Answer to Referee #1

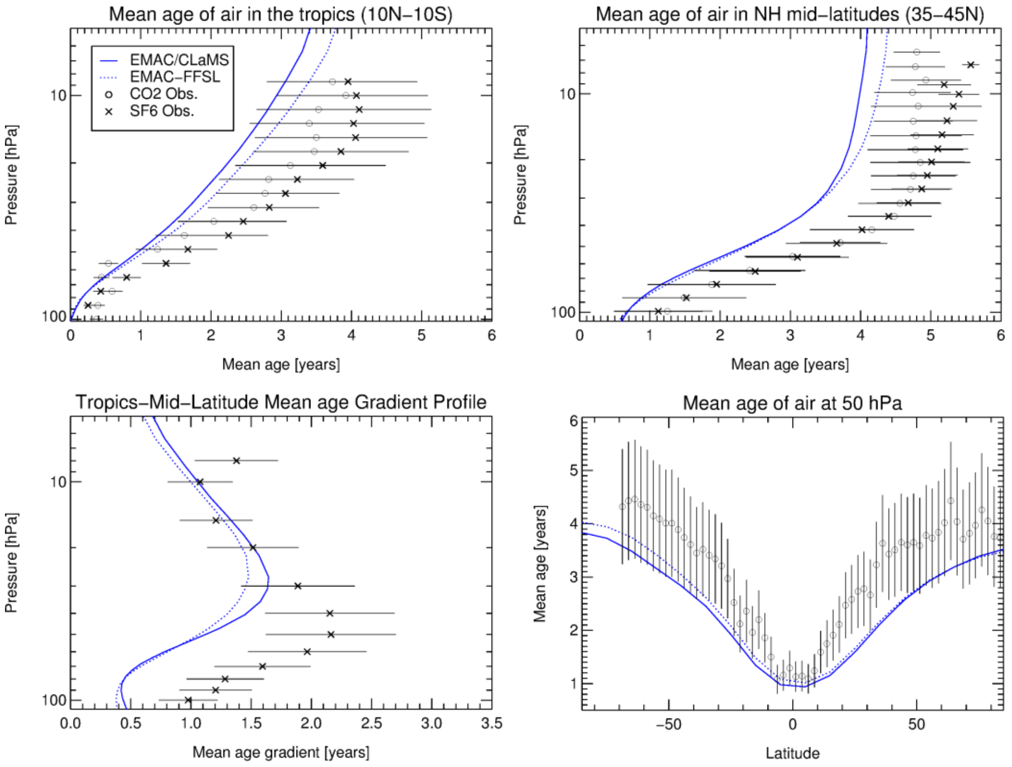
We thank the reviewer very much for the helpful comments on our manuscript. All comments were taken into account in the revised version and we feel that the manuscript has improved substantially through these revisions. Below, we reply point-by-point to the referee comments:

Section 3.2 Age of Air
p. 1771, lines 15-17. Use simulated mean age relative to the tropical tropopause in Figure 2. It is acknowledged that the observational age is referenced to the tropical tropopause yet the plot shows simulated ages since surface emission. This comparison must be made fairly. Please note the transit time for each model from the surface to the tropical tropopause.

Answer:
Figure 2 has been changed and expanded as suggested by the referee. In the new version of the plot, the model data is also shown relative to the tropical tropopause. For the discussion of the age of air profiles, see our answer to *p. 1771, lines 19-22.*

p. 1771, lines 19-22. Figure 2, once simulated age is correctly plotted, will show simulations whose extratropical age is quite young. Twelve of the 15 of the CCMs compared in this same way (SPARC 2010) are in better agreement with the observations than these simulations, so I do not agree this is a 'typical feature which can be much more pronounced than in the example shown here.'
The evaluation of mean age is inadequate but can be easily improved. The SPARC CCMVal report, referenced by the authors, shows four mean age diagnostics that together reveal much more about transport. Please use the 4 diagnostics shown in Figure 5.5 of this report. For example, the tropical-midlatitude mean age gradient profile is useful for evaluating credibility of the tropical ascent rate. From Figure 2 I calculate this gradient at 50 hPa to be about 1.3 years for both simulations. That indicates much faster than observed circulation, and faster than most CCMVal CCMs. The mean ages derived from observations used in these diagnostics are publically available.
Better still for diagnosing ascent rate, compare directly with the observationally (tape recorder) tropical vertical velocities shown in Fig. 5-6 of the CCMVal report.

Answer: As suggested, we now show additional age of air diagnostics in the revised version of the paper, according to Figure 5.5 in SPARC (2010):



We replaced the discussion of the age of air (p.1771, ll. 6-27) by:

We show mean age of air for the EMAC/CLaMS and EMAC-FFSL climatologies in comparison to mean age of air derived from measurements in Fig.2. The age of air values are derived from CO2 and SF6 measurements (Andrews et. al. 2001, Engel et. al. 2009). Figure 2 shows that both models produce a similar and reasonable age of air distribution, which lies in most cases at the lower boundary of the one-sigma uncertainty range of the age of air values derived from measurements. The tropical profile for EMAC/CLaMS shows slightly younger age than EMAC-FFSL due to stronger upwelling in the tropics. In mid-latitudes, the age of air profiles from the simulation climatologies are about 1 to 1.5 years younger than the age of air profile derived from CO2 and SF6 measurements. This is a typical feature in models, thus the profiles shown here are comparable to many models, which are compared in a similar way in SPARC (2010).

The gradient (here: the difference) between the mid-latitude and tropical profile fits well to the measurements at high altitudes down to 30hPa, whereas in the lower stratosphere the gradient in the model climatology is lower than in the observations (Fig. 2, bottom left panel). The difference between the tropical and mid-latitude profiles is slightly higher in the EMAC/CLaMS simulation, which indicates a stronger transport barrier at the edge of the tropical pipe in EMAC/CLaMS.

Figure 2 also shows zonal, annual mean age at 50hPa for all latitudes. The simulated age of air pattern is consistent with the general features of the stratospheric circulation. There is upwelling of young air masses in the tropics, and downwelling of old air masses in the polar regions. Both simulations exhibit slightly older air in the Southern Hemisphere compared to the Northern Hemisphere. The age in EMAC/CLaMS is slightly younger in the Southern hemisphere than EMAC-FFSL due to stronger upward vertical velocity in the EMAC/CLaMS representation and different properties of the subtropical barrier.

The differences in the zonal, annual mean between the two model representations do not exceed three months, but regional and seasonal differences may be larger.

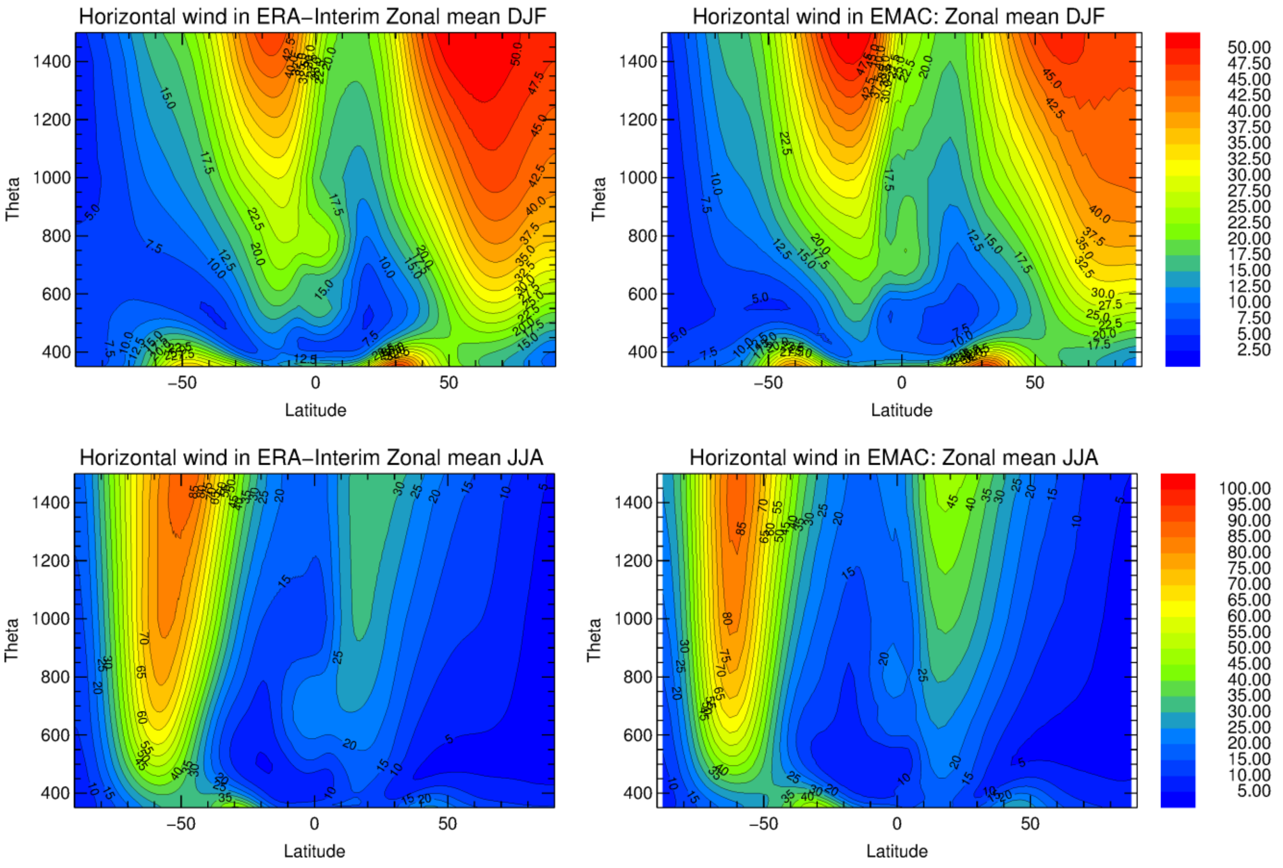
p. 1772, lines 1-6. Please report the simulated lifetimes and compare with the SPARC lifetimes report. Here the authors make a qualitative, subjective statement (‘compare very well’) and reference an unpublished source (Hofmann et al, 2014 – in prep). The recently published SPARC Lifetimes report has new recommended *values for CFC11 and CFC12 which should be used in your comparison. Please state the simulated lifetimes too. The lifetimes report is available at http://www.sparcclimate.org/publications/sparc-reports/sparc-report-no6/. The new lifetimes were derived from many observations and models, carefully evaluated and combined to derive the most likely lifetimes and uncertainties. The recommended values and most likely uncertainties are 52 (43-67) years for CFC11 and 102 (88-122) years for CFC12. In Chapter 4 of this report, lifetimes derived from satellite or space-based observations showed a wide range of values with large uncertainties.*

Answer: The article by Hoffmann et. al. is now published in *Atm. Chem. Phys. Discuss*. We changed the paragraph in the following way:

Zonal mean trace gas climatologies from EMAC/CLaMS were also used to derive the relative lifetime of CFC-11 and CFC-12 (Hoffmann et al. 2014). The results compare very well with lifetimes derived independently from various satellite climatologies. The lifetime ratio of CFC-11 and CFC-12 from the EMAC/CLaMS simulation yields 0.48 (0.40-0.55). The satellite climatologies Hoffmann et al. (2014) deduce lifetime ratios ranging from 0.46 (0.39-0.54) to 0.47 (0.39-0.54). The good agreement between the model deduced and observationally deduced lifetimes provides further confidence in the representation of transport and chemistry of long-lived tracers in the EMAC/CLaMS model system. The results also correspond very well with the recommendations for the lifetimes of CFCs by SPARC 2013 which provide a lifetime ratio of CFC-11 and CFC-12 of 0.51 (0.35-0.76).

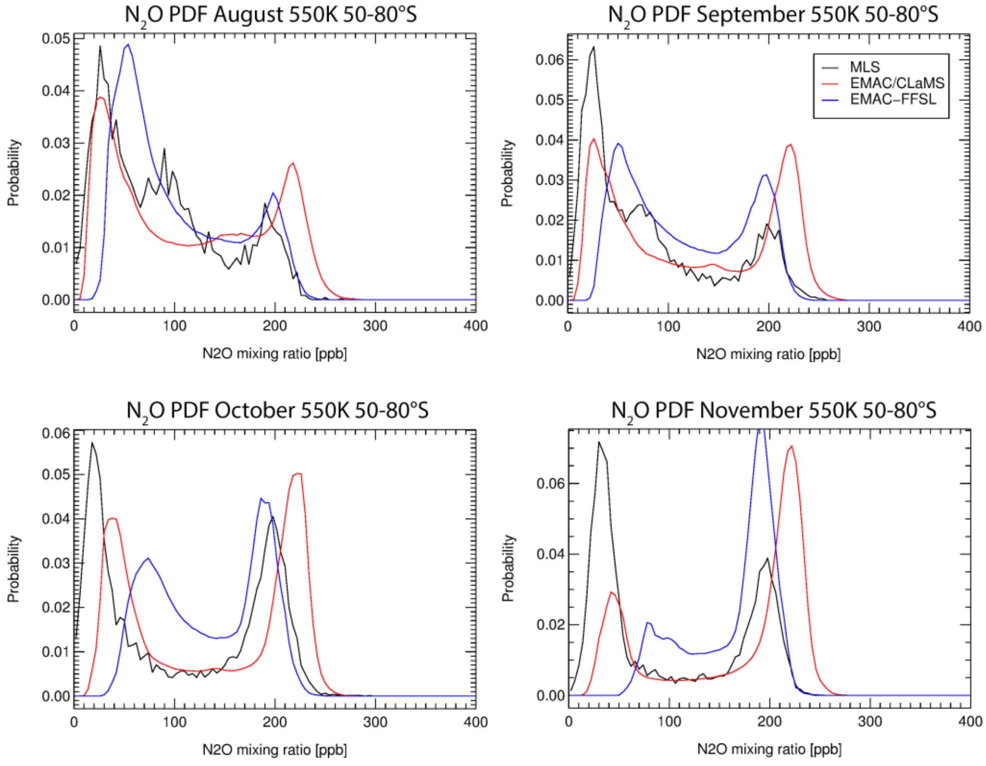
p. 1772, last paragraph. It is stated that the simulated horizontal wind (u and v) matters for vortex isolation (agreed), but then you compare only the zonal wind, u – that ignores the v wind which actually matters for meridional transport (i.e., vortex isolation). Please use a better meteorological metric for vortex isolation in the two simulations, perhaps the mean u'v' (horizontal momentum transport) or potential vorticity.

Answer: We followed this suggestion and changed Figure 3 to horizontal instead of zonal wind.



Section 3.3.1 Antarctic Polar Vortex
p. 1774, lines 5-7. A mean trace gas profile, especially a multi-year mean, does not reveal anything about processes (fig. 5). The profile is the net result of two ongoing processes, descent and mixing, not either one individually. (Those processes differ significantly across the vortex edge. Since these profiles are multi-year means from 70-80S, they may include profiles from outside the vortex although I agree that these are probably mostly vortex profiles.) The differences in September CFC11 profiles can be caused by several things: 1) differences in descent rates during fall and winter, 2) different initial profiles when descent began, 3) differences in vortex isolation at higher altitudes that affected vortex profiles during descent. In addition, the simulated CH4 profiles have small differences, and they are more similar to each other than to the observations.
Even combined with statements about age of air gradients in Fig. 4, this discussion does not unambiguously separate the effects of descent and mixing. I suggest another of the CCMVal transport diagnostics, the Antarctic CH4 pdf profiles shown in Fig. 5-15. The contoured vertical pdfs (68-78S) show vortex isolation as a function of height. As you are trying to make the point about process differences between August and October you could compare PDFs for these months. This could clearly identify downwelling differences (by changes in the polar most probable profile) while at the same time identifying mixing differences (i.e., how well separated the midlatitude and polar most probable profiles are and how their separation changes between August and October). Although Fig. 5-15 used HALOE CH4 data from mid-October to mid-November (HALOE did not sample high latitudes in August), the model-model comparison would reveal process differences between the FFSL and Clams transport. One could take this a step further by using MLS N2O data, available daily during all months over the past 10 years, for the contoured PDF calculations. This could be used to evaluate descent and mixing over the entire winter. This would most useful for processes occurring below 700 K.

Answer: We performed the suggested PDF analysis and have now included the following PDF plots and the accompanying discussion in the paper.



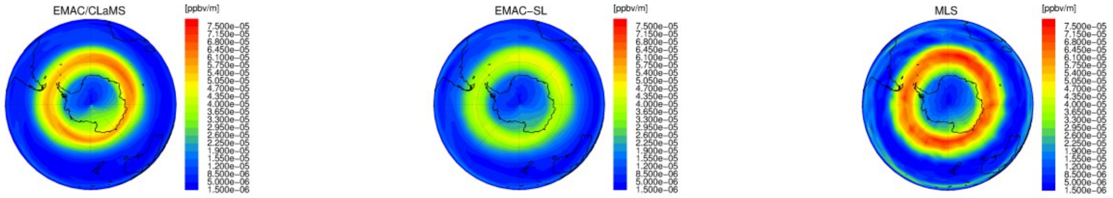
In Figure 6 we show N2O PDFs at 550 K from 50-80°S for the months August to November. Here,

EMAC/CLaMS and EMAC-FFSL results are compared to MLS satellite data. The PDFs show a two-peak structure, indicating the separated air masses inside and outside the polar vortex. The peak at lower N2O mixing ratios of about 30 ppb characterizes the air inside the vortex. In EMAC-FFSL the lowest observed values are not reached, which indicates that the downweling in this model representation is too weak. It is also visible that the vortex breaks up too early in EMAC-FFSL, since in October the vortex peak has nearly vanished completely. In EMAC/CLaMS, the peak position is captured well in in most months except for October. The peak in EMAC/CLaMS is less pronounced than in the MLS data, but the vortex is more stable than in EMAC-FFSL.
The second peak around 200 ppb indicates mid-latitude air. The mid-latitude peak is well captured in EMAC-FFSL. In EMAC/CLaMS, the peak value is about 20 ppb higher than in the measurements. This corresponds to the faster upwelling in the Southern hemisphere in EMAC-CLaMS, which is also visible in the age of air distribution.
The separation (i.e. the range of low probability values) between the two peaks of the PDF is an indicator for the strength of the transport barrier at the edge of the polar vortex. Here, using EMAC/CLaMS leads to a clear improvement compared to EMAC-FFSL. The separation between the two peaks is well captured in the Lagrangian transport representation.
The comparison of CH4 PDFs of EMAC/CLaMS and EMAC-FFSL with HALOE measurements (Grooß and Russell, 2005) shows similar results (not shown).

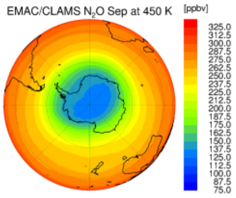
p. 1775, lines 1-4. Since Figure 3 showed zonal, not horizontal wind (or momentum flux or PV), it is not clear which simulation has a stronger vortex, greater descent, or differences in vortex isolation.
Answer: The plot has been changed to horizontal wind, and the effects of a stronger vortex, greater descent, or differences in vortex isolation are mentioned in the discussion of the PDF plots.

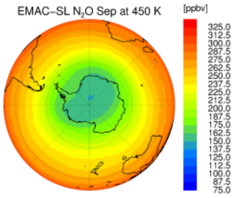
Figure 6 is a nice comparison and would be complementary to contoured vertical pdfs. But what is going on with the MLS data plotted? There are yellow 'stripes' running from the high gradients (from 65oS) to the pole. This does not look at all physical. Also, MLS observations are not made poleward of 82o so 'data' should not be plotted all the way to the pole.
The curious yellow stripes are also found on the EMAC/FFSL panel (but not the Clams panel). The Emac/ffsl gradients are noisy with spuriously high values around latitude circles (numerous yellow stripes along meridians in a sea of blue).Is this a plotting problem or is there a mysterious source of polar N2O in the FFSL simulation (i.e., transport implementation error)? I don't understand why this unphysical feature also appears in the MLS panel. I would like to see how the N2O fields compare between MLS and the simulations.

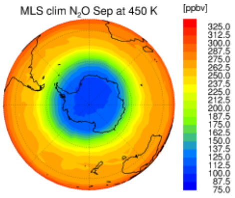
Thanks for pointing this out. The stripes that occured in the original version of the plot were due to a plotting error, i.e. they had no physical meaning. We corrected the plot program, so that this error does not occur any more.



The respective N2O distributions at 450K look like this:

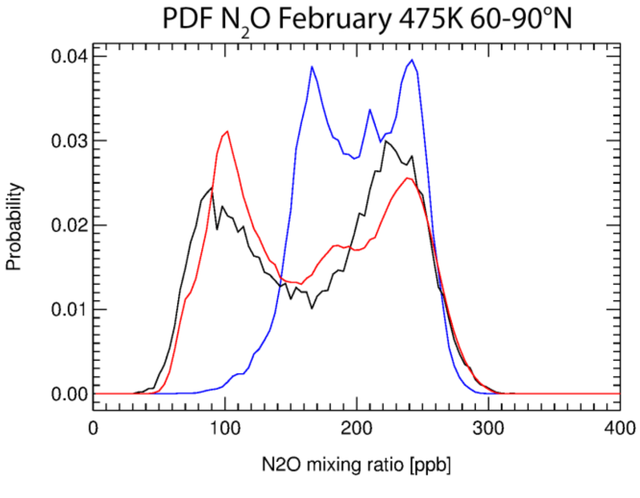


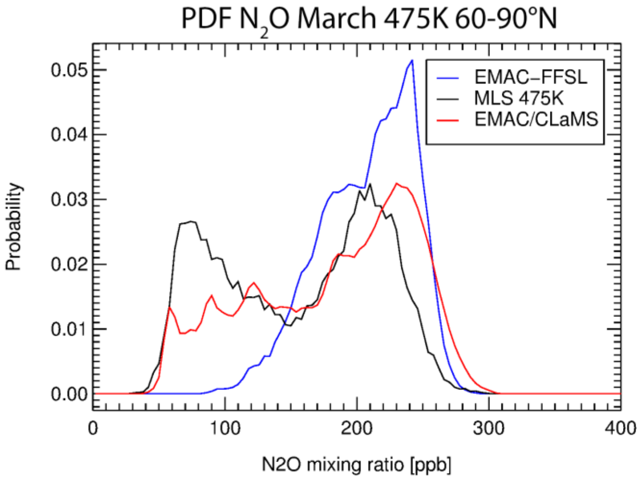




Section 3.3.2 Arctic polar vortex
While the Clams simulation is older and it might be because the vortex barrier is stronger, the steepness of edge gradients that would show this can't be discerned in this figure. (The MLS N2O gradient plot was better in this respect.) This could be better compared using PDFs of age or N2O (say 60-90N) as they would clearly show mixing by the height of the minimum between the mid and high latitude peaks. Comparing PDFs for February and March would show if the degree of mixing in each model were changing during Arctic winter, plus the change in the high latitude peak would indicate whether additional descent were occurring (unless there was a lot of mixing, in which case you could not distinguish the separate effects of descent and mixing). p. 1776, lines 15-16 claim that greater mixing has taken place in the FFSL, but this isn't something that can be discerned from these maps. PDFs would show this. If Arctic transport were analyzed using N2O instead of age, this would have the advantage of allowing a comparison with MLS N2O. While one of the primary goals of this paper is to show transport differences between the two model versions, it is also of value to show how realistic they are. At present I don't agree with line 28, that these results are consistent with the SH analysis given the odd stripes shown in the N2O gradient figure.

Answer: Following this suggstion, we added PDF plots for the Arctic for the months February and March as well. The model results are compared with MLS observations. The plots containing the stripes have been corrected.





The following text has been added to the paper:
In Figure 8 we compare PDFs from 60-90°N for February and March with MLS measurements, similar to the analysis for the Southern hemisphere. The peaks of the NH PDFs are wider than in the PDFs for the Antarctic, which illustrates the larger variability of the Artic polar vortex. The PDFs picture the problems of EMAC-FFSL in representing the Arctic polar vortex. In February, the peak N2O mixing ratio in EMAC-FFSL of 170 ppb is much higher than in the measurements, for which the peak value is located around 100 ppb. The separation between the polar vortex air and the mid-latitude air is very weak in EMAC-FFSL. In March, the two-peak structure vanishes in the EMAC-FFSL PDF. EMAC/CLaMS shows improved downwelling and vortex isolation compared to EMAC-FFSL. In February, the structure of the N2O PDF from measurements is well represented by EMAC/CLaMS. In March, low N2O mixing ratios below 100 ppb

inside the vortex are possible in EMAC/CLaMS, but they do appear less often than in the measurements. Nonetheless, this is a clear improvement to the EMAC-FFSL simulation, where no vortex structure is visible in the N2O PDF in March.

Minor comments

p. 1767, line 24. 'There are two ways to use: : :' Instead of this being the last sentence of a paragraph it should be the first sentence of the next paragraph.
changed

p. 1769, line 7. Change 'a profound discussion' to 'an in-depth discussion'
changed

p. 1773,
line 5. 'In contrast' instead of 'on the contrary'.
changed

p. 1773, lines 10-12. A problem with the wording, it's downwelling (of air) not downwelling of vertical velocities you're talking about. I think you're meaning is that diabatic vertical velocities are larger in the vortex than the kinematic velocities, yes?
Yes, the sentence has been changed to:
It is found that in August the diabatic vertical velocities in EMAC/CLaMS are larger in the downwelling region of the polar vortex than the respective kinematic vertical velocities in EMAC-FFSL.

p. 1773, line 14. suggest changing 'downwelling decreases in strength' with 'downwelling weakens'
changed

p. 1773, line 15. It's the relative impact that you're talking about. Just because descent has slowed to about zero in October does not mean mixing across the vortex barrier has increased or is strong, but the way this is worded increased mixing is implied.
We replaced
'However, in October downwelling decreases in strength and becomes less important for the age of air distribution so that the impact of horizontal transport through the vortex edge and mixing increases.'
with
'However, in October downwelling weakens so that the relative impact of horizontal transport through the vortex edge and mixing on the age of air distribution increases.'

p. 1773, line 20. The pattern of trace gases and age of air are both controlled by transport – I think this is your intended meaning. 'Age of air distribution' is not a force and cannot dominate control.
changed

p. 1776, lines 8-9. Older mean age could indicate a strong vortex but not necessarily. Age is a function of both circulation and mixing.
We replaced the sentence by:
In February, the EMAC/CLaMS simulation shows a more pronounced pattern in the age of air distribution, which is a result of a stronger polar vortex barrier and downwelling.