all longitudes and time periods."

RC P13L17: No comma before between.

AR: This has been fixed.

RC P14L5-9: I found this section strange. "readily observed" where? The other information seems redundant. Perhaps just remove all together?

AR: This section on P14 L5-8 has been removed.

Text removed: "Most long-lived tracers with tropospheric sources exhibit quantitatively similar behaviour in the upper troposphere and lower stratosphere. In the context of a zonally-averaged tracer morphology, the equator-to-pole gradients of tracer isopleths that are created by the diabatic circulation in the stratosphere are readily observed. By choosing to sample the CMAM30HR

output as described above, the behaviour of N_2O, CFC-12, and CFC-11 can be investigated thoroughly."

RC P14L16-17: Redundant … rephrase.

AR: We agree that the sentence could be clearer. It has be edited.

Original sentence: "The distributions show a decrease in concentration of N_2O with altitude at all latitudes, and also moving from the equator poleward at each pressure level and in each hemisphere.

New sentence: "The southern extratropical and Antarctic concentrations of N_2O tend to decrease with altitude more rapidly than those in the Northern Hemisphere."

RC P14L17-18: "Likely caused by significant differences in the conditions of the influence of downwelling…" This is confusing. Please rephrase or explain further. There is more downwelling in the SH vortex and you've mentioned later that N2O has a source higher up, so shouldn't there be more N2O in the SH vortex than the NH vortex?

AR: This sentence has been edited to provide more clarity. The authors were referring to the reflection of the differences in downwelling in the two hemispheres. The source in the upper stratosphere doesn't significantly impact depletion of N_2O in older air and there isn't necessarily a hemispheric difference in this source. Therefore, the hemispheric asymmetry in the differences between the observations and simulations is primarily driven by the differences in downwelling and the related differences in the isolation of the vortex.

Original sentence: "This is likely caused by significant differences in the conditions of the influence of downwelling within the polar vortex between the two hemispheres."

New sentence: "This asymmetry is likely driven by differences in the isolation of the polar vortex in each hemisphere and the large-scale downwelling that is largely dependent on this isolation."

RC P15L11: Another place to address 1). The BDC is strongest in the NH winter because of the climatological westerlies and the enhanced wave driving from the troposphere both. If there were more waves in the NH summer, they wouldn't do any good because they can't propagate up into the stratosphere when there are climatological easterlies.

AR: We have clarified the sentence at P15L11-12.

Original sentence: "In general, the BDC is strongest in the Northern Hemisphere winter due to wave driven enhancements initiated by topography (e.g., Rosenlof, 1995; Plumb and Eluszkiewicz, 1999). "

New sentence: "In general, the BDC is strongest in the Northern Hemisphere winter because of wave driven enhancements initiated by topography, and because, during that time of year, climatological westerlies facilitate wave propagation into the stratosphere (e.g., Rosenlof, 1995; Plumb and Eluszkiewicz, 1999)."

RC P15L28: not sure what you mean by "robust"

AR: The sentence has been changed to clarify the meaning.

Original sentences: "However, the shifting of the agreement in the tropical region through the seasons indicates a robust simulation."

New sentences: "Meanwhile, the shifting of the agreement in the tropical region through the seasons indicates that the simulation is consistent with the spatial distribution of the observations in this region."

RC P16L1: "significantly" is confusing. Significant with respect to what?

AR: The comparisons are different in each polar region. The text has been edited to provide more clarity.

Original sentence: "The polar regions of each hemisphere in the comparisons of Fig. 8 are significantly different."

New sentence: "In the comparisons shown in Fig. 8, the Northern polar region measurement-model differences are significantly different compared to those in the Southern polar region."

RC P16L5: "particular" → "particularly"

AR: This has been fixed.

RC P16L5-6: Please explain more why these differences would be due to the polar vortex behavior. I agree with you, but a discussion of the mechanism would be helpful.

AR: The last sentence of this paragraph has been modified to clarify the cause of these differences.

Original sentence: "These differences are likely due to the behaviour of the polar vortex in each hemisphere.

New sentence: "These differences are likely due to the behaviour of the polar vortex in each hemisphere; in particular, they may be related to the models' (either CMAM30HR or the ERA-Interim model used for the nudging or both) ability to represent transport processes in the strong, cold, quiescent Antarctic vortex versus the warmer and more variable Arctic vortex."

RC P17L4: no commas offsetting "with a stratospheric sink"

AR: This has been fixed.

RC P17L26: no "was" before "passed"

AR: This has been removed.

RC 5.2: This discussion needs to be revised. Specifically, please review the literature that treats the tropopause as a "barrier", review the recent literature on transport across the tropopause (Randel 2017 tropospheric dry layers or Randel 2016 asian monsoon transport, for example), discuss the difference between what has been defined here and the more typical treatment of stratospheric intrusions (tropopause folding events that cause deep stratospheric intrusionsâ˘A˘ Tsee work by Meiyun Lin, for example). Compare to stratospheric intrusion climatology (Skerlak et al. 2014). Finally, please validate your method for defining intrusion events by looking at the colocated water vapor and ozone concentrations in the model. (Or other stratospheric tracers, you could use PV.)

AR: We have decided to continue to use the term intrusion in this context. The calculation presented in the manuscript is intended to provide a diagnostic for comparison of the observations and the simulations. While water vapor, ozone or other stratospheric tracers could illustrate what is happening in the atmosphere, using the algorithm employed here on other species would not necessarily validate the method. CFC-11 was chosen because of its distinctly different stratospheric loss rates. The algorithm would have to be adjusted for water vapor or ozone. In this work, we are demonstrating that CFC-11 is species that can provide another diagnostic to identify the origin of air similar to water vapour or ozone. In addition to this, utilizing the ozone and water vapour ACE-FTS products would require extensive comparison to the CMAM30 simulations as was done for the halocarbon simulation. The authors have decided to not perform this additional work as it would be a significant undertaking and that is not within the scope of this project. However, we have added a sentence after the first sentence in paragraph 2 of this section to clarify the definition of intrusion used in this work.

New sentence: "The diagnostic developed for this analysis is the frequency of intrusions, which signifies the frequency of stratospheric (tropospheric) air penetrating into the troposphere (stratosphere)."

RC P20L19-20: Which differences and how? I must have missed the discussion previously, so a brief repetition here couldn't hurt.

AR: This sentence has be changed to clarify this point.

Original sentence: "This difference between the measurements and simulations is likely due to the overly rapid BDC in the model simulation, as previously discussed."

New sentence: "These differences between the measurements and simulations are likely due to the overly rapid BDC in the model simulation, leading to higher concentrations of the trace gases in the simulation, which is consistent with the zonal mean comparisons discussed in Section 4."

RC P20L25-6: Have you demonstrated this? If so, how?

AR: While we have not demonstrated this explicitly in this manuscript, this was demonstrated by Hegglin and Shepherd (2007). This reference has been added to this sentence.

Original sentence: "It is the atmospheric variability that contributes to the variability observed in the ACE-FTS JPDF around 150-200 ppbv of N_2O."

New sentence: "It is the atmospheric variability that contributes to the variability observed in the ACE-FTS JPDF around 150-200 ppbv of N_2O (Hegglin and Shepherd, 2007)."

RC P21L1: add "timescale" after transport.

AR: "Timescale" has been added.

Original sentence: "This relationship is observed because the photochemical lifetime of CFC-11 is shorter than the time scale for mixing by horizontal eddy transport …"

New sentence: "This relationship is observed because the photochemical lifetime of CFC-11 is shorter than the time scale for mixing by horizontal eddy transport timescale …"

RC P21L7: This needs to be more specific. How were the turn around latitudes determined? Were they monthly mean or instantaneously calculated?

AR: They were determined from the monthly mean vertical velocities. We have added "monthly mean" to sentence.

Original sentence: "The data were selected from the tropical latitude region using estimates of the turn-around-latitude, the height-dependent latitude where the tropical upwelling is zero, determined from CMAM30HR vertical velocities."

New sentence: "The data were selected from the tropical latitude region using estimates of the turn-around-latitude, the height-dependent latitude where the tropical upwelling is zero, determined from CMAM30HR monthly-mean vertical velocities."

RC P21L18-9: Wording is informal

AR: The language has been formalized.

Original sentence: "However, the differences between the measurements and simulations are primarily in the steepness of this segment of the JPDF, which is a sign of having not enough mixing into the tropics, rather than a product of a too rapid tropical ascent."

New sentence: "However, the differences between the measurements and simulations are primarily in the steepness of this segment of the JPDF, which is an indication of insufficient mixing into the tropics, rather than too rapid tropical ascent."

P21L23-5: You've already talked about Fig. 14, so why is this here?

AR: This has been deleted here and we have added a reference to Fig. 14 in next sentence.

Original sentences: "Figure 14 isolates the N_2O/CFC-11 JPDFs to the tropical region only, as defined by the turn-around-latitudes, for both the ACE-FTS measurements (left column) and the CMAM30HR simulations (right column) for each season. There is an evolution of the characteristics of the JPDFs shown in this figure."

New sentence: "There is an evolution of the characteristics of the JPDFs shown in Fig. 14."

RC Section 6: More discussion of the TLP would be useful. I am very familiar with it, but Eric's paper is complicated, so a brief discussion of why it's great and useful here would be helpful for the average reader who isn't going to want to read through his whole paper.

AR: As noted in response to Reviewer #2's comments, more discussion of the TLP model has been added and a portion of Section 2.3 was moved to Section 6. The authors think the descriptions are now sufficient for the average reader.

RC P22L14: "mixing levels" I thought it was mixing efficiency.

AR: This phrase has been changed to "mixing efficiency" throughout the manuscript.

RC P22L17: Be more specificâ˘A˘Te.g. " As both the residual circulation and the mixing are driven by wave breaking, a weaker residual circulation likely correlates with less mixing and thus longer mixing timescales"

AR: We agree with your suggested change.

Original sentence: "A reduced mean circulation would likely correspond to less mixing and longer mixing times."

New sentence: "As both the residual circulation and the mixing are driven by wave breaking, a weaker residual circulation likely correlates with less mixing and thus longer mixing timescales."

RC P22L25-6: The implication here and elsewhere in this section is that there is some knob to turn to "change" w or epsilon. There obviously isn't, and so while this paper has diagnosed that the mixing is too weak in JJA, for example, it hasn't come close to determining changes required in the CMAM30HR simulations to match the observations. Certainly changing the language here and P23L14, L29, P24L7 is necessary. If feasible, some discussion of what does set w* (epsilon is probably harder) would be great. If anyone has looked at EP flux divergence esp. broken down by wavenumber, even in CMAM and not necessarily CMAM30, that would be a great thing to discuss here. The authors have done plenty to warrant publication and so they don't need to do it if it hasn't been done. Regardless, some discussion of what does drive the circulation and the mixing in the model would be appropriate here.*

AR: While we cannot pinpoint an exact mechanism that would improve the CMAM30 simulations, the TLP model does provide a pathway or blueprint to make changes in CMAM that you wouldn't have otherwise. The end of Section 6 has been edited to provide more explanation of what could lead to improvements in CMAM.

Original sentences: "In particular, JJA appears to require increased mixing in all regions studied, implying that there is a substantial deficiency in the CMAM30HR simulation during this season. The most significant physical mechanism for mixing during this season is the Asian monsoon. The quality of this transport mechanism has not been directly assessed in CMAM30. It is unclear as to whether the mechanism has a direct or indirect effect but the Asian monsoon is a prominent climatological feature of the upper troposphere and lower stratosphere at this time of the year and it can be speculated that the required additional mixing may be related to the strength or extent of the simulated monsoon."

New sentences: "It follows that if *w needs to be reduced in the model then a reduction in wave activity is required. There are specific waves that break in the lower, middle, and upper stratosphere that could be investigated for possible sources of increased *w**. For mixing changes, the background state of the winds and corresponding critical layers for wave breaking could be investigated for critical layers that extend too far into the tropics."**

RC P24L23: If you've proven that the BDC is too rapid, say so here. As far as I recall, you said that it might be too strong.

AR: The sentence has been modified to more clearly state that the BDC is too rapid in CMAM30.

Original sentence: "The polar vortex comparisons reveal issues in both the timing and strength of the downwelling portion of the deep branch, which is likely directly related to the too-rapid overturning nature of CMAM's BDC."

New sentence: "The polar vortex comparisons reveal issues in both the timing and strength of the downwelling portion of the deep branch, which is related to the too-rapid BDC in CMAM30 simulations observed in the zonal mean comparisons."

RC P24L33-4: Nonsense. Insufficient mixing does not cause any changes to the BDC (at least to first orderâ˘A˘ Tsecond order effects on ozone and the corresponding heating due to ozone might be minor but that has definitely not been addressed in this paper).

AR: We agree that this is not an accurate statement based on the analysis presented here. We have decided to review the entire paragraph to make it a more accurate summary of our findings.

Original paragraph: "Insufficient mixing during JJA may be related to the poorly simulated tropospheric intrusions during the same season and this same issue may be directly related to the younger air in CMAM30HR. Garny et al. (2014) found that, in the subtropical lower stratosphere, younger air is the result of a combination of a speeding up of the overturning circulation and weaker mixing or recirculation of the stratospheric air between the tropics and midlatitudes. This may be evidence for insufficient mixing in the specified dynamics simulations being the cause of a too-rapid BDC. It is important to scrutinize the mixing levels in CCMs and GCMs since it appears to be related to the mechanisms driving the projected trends in stratospheric circulation, thereby influencing the simulations of stratospheric ozone recovery and climate change. The techniques used in this work, including the advanced sampling and use of the tropical leaky pipe model, have proven illuminating. It is suggested that other CCMs and GCMs investigate the use of these techniques in future studies."

New paragraph: "The analysis presented here highlights the importance of scrutinizing the mixing efficiency in CCMs and GCMs since it may be related to the mechanisms driving the projected trends in stratospheric circulation, thereby influencing the simulations of stratospheric ozone recovery and climate change. The techniques used in this work, including the advanced sampling and use of the tropical leaky pipe model, have proven illuminating. It is suggested that other CCMs and GCMs investigate the use of these techniques in future studies."