Response to interactive comment on "Assessing stratospheric transport in the CMAM30 simulations using ACE-FTS measurements" by Felicia Kolonjari et al. from Referee #1

RC = Reviewer Comment
AR = Author Response

RC: The authors compare free-running and nudged simulations using the CMAM middle-atmosphere model to ACE-FTS measurements of long-lived tracers. The analysis is generally well grounded and based on established analysis techniques, such as tracer-tracer correlation plots. Generally my impression is that the paper takes in a lot of information, making this a fairly dense read. For the future, I recommend to the lead author to break up such works into into smaller, separately publishable pieces. I don't think it would be adequate to recommend this course of action for the present paper as this is only a matter of presentation. The captions of some figures could be more detailed; for examples see below. In the comparison of models versus satellite measurements, a more thorough discussion of the effect of measurement uncertainty on tracer-tracer plots would be desirable. For example, the density plots of N_2O versus CFC-11 and CFC-12 (figures 13a and c), in the case of the satellite, are probably affected by measurement noise giving the JPDFs a fuzzy appearance. Such noise is absent in the model, making for a skewed comparison. Possibly averaging kernels of the ACE-FTS measurements could be used to define random noise to be added to the model data, making them more comparable to the measurements. If this is not practical, at least some text to this effect would be good to have.

AR: We thank Referee #1 for their thoughtful comments on our manuscript. We recognize the manuscript is long but feel that to have a complete discussion, the various components of the paper that have been included here are necessary. We will consider this feedback on future manuscripts. As per your suggestions below, some figure captions have been edited. The ACE-FTS retrieval does not routinely produce averaging kernels so these are not available for the analysis. Also, a discussion of the effects of measurement uncertainty has been added to the JPDF discussion (detailed below).

RC: Similarly, the discussion of differences in stratosphere-troposphere intrusions / extrusions mentions that resolution might factor into this comparison. At least for the detection of such structures in the data, this can be accounted for by removing small scales from the satellite data using a low-pass filter. Then the two datasets are nominally at the same resolution. This would however not address that the simulation of cut-off systems is fundamentally sensitive to numerical diffusivity in the model, causing reduced incidences of such systems.

AR: Philosophically, we have tried to approach the measurement-model comparisons in this study as directly as possible, knowing that the model is always representing a smoother version of reality than ACE-FTS does. We decided to not filter the measurements in any way so that the differences due to small scale features could be identified. If we were to apply a low-pass filter to the observations, we would not be able to answer the questions we posed. In addition to this, we would not be able to apply a three-dimensional low-pass filter on the ACE-FTS profiles (which would be

the only appropriate way to do such a comparison since the model fields are effectively smoothed across all three dimensions). The point about the finite resolution of the model is well taken and should be mentioned, but to start to modify the observations seems like a slippery slope.

P18L13-14 has been changed to reflect this sentiment:

Original sentence: "Understanding how CMAM30HR simulates this exchange assists in the interpretation of mixing effectiveness in the model and the impact of its vertical resolution on the comparisons."

New sentence: "Understanding how CMAM30HR simulates this exchange assists in the interpretation of mixing effectiveness in the model and the impact of its finite resolution on the comparisons."

RC: Regarding the differences in age-of-air between the free-running and nudged version of the model, my impression is that this is partly caused by mass non-conservation in the nudging fields, whereby artificial divergence caused by relaxation towards reanalyses causes noise in the vertical motion fields. The effect of this might be increased numerical diffusion and a reduced age. Since CMAM is based in a spectral dynamical core, one could consider, in separate experiments, to only nudge divergence or only vorticity, to try to control this behaviour.

AR: This is an interesting idea that we will pass along to the CMAM development team.

RC: On the whole, the above amounts to a recommendation to publish after a minor revision.

Minor comments:

RC: P3L3: This sentence reads a little awkwardly – modelling and observations are independent activities. How about "The BDC is well characterized in models but remains poorly constraint in obs" or so?

AR: We agree that this sentence should be rephrased.

Original sentence: "Despite significant progress in modelling, the BDC has been poorly constrained by observations (Butchart, 2014)."

New sentence: "The BDC is well characterized in models but remains poorly constrained by observations (Butchart, 2014)."

RC: P7L10: It's certainly possible to rescale the fields to construct approximations for the other tracers. But this requires further assumptions.

AR: We chose to use a parallel set of halocarbon species with adjusted boundary conditions because we felt the assumptions that would be required to construct approximations would complicate the interpretations of the measurement-model comparisons. Given the time-varying relative contributions of individual halocarbons to a particular model species in the troposphere, as air parcels enter the stratosphere and air parcels of different ages are mixed together, to untangle the contribution would require assumptions about both the mean age and the full age spectrum. It was felt that such an approach would introduce significant uncertainties.

Original sentence: **"Because of the time-varying contribution of the individual halocarbons to the tropospheric concentration of the model species, it is not possible to re-scale the concentration of the model species to recover a concentration that could be compared with observations."**

New sentence: "Because of the time-varying contribution of the individual halocarbons to the tropospheric concentration of the model species, the numerous assumptions that would be required to rescale the model species concentration to recover a concentration that could be compared with observations would introduce significant uncertainties."

RC: P7L20: Worth mentioning / discussing Meinshausen et al., Geosci. Model Dev., 2017 here. They have constructed boundary conditions for CMIP6 simulations that follow very similar ideas.

AR: The following sentence has been added after the sentence on P7L20:

"The application of hemispherically-defined lower boundary conditions based on observations is consistent with the proposed approach for the upcoming sixth phase of the Coupled Climate Model Intercomparison (CMIP6) project (e.g. Meinshausen et al., 2017)."

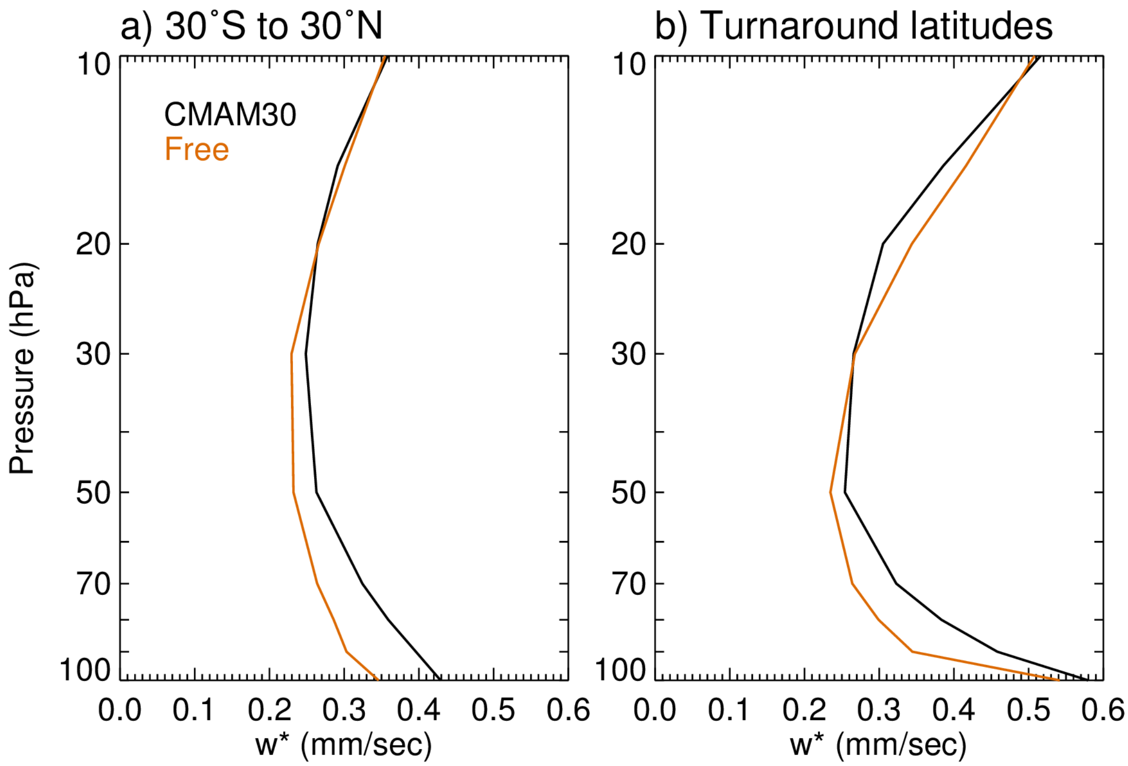
RC: P8L6: It remains a little unclear to me how you can have systematic differences between the LBCs used to constrain the simulations and the long-term observations, when the obs were used to construct the LBCs.

AR: The halocarbon LBCs were hemispheric averages obtained from the HATS network. While the data from the individual stations contributed to the hemispheric average, there are latitudinal dependencies in the surface measurement data. These comparisons indicate that differences exist but they are very small. The N_2O comparisons exhibit larger differences because the LBC was a global average.

RC: P8L27ff: Perhaps not drag but noise in the vertical motion. The w fields in the nudged and free-running model might show some differences.

AR: We agree that differences in the vertical motion between the free running and nudged simulations could certainly be contributing to the differences in the residual circulation. A figure has been provided below to demonstrate the differences in *w in the tropical region. Our aim with 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 be a completely separate study and is well outside of the subjects addressed here. 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 moment and are outside of the scope of the paper.



RC: P14L10: Perhaps insert "annual-mean" and some time information here (which period does the average represent?) Likewise in the caption, here and elsewhere.

AR: Annual-mean has been added to this line as well as the Fig.6 caption.

Original sentence: "The zonally-averaged distribution of N_2O is presented in Fig. 6.".
New sentence: "The zonally-averaged annual-mean distribution of N_2O is presented in Fig. 6.".

Original line in Fig 6 caption: "Zonally averaged latitude-altitude distributions of N_2O."
New line in Fig 6 caption: "Zonally-averaged annual-mean latitude-altitude distributions of N_2O."

RC: P20L3: Here's where the above comment on model resolution applies. The key difference is not that the two fields are at different resolutions (that could be easily fixed) but that the finite resolution of the model leads to differences in the formation and lifetimes of the cut-off systems.

AR: We agree with the reviewer that the capability of a relatively low-resolution model such as CMAM (at T47) to correctly model the dynamical evolution of cut-off systems that produce intrusions would be a consideration for a freely running model. In this analysis, however, the dynamical fields are nudged to reanalysis fields derived from a high resolution atmospheric model (T255 for ERA-Interim). Since the synoptic scale and larger (to T21) in CMAM30 are nudged to the ERA-Interim data, there should be, though admittedly it has not been shown, a good representation of the formation and lifetimes of cut-off systems. As for homogenizing the resolution of the observations and model, we note that ACE provides high resolution (up to ~3 km) vertical profiles but only about 15 profiles a day, so these profiles are widely spaced horizontally. While it would be possible to perform vertical smoothing of the observations, it is not clear whether the smoothed observations would be more comparable with the model, as the model fields are the result of a finite horizontal and vertical resolution.

Original sentence: "Therefore, it is unlikely that there is a physical mechanism or deficiencies in the model leading to the differences observed in Fig. 11 and the differences are primarily due to the model resolution."

New sentence: "Therefore, the differences are primarily due to the finite horizontal and vertical resolution of the model, which leads to differences in the representation of stratosphere-troposphere exchange events."

RC: P34: More detail in the caption please. Which species, which network, which measurement principle, why are there these systematic differences when the measurements had been used in constructing the LBC for the model?

AR: The boundary conditions for N_2O were not derived in a special manner for the CMAM30 run. The global averages of N_2O were based upon an older IPCC Assessment Report - the A1b scenario for the 4th Assessment Report and also used for CCMVal-2. The N_2O time series uses observations only up to 2000, then it becomes a projection so there are differences to be expected there. The CFC-11 and CFC-12 measurements used in the comparisons in Fig. 1 are an updated version compared to the data that was used for the boundary conditions in the CMAM30 simulations. The minimal differences observed are due to updated values in the observations. Further detail has been added to the caption of Figure 1.

Original Figure 1 caption: "Comparison of CMAM30HR run to surface measurements, relative differences calculated as the site subtracted from the CMAM30HR simulation, divided by the average of the two, as described in the text. The differences and the uncertainties included are the mean and standard deviation of relative differences over the time series."

New Figure 1 caption: "Comparison of CMAM30HR simulations of CFC-11 (blue x), CFC-12 (green diamond), and N_2O (black circle) to the HATS surface flask network of measurements at various locations around the world. Locations of measurement sites are indicated by latitude. Relative differences are calculated as the difference between the concentration at the surface site and the lowest model layer of the nearest neighbor gridbox to the site in the CMAM30HR output,

divided by the measured concentration. The relative differences were calculated based on the monthly averaged observations and simulations. Shown here are the mean of the differences between May 2004 and June 2010 and the error bars indicate one standard deviation of the mean of the relative differences over the time period."

RC: P46: Here's where I think measurement uncertainties make this a skewed comparison. The model output would ideally be folded with the averaging kernels and a-priori assumptions used in the retrievals of the ACE-FTS measurements before comparison with those measurements.

AR: We agree that it would be ideal to incorporate averaging kernels and a-priori assumptions in measurement-model comparisons. However, we are unable to do this because there are no averaging kernels available for the ACE-FTS dataset because we do not use optimal estimation in the retrieval process, and so we have no averaging kernels.

To address the measurement uncertainty concerns, the following text has been added to the end of section 3.1.2:

"Hegglin and Shepherd (2007) have shown the impact of ACE-FTS measurement uncertainties in joint PDFs by comparing the full model output, subsampled model output, and ACE-FTS measurements. They found that there was larger variability in the ACE-FTS joint PDFs compared to those of the subsampled CMAM output."

<u>Response to interactive comment on "Assessing stratospheric transport in the CMAM30 simulations using ACE-FTS measurements" by Felicia Kolonjari et al. from Referee #2</u>

RC = Reviewer Comment
AR = Author Response

RC: This manuscript describes a detailed set of comparisons between specified dynamics simulations with the CMAM chemistry climate model and satellite-based observations of stratospheric long-lived tracers, with the goal of identifying discrepancies and therefore errors in the simulated stratospheric dynamics. The topic is suitable for ACP, and the conclusions reached by the authors are generally reasonably well supported by the analysis presented.

AR: The authors thank Reviewer #2 for their helpful comments on the manuscript. We have addressed both the general and specific comments below.

General comments:
RC1. I have no doubt that the "advanced" comparison technique–which samples the model data along the actual line of sight of the satellite instrument–is the best way of minimizing sampling errors in the CMAM vs. ACE-FTS comparison. But, the explanations given in Sec. 3.3 don't quite make sense to me. If the difference between the advanced and intermediate methods is so small, as shown in Fig. 4c, this implies that the "line of sight" sampling is actually making very little difference to the sample means. The difference between advanced and basic (fig 4b) is much larger, which means that the most important source of sampling error has to do with a bias in the distribution of samples within each 5deg latitude bin (which is fixed by performing the 2D horizontal interpolation rather than using the closest neighbor gridpoint). It's possible these issues would be easier to sort out if Fig 4 showed differences between the 3 methods and the full model sampling, rather than differences between the advanced sampling and the other sampling methods. Or perhaps BASIC-FULL, INTERMEDIATE-BASIC, and ADVANCED-INTERMEDIATE. In any case, I believe the logic of the explanations here could be sharpened.

AR: We chose to represent this comparison as the difference between the Advanced method and the others to illustrate the differences between what would be "observed" by ACE-FTS compared to the full model and common sampling techniques. For a model at T47 resolution, it is true that the intermediate and advanced sampling are very similar. However, for models with finer spatial resolution, the differences would be more apparent. Therefore, we have used the advanced method for ACE-FTS/model comparisons. The text has been clarified to minimize the expectation for a comparison of each method to the full model output.

Original sentence: "To determine the impact of sampling the model output at varying levels of detail, three methods were tested and compared to the full model output (the CHAM30HR output at all latitudes and longitudes for each 5 ° latitude bin) between June 2004 and May 2010."

New sentence: "To determine the impact of sampling the model output at varying levels of detail, three methods were tested by sampling the full model output (the CMAM30HR output at all latitudes and longitudes for each 5° latitude bin) between June 2004 and May 2010."

RC2. No doubt there is much more information given in the Ray et al. (2016) paper, but Sec. 6 requires a little more guidance on the set up of the TLP simulations. At some point, in passing, we learn that there were 480 TPL simulations, but it is not said how these simulations differ; presumably certain input assumptions are varied in the different simulations, but not, it seems, w and epsilon directly. Also, the terms w* and epsilon should be better defined. At some point, w* is introduced as a mean tropical upwelling, but it is later used to quantify vertical motion in the extratropics. It's also not really clear if epsilon is a prognostic or diagnostic variable, and how it depends on height, latitude, time, etc.*

AR: To maximize clarity and minimize additional text, the second paragraph of section 2.3 has been moved to the beginning of section 6 and the first sentence of section 6 has been moved to the end of section 2.3.

The sentence moved to section 2.3 from the first sentence of section 6 has been changed to: "The TLP model is used here to identify the changes to the CMAM30HR tropical upwelling and effective mixing that may improve the comparisons between ACE-FTS and CMAM30HR. However, this tool does not identify a specific mechanism but it can isolate seasons and regions in which changes are required."

The following sentences were added to the end of the paragraph moved (now the first paragraph of section 6): "There were 480 simulations initialized with different combinations of w^* (velocity of tropical upwelling) and ε (the mixing efficiency) settings. The fraction of the CMAM30 w^* used to initialize the TLP model ranged from 0.20 to 1.24 and the ε ranged from 0.18 to 1.50. In each TLP simulation run, the relationship between mean circulation and mixing is constrained by the vertically-averaged mixing efficiency [Ray et al., 2016]. The mixing efficiency in the TLP model is defined as $\varepsilon = \alpha / \lambda\tau$, where α is the ratio of tropical to extratropical mass, λ is the rate of the mean circulation influence, specifically the mass flux out of the tropics due to mean circulation, and τ is the mixing time or time scale for mass flux between the tropics and extratopics [Ray et al., 2014; Garny et al., 2014]."

RC3. I have trouble following the logic from Figure 15 to 16. Figure 15 seems to show that best agreement with the ACE-FTS measurements is achieved running the TLP model with parameter settings which produce the smallest w and largest epsilon values (at least in Fig 15a,b,c). In fact, it seems that the range of TLP simulations is not large enough to find the actual best agreement with ACE-FTS–a point which could be discussed. But then, in Figure 16, it is implied that e.g., best agreement with ACE-FTS is produced with no significant change in w* values in the tropics. Something seems inconsistent here.*

AR: To clarify the link between changes in CMAM that may improve the comparisons to ACE-FTS, we have updated Fig. 16 to include the individual contributions of both CFC-11 and CFC-12. The thresholds used for all latitude ranges in the new Fig. 16 have also be reduced to 0.2 (as was used for the tropics) to make the regional comparisons more consistent. Deciding to use a more

restricted but consistent threshold leads to no result in some seasons for some geographic regions, perhaps suggesting that the range of *w and *ε* values should be extended.**

RC4. For many of the difference contour plots, it would be helpful if the colorbars were chosen such that positive differences could be more easily differentiated from negative differences. An example is Fig 6, where it is very difficult to know whether the CMAMACE differences in the UTLS are positive (like the middle stratosphere differences) or negative.

AR: The colorbars of the difference contour plots have all been changed to a red/white/blue scheme where white is zero.

Specific comments

RC P1, l11: "The model consistently: : :" could be taken out of context–this conclusion is specific to the trace gases examined in this study (and probably wouldn't apply to ozone, for example).

AR: We have clarified this statement.

Original sentence: "The model consistently overpredicts tracer concentrations in the lower stratosphere, particularly in the Northern Hemisphere winter and spring seasons."

New sentence: "The model consistently overpredicts tracer concentrations of CFC-11, CFC-12, and N_2O in the lower stratosphere …"

RC P1, l14: the "too little isentropic mixing" should probably be connected if possible to a height or range of heights.

AR: We have clarified this statement.

Original sentence: "In particular, the CMAM30 simulation exhibits too little isentropic mixing in the June-July-August season."

New sentence: "In particular, the CMAM30 simulation exhibits too little isentropic mixing in the tropical lower stratosphere during the June-July-August season."

RC P2, l1: I'm not sure this is the only reason for the increase in interest in stratospheric transport–and it's a bit of a chicken and egg problem.

AR: We have included a reference to Butchart (2014) to support this statement.

Original sentence: "Interest in stratospheric transport has increased over the last 20 years as a result of significant developments in stratosphere-resolving general circulation models (GCMs) and chemistry-climate models (CCMs) (e.g., Pawson et al., 2000; Eyring et al., 2005; SPARC-CCMVal, 2010; Gerber, 2012)."

New sentence: "As highlighted by Butchart (2014), interest in stratospheric transport has increased over the last 20 years as a result of significant developments in stratosphere-resolving general circulation models GCMs (e.g. Pawson et al., 2000; Gerber, 2012) and chemistry-climate models (CCMs) (e.g. Eyring et al., 2005; SPARC-CCMVal, 2010)."

*RC P2, l4: *Accurate* projections rely on good models.*

AR: Accurate has been added.

Original sentence: "Projections of stratospheric ozone and climate rely on the ability of these models to simulate stratospheric transport and chemistry."

New sentence: "Accurate projections of stratospheric ozone and climate rely on the ability of these models to simulate stratospheric transport and chemistry."

RC P2, l5: Definitely the distribution of long-lived trace gases depends on the BDC…for short-lived species it may not have that much influence.

AR: The term "long-lived" has been added to the sentence.

Original sentence: "It is clear that the transport of chemical tracers will be impacted by changes in the BDC, which will in turn influence ozone recovery projections, lifetimes of ozone depleting gases, and mass exchange between the troposphere and stratosphere (Butchart, 2014)."

New sentence: "It is clear that the transport of long-lived chemical tracers will be impacted by changes in the BDC, which will in turn influence ozone recovery projections, lifetimes of ozone depleting gases, and mass exchange between the troposphere and stratosphere (Butchart, 2014)."

RC P3, l14: "has" or "have"? The word choice depends on whether the models project changes in the BDC (plural) or evidence of those changes (singular).

AR: Based upon the feedback of Reviewer #3, this sentence has been changed.

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 P4, l31: ACE-FTS measurements have high vertical resolution, not the instrument itself.

AR: That is correct, the ACE-FTS retrieved profiles have a high vertical resolution. The sentence has been clarified.

Original sentence: "ACE-FTS is ideal for studying 30 the vertical structure of constituent gases from cloud tops to 100 km; it is particularly useful in the upper troposphere and lower stratosphere because of its high vertical resolution (Hegglin et al., 2008)."

New sentence : "ACE-FTS is ideal for studying the vertical structure of constituent gases from cloud tops to 100 km; the retrieved profiles are particularly useful in the upper troposphere and lower stratosphere where the vertical resolution is approximately 3 km (Hegglin et al., 2008)."

RC P8, l2: For comparisons of model to measurements, it's probably more intuitive to treat the measurements as the truth, and show relative differences of the model with respect to the measurements, rather than the mean of the measurements and model. It's of course not a big deal as long as it is clear how the calculation is being done, but I feel the simpler the calculation, the easier it is to interpret.

AR: The comparisons between measurements and model output of tracer concentrations throughout the paper used the same method of calculating comparisons with respect to the mean of the measurement and model. The authors agree that the comparison would be more intuitive with respect to the measurements. Therefore, all relative difference calculations have been updated throughout the paper and are now relative to the measurement.

RC P8, l8: this statement of significance applies to the 1 sigma confidence level. For 2 sigma, I guess all differences would be not significantly different from zero.

AR: True, we have clarified the sentence to reflect this detail.

Original sentence: "Over the time period compared, CMAM30HR appears to overpredict CFC-11 at all HATS sites while the CFC-12 comparisons are not significantly different from zero for all but two sites in the Southern Hemisphere."

New sentence : "Over the time period compared, CMAM30HR appears to overpredict CFC-11 at all HATS sites while the CFC-12 comparisons are not significantly different from zero within one standard deviation for all but two sites in the Southern Hemisphere"

RC P9, l24-26: Apparent contradiction between "increased isentropic mixing" and "slower shallow branch".

AR: You are correct. There is a problem with this description. Relatively old mean age in the lowermost stratosphere is due to either a relatively strong upper branch of the BDC compared to the lower branch, and/or less isentropic mixing since that process brings young tropospheric air

into the lowermost stratosphere. In the LMS, increased mixing brings in young air because it's primarily tropospheric.

Original sentences: "... CMAM30 clearly has older air in the extra-tropical lowermost stratosphere. It is potentially caused by either stronger downwelling of the older air from above, consistent with a stronger BDC, or increased isentropic mixing of tropospheric air from lower latitudes (e.g. Hegglin and Shepherd, 2007). Therefore, the differences in age appear to suggest a slower shallow branch or a faster deep branch of the BDC."

New sentences: "... CMAM30 clearly has older air in the extra-tropical lowermost stratosphere. It is potentially caused by either stronger downwelling of the older air from above, consistent with a stronger BDC, or reduced isentropic mixing of tropospheric air from lower latitudes (e.g., Hegglin and Shepherd, 2007). Therefore, the differences in age appear to suggest a slower shallow branch or a faster deep branch of the BDC."

RC P12, l19: the end of this sentence could be misconstrued, as of course there has been vertical interpolation applied in the translation to the vertical levels of the ACEFTS retrievals. I would suggest to remove this last part, and write ": : : location of the tangent points with altitude".

AR: We have made the change suggested.

Original sentence: "... no consideration of the variation in geographical location of the tangent points ..."

New sentence: "… no consideration of the variation in geographical location of the tangent points above or below 30 km."

RC P13, l25: But the INTERMEDIATE sampling technique is also limited to the 30 km tangent altitude, and the differences to the advanced method are much smaller. Indeed, by construction, differences between a method using the variable tangent heights and one using only the 30 km tangent height should be zero at 30 km (which appears to be the case in the advanced-intermediate comparison). Therefore, the differences between basic and advanced, which are strongest around 30 km, cannot be due to the line of sight sampling.

AR: The basic sampling uses a 30 km tangent height but the nearest model column and the intermediate sampling produces a profile by computed the bilinear interpolation of the concentrations of the nearest 4 grid boxes at each point along the column. It is the intermediate sampling that we should expect to have a zero difference with the advanced sampling at 30 km. This is the case in Fig. 4c.

RC P15, l26: how are large disagreements in the north polar region so confidently connected to problems with tropical upwelling? Could this not be an issue with mixing across the polar vortex?

AR: The sentence has been changed.

Original sentence: "The large disagreements in the north polar region during winter and spring indicate that the upwelling portion of the BDC across the different seasons is not well characterized by CMAM30HR."

New sentence : "The large disagreements in the north polar region during winter and spring indicate that the downwelling portion of the BDC across the different seasons is not well characterized by CMAM30HR."

RC P22, l17: I would avoid the term "mixing levels" as it could be taken as meaning isentropic surfaces.

AR: This has been changed to mixing efficiency throughout the manuscript, which is more accurate.

RC P22, l19: Is this result based on looking at the best agreement between CMAM and ACE-FTS under the natural variability of the CMAM simulations? Located here within the discussion of the TPL, it comes across a little as one has tuned CMAM, which I think is not the case. Also, "a reduction from the fitting estimate" is unclear to me, is this the best fit of the TPL parameters to the CMAM climatology?

AR: The TLP model was tuned to the CMAM30 output before it was used to test a range of w^* and ε values. Line 16 of page 22 states "... the TLP model was tuned to be representative of CMAM30HR by fitting estimates of the mean tropical upwelling (w^*) and mean mixing levels (ε) …". The reduction from the fitting estimate refers to the change required in the CMAM30HR run based on the tuned TLP model.

RC P22, l21,22: First sentence implies best agreement when epsilon is increased–which, based on previous description I take to mean a mixing rate should be increased. But the following sentence says mixing "times" should be increased, which actually means rates should decrease. Some clarification here would be useful.

AR: We have decided to revise the paragraph to clarify our meaning.

Original paragraph 2 of section 6:

"A reduced mean circulation would likely correspond to less mixing and longer mixing times. Mixing levels are defined in the TLP by the ratio of horizontal mixing mass flux to horizontal mean mass flux scaled by the width of the tropical pipe region (Garny et al., 2014). Ray et al. (2016) found that the CMAM30HR simulations best match the ACE-FTS measurements when the w^* is between 0.27 mm/s and 0.32 mm/s (a reduction from the fitting estimate of 0.4 mm/s) and ε ranges from 0.7 to 1.2 (an increase from the fitting estimate of 0.55). Based on the in-mixing time profiles shown by Ray et al. (2016), it is apparent that the CMAM30HR mixing times need to increase to slow down the mixing at all levels, although the differences are only significant in the lower part of the stratosphere, below 20 km, and above 24 km. These are physically consistent changes since the mean circulation is driven by wave breaking, which also causes mixing between the tropics and extratropics."

New paragraph 2 of section 6:

"Ray et al. (2016) found that the CMAM30HR simulations best match the ACE-FTS measurements when the *w is between 0.27 mm/s and 0.32 mm/s (a reduction from the fitting estimate of 0.4 mm/s) and ε ranges from 0.7 to 1.2 (an increase from the fitting estimate of 0.55). Ray et al (2016) found that since ε is inversely proportional to both λ and τ, there is a compensating effect with changes in *w** (λ) or τ. For the CMAM30HR changes derived, *w** needed to be slowed down significantly below 20 km, and above 24 km. For constant ε that would result in larger τ (less mixing). However, Ray et al (2016) found that ε also needed to be increased so there needed to be more mixing than would result from slower *w** and constant τ, but not enough of an increase in ε so that the mixing times were less (more mixing) than CMAM30HR has currently. With the increase in ε, mixing times are reduced but still longer than the current CMAM30HR mixing times."**

RC P23, l5: Figure 15 is quite dense, and really could use better description in the main text and figure caption. The "level of agreement" between CMAM and ACE-FTS needs to be explained fully, what kind of quantity is this? It took me some time to determine that the white-to-black shading and the white isolines were describing the same quantity. Also, there are bluish boxes which are barely detectable in the plots, are these the "agreement matrices"? As mentioned in the general comments, these agreement boxes don't seem consistent with the ACE-FTS "agreement" scale.

AR: We have clarified the text to make it clear that the white isolines are the same quantity as the grey color contours but we are not sure what the blueish boxes the reviewer is referring to. We do recall that on printing the paper sometimes, weird blueish boxes appeared on these plots. Please refer to the electronic version of the paper and check to see if the blue boxes are there as well.

Original sentence: "The white contours illustrate the agreement isolines."

New sentence: "The white-to-black shading is reinforced by white contours of the same quantity to illustrate the comparison more clearly."

RC P23, 19: How are these thresholds chosen? 0.65 seems like a rather lenient agreement threshold, since the agreement values shown in Fig 15 go as low as 0.1.

AR: As noted above, the thresholds have been changed to be 0.2 for both CFC-11 and CFC-12 in all regions (tropics and extratropics). This is reflected in the updated Figure 16.

RC P23, l20: what chemical species is Fig 16 based on?

AR: CFC-11 and CFC-12 were both used in Fig 16. Figure 16 has been changed, as previously noted.

RC P24, l32: If mixing and the meridional residual circulation are both driven by Rossby wave breaking, it is hard to see how insufficient mixing could be the cause of a too-rapid BDC.

AR: Our results show that the CMAM30 circulation needs to be slower but the mixing efficiency, *ε*, needs to increase. As explained in response to previous comments, increasing ε does not necessarily mean increasing mixing.

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."

Assessing stratospheric transport in the CMAM30 simulations using ACE-FTS measurements

Felicia Kolonjari[1], David A. Plummer[2], Kaley A. Walker[1], Chris D. Boone[3], James W. Elkins[4], Michaela I. Hegglin[5], Gloria L. Manney[6,7], Fred L. Moore[8,9], Diane Pendlebury[10], Eric A. Ray[8,9], Karen H. Rosenlof[8], and Gabriele P. Stiller[11]

[1]Department of Physics, University of Toronto, Toronto, Canada
[2]Climate Research Division, Environment and Climate Change Canada, Victoria, Canada
[3]Department of Chemistry, University of Waterloo, Waterloo, Canada
[4]Global Monitoring Division, NOAA Earth System Research Laboratory, Boulder, USA
[5]Department of Meteorology, University of Reading, Reading, UK
[6]NorthWest Research Associates, Socorro, USA
[7]Department of Physics, New Mexico Institute of Mining and Technology, Socorro, New Mexico, USA
[8]Chemical Sciences Division, NOAA Earth System Research Laboratory, Boulder, USA
[9]Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, USA
[10]Air Quality Research Division, Environment and Climate Change Canada, Toronto, Canada
[11]Institute of Meteorology and Climate Research, Karlsruhe Institute of Technology, Karlsruhe, Germany

Correspondence to: Kaley A. Walker (kaley.walker@utoronto.ca)

Abstract.

Stratospheric transport in global circulation models and chemistry-climate models is an important component in simulating the recovery of the ozone layer as well as changes in the climate system. The Brewer-Dobson circulation is not well constrained by observations and further investigation is required to resolve uncertainties related to the mechanisms driving the circulation. This study has assessed the specified dynamics mode of the Canadian Middle Atmosphere Model (CMAM30) by comparing to the Atmospheric Chemistry Experiment Fourier Transform Spectrometer (ACE-FTS) profile measurements of CFC-11 (CCl_3F), CFC-12 (CCl_2F_2), and N_2O. In the CMAM30 specified dynamics simulation, the meteorological fields are nudged using the ERA-Interim Reanalysis and a specified tracer was ~~used~~ employed for each species, with hemispherically-defined surface measurements used as the boundary condition. A comprehensive sampling technique along the line-of-sight of the ACE-FTS measurements has been ~~employed~~ utilized to allow for direct comparisons between the simulated and measured tracer concentrations. The model consistently overpredicts tracer concentrations of CFC-11, CFC-12, and N_2O in the lower stratosphere, particularly in the Northern Hemisphere winter and spring seasons. The three mixing barriers investigated, including the polar vortex, the extratropical tropopause, and the tropical pipe, show that there are significant inconsistencies between the measurements and the simulations. In particular, the CMAM30 simulation exhibits too little ~~isentropic~~ mixing in the ~~June-July-August~~ tropical lower stratosphere during the June-July-August season.

1 Introduction

Interest As highlighted by Butchart (2014), interest in stratospheric transport has increased over the last 20 years as a result of significant developments in stratosphere-resolving general circulation models (GCMs) (e.g., Pawson et al., 2000; Gerber, 2012) and chemistry-climate models (CCMs) (e.g., Pawson et al., 2000; Eyring et al., 2005; SPARC-CCMVal, 2010; Gerber, 2012). Projections (e.g., Eyring et al., 2005; SPARC-CCMVal, 2010). Accurate projections of stratospheric ozone and climate rely on the ability of these models to simulate stratospheric transport and chemistry. The distribution of long-lived trace gases in the stratosphere is primarily controlled by the Brewer-Dobson Circulation (BDC), which is generally characterized by tropospheric air entering the stratosphere in the tropics, poleward transport, and descent in the midlatitude and polar regions of the winter hemisphere (e.g., Plumb, 2002; Butchart, 2014, and references therein). The BDC describes the primary features of the stratospheric circulation, based upon a conceptual model proposed to explain observations of ozone and water vapour (Dobson et al., 1929; Brewer, 1949; Dobson, 1956). Over the last decade, BDC modeling studies have reached a consensus regarding how the stratospheric circulation will respond the response of stratospheric circulation to changes in anthropogenic climate forcing has been studied (Butchart, 2014). It is now understood that the residual circulation as well as and quasi-isentropic mixing are key factors to understanding the structure of the BDC (McLandress et al., 2011; Butchart, 2014; Abalos et al., 2015; Ploeger et al., 2015a, b; Oberländer-Hayn et al., 2016). Plumb (2002) and Shepherd (2007) state that two-way mixing between the tropics and extratropics and the rapid stirring of air parcels are important components of stratospheric transport. The influence of planetary waves on stirring is predominately in the winter midlatitude surf zone (McIntyre and Palmer, 1983, 1984) but synoptic-scale wave activity occurs throughout the year in the subtropical lower stratosphere and its influence can extend upwards to 25 km (Haynes and Shuckburgh, 2000). It has been argued by Shepherd and McLandress (2011) Shepherd and McLandress (2011) argued that the greenhouse gas induced warming of the climate system has led to an upward displacement of the critical layers for wave breaking. Subsequently, it has been suggested that the BDC changes are characterized more by a vertical lifting of rather than an acceleration of the meridional circulation (Oberländer-Hayn et al., 2016).

Plumb (2002) and Birner and Bönisch (2011) identified two distinct pathways within the BDC. These are the “deep branch”, defined as the poleward transport in the winter hemisphere extending into the middle and upper stratosphere, and the “shallow branch(es)”, defined as multiple pathways of faster poleward transport that are observed in both hemispheres throughout the year and are generally restricted to the lower-to-middle stratosphere. It is likely that the shallow branches are driven by Rossby-wave pumping on a synoptic scale (Plumb, 2002; Butchart, 2014). As part of the Climate Chemistry Model Validation (CCMVal) project, Lin and Fu (2013) investigated simulated changes in the BDC by considering three branches separately; the transition, shallow, and deep branches. They found that changes in the transition and shallow branches of the BDC were consistent with the increase of greenhouse gas concentrations and the trends were associated with changes in subtropical jets and tropical upper tropospheric temperatures, which is also consistent with the mechanism described by Shepherd and McLandress (2011). The acceleration of the deep branch is consistent with that of the transition and shallow branches but is seasonally modulated by changes in ozone concentrations with the exact mechanisms yet to be determined (Lin and Fu, 2013). The mechanisms that lead to an acceleration or deceleration in the deep branch remain unresolved (Shepherd and McLandress, 2011; Lin and Fu, 2013). Observational evidence seems to indicate a deceleration in the deep branch (Engel et al., 2009; Hegglin et al., 2014). However, there is a fair degree of confidence that the acceleration of the shallow branch is likely since it is driven by the vertical lifting mechanism proposed by Shepherd and McLandress (2011) that is related to large-scale changes and also has been diagnosed from changes in stratospheric constituent distributions (Hegglin et al., 2014).

Despite significant progress in modeling, the BDC has been The BDC is well characterized in models but remains poorly constrained by observations (Butchart, 2014). A direct comparison to determine how the BDC and quasi-horizontal mixing combine to produce the distribution of long-lived tracers with tropospheric sources is not possible. Therefore, a number of observational techniques have been used to investigate stratospheric transport characteristics, such as age of air diagnostics (Stiller et al., 2008; Engel et al., 2009), tropical lower stratosphere ascent rates (Mote et al., 1996; Niwano et al., 2003), and descent rates in the Antarctic polar vortex (Abrams et al., 1996b; Allen et al., 2000; Kawamoto and Shiotani, 2000), and in the Arctic polar vortex (Abrams et al., 1996a; Greenblatt et al., 2002; Ray et al., 2002; Greenblatt, 2003). Except for Stiller et al. (2008), these observations do not provide global seasonally-resolved quantitative estimates of the BDC (Butchart, 2014). Thus, it has proven Butchart (2014) identified that it is difficult to deduce changes in the strength of the BDC using available measurements (e.g., Engel et al., 2009; Diallo et al., 2012; Seviour et al., 2012; Stiller et al., 2012; Haenel et al., 2015).

Fundamental questions remain as to the mechanisms driving the stratospheric circulation because there has been evidence of changes Recently, Hardiman et al. (2017) determined that, for the BDC, 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 that has not been projected by models (e.g., Butchart, 2014; Mahieu et al., 2014). if it is based on less than 12 years of data (Hardiman et al., 2017).

It is clear that the transport of chemical tracers will be impacted by changes in the BDC, which will in turn influence ozone recovery projections, lifetimes of ozone depleting gases, and mass exchange between the troposphere and stratosphere (Butchart, 2014). Understanding how the structure of the BDC will change depends greatly upon the ability to simulate its current behaviour. This is typically assessed by investigating how capable a model is at simulating long-lived tracerstracer concentrations (Jin et al., 2005; Allen et al., 2009; Park et al., 2013; Pendlebury et al., 2015); in particular, assessing the characteristics of the simulated tracers such as concentrations and behaviour at the mixing barriers in the stratosphere (i.e. the polar vortex edge, the extratropical tropopause, and the tropical pipe).

In this study, measurements of the long-lived chlorofluorocarbons, CCl_3F (CFC-11) and CCl_2F_2 (CFC-12), and N_2O from the Atmospheric Chemistry Experiment Fourier Transform Spectrometer (ACE-FTS) are used to evaluate the specified dynamics simulation mode of the Canadian Middle Atmosphere Model (CMAM30). Using these global measurements, areas in which simulated tracers agree with observations and where improvements are needed have been investigated. Since ACE-FTS measurements are vertically resolved, these data are useful for testing model simulations (e.g., Hegglin and Shepherd, 2007; Manney et al., 2009; Strahan et al., 2011); however, care must be taken in the methods used for these comparisons. In addition to comparison methods, external tools are useful to for the interpretation of differences between CMAM30 and ACE-FTS. An idealized stratospheric model, the Tropical Leaky Pipe (TLP) model, described by Ray et al. (2016) has been used to test factors contributing to the differences observed between ACE-FTS and CMAM30.

Many limb-viewing satellite missions, including ACE, are contributing data to the SPARC Data Initiative, whose purpose is to compile and assess a repository of climatologies for comparison with model output (Hegglin et al., 2013; Tegtmeier et al., 2013; Neu et al., 2014; Tegtmeier et al., 2016). Some of the instruments involved in the initiative, such as ACE-FTS, do not cover all latitudes and altitudes each month. This matter has been the subject of recent studies (Toohey and von Clarmann, 2013; Toohey et al., 2013; Millán et al., 2016). They have shown that the impacts of sampling patterns of various instruments must be considered when performing comparisons because the model simulates data points that are evenly spaced throughout each latitude range in the zonal mean at each pressure level while the measurements can represent a subset of that domain.

This work addresses issues related to climatological comparisons in several ways: first, by using a nudged version of the CMAM; second, by breaking out the species of interest from the halocarbon family arrangement in the model; and third, by sampling the model output along the individual measurement profile pathways through the atmosphere. The latter addresses sampling issues identified by Toohey and von Clarmann (2013), Toohey et al. (2013), and Millán et al. (2016). By isolating the model output in this way, the transport and chemical processes in CMAM can be evaluated. This type of assessment of CMAM simulations has not been possible until recently because of the typically free-running nature of the model simulations. The CMAM specified dynamics simulation has been investigated in a few recent studies: McLandress et al. (2014) evaluated the polar cap mesospheric transport and midlatitude mean zonal winds, and long term observational records of water vapour and ozone were also used to evaluate the CMAM30 run by Hegglin et al. (2014) and Shepherd et al. (2014), respectively. Additionally, Pendlebury et al. (2015) investigated the CMAM30 polar regions using satellite data, including ACE-FTS. Comparisons of the CMAM30 simulations to observations remain limited in their extent, a gap in which this work attempts to fill.

This paper is structured as follows: Section 2 describes the tools used in this study including measurements from ACE-FTS, CMAM30 simulations, and the TLP model simulations. Section 3 examines methods of sampling and comparison techniques and considers the impact of the sampling of ACE-FTS. Section 4 examines the measured and simulated zonally-averaged morphologies of CFC-11, CFC-12, and N_2O. Section 5 investigates the three barriers to mixing: the polar vortex, the extratropical tropopause, and the tropical pipe. In section 6, the TLP model is used to investigate changes in tropical upwelling and quasi-isentropic mixing required to allow the model to more effectively simulate the lower stratosphere. Finally, the results are summarized and discussed in Section 7.

2 Tools

2.1 ACE-FTS

The ACE mission on-board the Canadian satellite SCISAT was launched to investigate the distribution of upper tropospheric and stratospheric ozone with the goal to further our understanding of the chemical and dynamical processes that influence its behaviour (Bernath, 2017). ACE entered a circular low-earth orbit (650 km, 74° inclination), on 12 August 2003, to observe the Earth's atmosphere using the solar occultation technique (Bernath et al., 2005). The high resolution Fourier transform spectrometer, ACE-FTS, is the primary instrument on SCISAT. It has a high spectral resolution (0.02 cm^{-1}) and a spectral range of ~~2.2-13.3 μm (750-4400 cm^{-1}~~2.2-13.3 μm (750-4400 cm^{-1}) (Bernath et al., 2005). ACE-FTS does not require filters for its operation, which allows it to measure solar absorption spectra for dozens of atmospheric constituents simultaneously. ACE-FTS is ideal for studying the vertical structure of constituent gases from cloud tops to 100 km~~; it is~~. The retrieved profiles are particularly useful in the upper troposphere and lower stratosphere ~~because of its high vertical resolution~~ where the vertical resolution is approximately 3 km (Hegglin et al., 2008). The ACE-FTS products are derived from the solar absorption spectra measured and include the vertical profiles of temperature, pressure, and the concentration expressed as a volume-mixing ratio (VMR) for several dozen molecules of atmospheric interest over latitudes from 85° N to 85° S (Bernath et al., 2005). These data products are useful for the ~~study~~ development of climatologies (e.g Allen et al., 2009; Jones et al., 2012; Koo et al., 2017), trends (e.g., Brown et al., 2011), and lifetimes (e.g., Brown et al., 2013), among other applications (e.g., Hegglin and Shepherd, 2007; Hegglin et al., 2009; Brown et al., 2014; Hoffmann et al., 2014; Hegglin et al., 2014).

The version 3.0 ACE-FTS retrievals of CFC-11 (CCl_3F), CFC-12 (CCl_2F_2), and N_2O based on spectra recorded between June 2004 and May 2010 have been used throughout this work. A description of the retrieval process used for ACE-FTS is provided by Boone et al. (2005) and Boone et al. (2013). The earlier work details the retrieval process for version 2.2 of the data while the latter describes improvements that have been implemented for the recent versions. The CFC-11 retrieval ranges from 5 km to a maximum of 28 km in the tropical latitudes but is limited in the stratosphere at higher latitudes due to low concentrations. CFC-12 is retrieved between 5 km and 36 km, while the N_2O retrieval covers 5 km to 95 km. Due to the vertical limitation of the CFC-11 retrieval, this work focuses on the upper troposphere and lower stratosphere (5 km to 30 km).

The validity of these measurements has been investigated in several studies. Mahieu et al. (2008) compared both the ACE-FTS CFC-11 and CFC-12 v2.2 products to the FIRS-2 and MkIV Interferometer measurements. They found ACE-FTS to be approximately 10% lower than FIRS-2 between 12 km and 16 km in the case of CFC-11. Using a non-coincident technique, Velazco et al. (2011) found agreement to be better than 20% between ACE-FTS and MkIV over the range of 17 km to 24 km. The CFC-12 comparisons made by ~~Velazco et al. (2011) and Mahieu et al. (2008)~~ Mahieu et al. (2008) and Velazco et al. (2011) show a consistent difference with ACE-FTS approximately 10% lower than MkIV. Using a climatological validation approach within the SPARC Data Initiative to compare the CFC-11 and CFC-12 products of HIRDLS, MIPAS, and ACE-FTS, Tegtmeier et al. (2016) found excellent agreement in the lower stratosphere (up to 50 hPa) and increasing positive deviations above this level to around 20% from the multi-instrument mean. Strong et al. (2008) provided an extensive validation of the ACE-FTS v2.2 N_2O product using satellite, aircraft, balloon and ground-based FTIR measurements. Differences observed were typically within ~~+~~±15% (Strong et al., 2008). The work of Velazco et al. (2011) was consistent with the results of Strong et al. (2008). Waymark et al. (2013) compared the previously validated CFC-11 and CFC-12 v2.2 products with the v3.0 products used in this work and found that around 15 km there is a slight increase in CFC-11 in the new product, bringing it closer to the correlative measurements used in the studies described above. Similarly for CFC-12, Waymark et al. (2013) found an increase of ~~2-5~~2–5% between 6 km and 22 km. They also found an approximately 10% decrease in the concentration of N_2O above 35 km in v3.0, bringing the differences found in Strong et al. (2008) to within approximately ~~+~~±5~~%~~%.

A set of quality control flags for the ACE-FTS products on the 1-km retrieval grid are available (Sheese et al., 2015). The version 1.1 flags were applied to the data used in this work by removing profiles that contained a flag between 4 and 7, as recommended by Sheese et al. (2015). The method rejects a maximum of 6% of data over all the species retrieved. For CFC-11, CFC-12, and N_2O, 1.7%, 2.1%, and 4.3% of the data are rejected (Sheese et al., 2015). In addition to the quality control flags, derived meteorological products from GEOS 5.2.0, based on the techniques described in Manney et al. (2007), are used here along with the geographic location information to account for the geographic extent and meteorological context of ACE-FTS profiles.

2.2 Canadian Middle Atmosphere Model

2.2.1 The model

CMAM is a freely-running CCM based on an upwardly-extended version of the Canadian Centre for Climate Modelling and Analysis (CCCma) third-generation Atmospheric General Circulation Model (Beagley et al., 1997; Scinocca et al., 2008). The chemistry includes the O_x, HO_x, NO_x, ClO_x, and BrO_x catalytic cycles that control ozone in the stratosphere; the chemistry of N_2O, CH_4, and seven long-lived halocarbon species; and a representation of heterogeneous chemistry on background stratospheric sulphate aerosols and on polar stratospheric clouds (deGrandpré et al., 2000; Jonsson et al., 2004). CMAM has been used extensively to investigate the middle atmosphere and to study complex processes of the climate system (e.g., Austin et al., 2003; Vyushin et al., 2007; Plummer et al., 2010; McLandress et al., 2011). Results from the CMAM have also been assessed during two phases of the CCMVal project and, more recently, the Chemistry-Climate Model Initiative (CCMI). The extensive investigation of dynamical and chemical processes in CCMs that took place during CCMVal-2 are detailed in the SPARC report (SPARC-CCMVal, 2010). For the simulations used here, CMAM was run at a T47 spectral resolution, equivalent to approximately 3.75° x 3.75°3.75° x 3.75°, with 71 vertical levels topping out at 0.08 Pa (approximately equivalent to 95 km in altitude).

Due to the chaotic nature of the atmospheric circulation, free running models are unable to reproduce the day-to-day evolution of the atmosphere. Therefore, simulated fields, such as tracer concentrations, cannot be compared to observations directly or on a day-to-day basis. There has recently been an effort to circumvent this limitation by constraining the evolution of the circulation and temperatures in CCMs with fields from reanalysis data sets through the use of Newtonian relaxation (e.g., McLandress et al., 2014; Shepherd et al., 2014), known colloquially as 'nudging'. The ability to constrain the dynamical fields to follow the observations more closely reanalysis, the best approximation of reality, enables direct model-measurement comparisons of chemical tracers in the model by eliminating the internal variability in the simulated circulation. This type of specified dynamics simulation has been used here to allow for space and time matched comparisons of output from CMAM with ACE-FTS observations over the June 2004 to May 2010 period.

2.2.2 The specified dynamics simulation

The CMAM30, the specified dynamics version of the CMAM, which is referred to here as CMAM30, uses meteorological fields from the ERA-Interim reanalysis (Dee et al., 2011) to constrain the dynamical fields while the chemical fields are allowed to freely evolve. The model horizontal winds and temperature are nudged towards the ERA-Interim fields from the surface to 1 hPa. The six-hourly reanalysis data is linearly interpolated in time to produce fields for nudging at intermediate time steps. The relaxation is applied at only the synoptic scales and larger by constraining only wavenumbers up to 21. The application of nudging on only the large scales and the use of a relaxation time constant of 24-hours has been found to produce root-mean-square differences between the CMAM30 and the reanalysis comparable to those found between different reanalysis datasets for fields such as temperature and vorticity (Merrifield et al., 2012). (?). McLandress et al. (2013) examined improvements in transient features of the model. In addition, some minor adjustments were made to the global average temperature in the ERA-Interim fields at 5 hPa and above to remove discontinuities associated with changes in the observing system as described by McLandress et al. (2014). Sea surface temperatures and sea-ice were specified using the HadISST dataset (Rayner et al., 2003). The nudging helped correct some large scale biases in CMAM and a comparison of more transient circulation features like sudden stratospheric warmings against independent observations was improved by nudging above 10 hPa.

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 diagnostics such as the evolution of total stratospheric chlorine.

The standard CMAM chemical mechanism uses a lumping approach for the halocarbon tracers to reduce the number of chemical species that must be transported by the model. A limited number of halocarbons are explicitly treated by the chemical scheme and the remaining long-lived halocarbons are combined into the model species based on their 'fractional release values' (Schauffler et al., 2003). The concentration specified as a lower boundary condition is increased so that the total amount of organic chlorine (or bromine) of all halocarbons represented by the model species is conserved. For example, the model explicitly treats the chemistry of CFC-12 (CCl_2F_2), but the concentration of CFC-12 was increased to account for the additional chlorine carried by CFC-113. Because of the time-varying contribution of the individual halocarbons to the tropospheric concentration of the model species, it is not possible to re-scale the concentration of the model species to recover a the numerous assumptions that would be required to rescale the model species concentration that could be compared with observations would introduce significant uncertainties. Therefore, to directly compare the model halocarbon concentrations with observations, a parallel set of halocarbons was added to the model that explicitly represents individual halocarbon species. These parallel species undergo the appropriate chemical reactions using the photolysis rates and concentrations calculated by the full model chemical mechanism, though the reactions of the parallel species do not feed into the concentration of other model species. The CMAM30 simulation including the additional explicit halocarbon species will be referred to as CMAM30HR for the remainder of this work.

2.2.3 Influence of the surface boundary conditions

In CMAM, global average concentrations are typically applied as the lower boundary condition for long-lived species, such as N_2O. To capture the inter-hemispheric differences in tropospheric concentrations of the halocarbon tracers, the parallel species have separate Northern and Southern Hemisphere surface mixing ratios imposed as lower boundary conditions. The application of hemispherically-defined lower boundary conditions based on observations is consistent with the proposed approach for the upcoming sixth phase of the Coupled Climate Model Intercomparison (CMIP6) project (Meinshausen et al., 2017). The lower boundary conditions were derived from the annual average hemispheric mixing ratios from the National Oceanographic and Atmospheric Administration's ~~HAlocarbon and other~~ Halocarbon and other Atmospheric Trace Species (HATS) program (Elkins et al., 1993; Montzka et al., 1996). The annual average values were linearly interpolated in time to calculate an instantaneous surface-layer mixing ratio for the model. The model mixing ratio in the lowest six model layers ~~,~~ (approximately the lowest 1 km~~,~~) was relaxed towards the specified concentration with a time constant that increased from 25 days near the equator to 12 hours at 25 degrees of latitude. In CMAM30HR, all important losses for the species of interest have been considered. The photolysis rates and reaction rates have been updated to the values from JPL-2010 (Sander et al., 2011). The chemical losses of CFC-11 are dominated by reactions with $O(^1D)$ and photolysis in the mid-to-lower stratosphere, particularly in the tropical region. The chemical losses of CFC-12 and N_2O are similar to that of CFC-11, except they generally occur ~~higher~~ at a higher altitude in the stratosphere. CFC-11, CFC-12, and N_2O losses are insignificant in the troposphere.

To ensure the ~~efficacy~~ consistency of the boundary conditions applied to the CMAM30HR simulation, the model output was compared to measurements at surface monitoring sites using data from the NOAA HATS program. The monthly mean measurements of N_2O, CFC-11, and CFC-12 have been compared to the monthly mean in the CMAM30HR output between 2004 and 2010. Because of the way the surface boundary condition was imposed, each site was compared to the lowest model level of the closest grid point in the CMAM30HR output. The relative differences have been calculated by subtracting the measurement from the simulation and dividing by the ~~mean of the two~~ measurement for each month in the time series. The differences over time were compared and no trend in the differences was observed. The comparisons of each trace gas at each site are summarized in Fig. 1, ordered by latitudinal location. Data included here are averaged over the 2004 to 2010 period and error bars (±) indicate one standard deviation.

Both CFC-11 and CFC-12 simulations have reasonably small differences, generally less than 1%, from the surface measurements. Over the time period compared, CMAM30HR appears to ~~overpredict~~ underpredict CFC-11 at all HATS sites while the CFC-12 comparisons are not significantly different from zero within one standard deviation for all but two sites in the Southern Hemisphere. The N_2O comparisons show that the model consistently ~~underpredicts~~ overpredicts the concentrations at 11 of the 13 sites. There appears to be some latitudinal dependence in the comparisons of N_2O which may be caused by the application of a globally-averaged boundary condition in the run. Generally the differences for all the species shown are within ±1%, which is to be expected since the model boundary conditions were derived from these measurements.

2.2.4 Influence of nudging on the age of stratospheric air

The mean age of stratospheric air, that is the average time elapsed since the last time an air parcel was in the troposphere, can provide a diagnostic for determining differences in isentropic transport and mixing between different model runs (Hall and Plumb, 1994; Waugh and Hall, 2002). The stratospheric age of air in the model is derived from an idealized SF_6 tracer whose lower boundary condition linearly increases over time. In this section, the CMAM30, rather than CMAM30HR, mean age of air is used since no transport changes were made between the two runs of the model. Averaged between 2004 and 2010, Fig. 2a is the zonal distribution of the CMAM30 mean age subtracted from the mean age from an identical, but freely-running, version of the CMAM using the same specified sea surface temperatures and sea-ice data. The differences in age range from approximately -1.25 year to $+1$ years, where the positive differences indicate areas where the air in CMAM is older than that in CMAM30, and the negative differences indicate areas in which the CMAM30 age is older than in the CMAM age. In general, the nudging of CMAM appears to affect the Southern Hemisphere more than the Northern Hemisphere. Below 50 hPa in the tropics and midlatitude regions, the difference between the age in the two versions of the model is close to zero. Above 50 hPa in the tropics and midlatitudes and above 150 hPa in the polar regions, air in CMAM is older than that in CMAM30, with peaks occurring around the surf zones in the stratosphere. This implies that for the majority of the lower stratosphere the process of nudging leads to an apparent decrease in the CMAM age of air. The cause of this has not been fully explained in the literature at this time; however, it can be speculated that nudging the model to the reanalysis could be a source of artificial drag that drives the BDC to be more rapid than in the free-running version of the model.

In the polar regions of Fig. 2a, the comparisons exhibit a different behaviour in the lowermost stratosphere relative to the rest of the stratosphere. The differences are close to zero at the lowest pressure level shown (400 hPa). Above approximately 300 hPa, the CMAM30 air is older than the CMAM air and younger above approximately 150 hPa. These differences, while strongest at the latitudes poleward of 60°, can extend to approximately 40° latitude in both hemispheres. Figures 2b and 2c show the monthly evolution at 80°S and 80°N, respectively. Below 150 hPa, the CMAM30 air appears to be older than the air in the free run and this tends to be pronounced during the respective summer months in each hemisphere. At 80°S, the pattern of differences in the age of air change in altitude over time. At approximately 150 hPa in May, the air in CMAM30 is older than the air in CMAM. This difference appears at approximately 100 hPa by October, with a larger magnitude. In November and December, during Austral spring, the age difference remains such that the CMAM30 air is older than the air in CMAM. In Fig. 2c, the evolution of the differences in age of air at 80°N appears restricted to the same altitudes but the seasonal timing of the pattern is similar. The difference peaks in spring/summer and dissipates through the fall and winter. The prevalence of the older air in the polar lowermost stratosphere in the nudged run is significant because, throughout the stratosphere, air in CMAM30 is younger than that in CMAM. It is known that the freely running CMAM has a cold bias inside the Antarctic vortex. McLandress et al. (2012) suggest that there may be missing gravity wave drag (GWD) in the Southern Hemisphere based on comparisons of the free running model simulations and reanalysis data. By effectively adding this missing GWD through the nudging to reanalysis data, downwelling between 70°S to 90°S is increased, leading to higher temperatures – a reduction in the cold bias – during September and October. The increased downwelling pushes the older air deeper into the lowermost stratosphere, causing the observed differences in age between the two versions of CMAM.

In general, synoptic-scale waves are filtered out close to the tropopause and only planetary-scale waves can propagate further up into the stratosphere (Dickinson, 1969). Plumb (2002) showed that synoptic-scale wave drag drives the lower branch of the BDC, while the drag that drives the deeper parts of the BDC are associated with planetary wave drag. CMAM30 appears to reproduce the upper troposphere simulated in CMAM and, to some extent, the lower branch of the BDC as well. This is evidenced by the near-zero differences in Fig. 2a between 50°S and 50°N and between 100 hPa and 50 hPa; however, the absolute ages in this region tend to be quite small. Understanding the impact of the nudging on the age of air provides the basis for an interpretation of the isentropic transport and mixing differences between the two model runs. While it is difficult to quantify the extent to which the differences in age of air would change tracer concentrations at a given location, it is necessary to consider these results when considering implications for the free-running model, particularly for the deep branch of the BDC. Based on Fig. 2, CMAM30 clearly has older air in the extra-tropical lowermost stratosphere. It is potentially caused by either stronger downwelling of the older air from above, consistent with a stronger BDC, or ~~increased~~ reduced isentropic mixing of tropospheric air from lower latitudes (e.g. Hegglin and Shepherd, 2007). Therefore, the differences in age appear to suggest a slower shallow branch or a faster deep branch of the BDC.

2.3 Tropical Leaky Pipe Model

A modified TLP model (Ray et al., 2016) is used to interpret the differences between the CMAM30HR simulations and the ACE-FTS measurements. The modified TLP is based on a set of three coupled one dimensional equations relating transport between the tropics and each hemispheric extratropical region (Plumb, 1996; Neu and Plumb, 1999; Hall and Waugh, 2000). The model includes advection, vertical diffusion, and horizontal mixing between the extratropics and the tropics. Significant changes to the modified version of the TLP model include common pressure coordinates in all regions and the addition of particle trajectories with photochemistry. The modification was done to allow for direct comparisons between TLP output and other models and/or measurements. The Lagrangian approach is described ~~in detail~~ by Ray et al. (2016). The tropical boundaries in the TLP model averages were chosen based on observational estimates of the upwelling region (Ray et al., 2016). The model was run with a vertical resolution of 200 m and a maximum altitude of 40 km above tropopause; however, the results included here are limited to 30 km in altitude above the tropopause. To ensure the effectiveness of the TLP as an interpretation tool, Ray et al. (2016) established that the TLP could accurately simulate the CMAM30HR output with its mean circulation and a TLP-derived mixing parameter as an input. The mixing parameter was derived from a suite of simulations conducted with the TLP at varying amounts of mixing. The resultant best match to the averaged 2004-2010 CMAM30HR CFC-11, CFC-12, and age of air profiles was the ~~level of mixing selected.~~ mixing efficiency selected to initiate the simulations. The TLP model is used here to identify the changes to the CMAM30HR tropical upwelling and effective mixing that may improve the comparisons between ACE-FTS and CMAM30HR by testing a range of tropical upwelling and mixing efficiency settings.

Diagnosing the causes of discrepancies between measurements and model output Diagnosing the biases in the model stratospheric circulation requires a complete separation of the effects of the strength of the BDC and the mixing. A simplified model, such as the TLP, is useful to interpret differences between measurements and CCMs because of the complexity of wave activity contributing to stratospheric mean circulation and mixing. It would not be prudent to adjust model parametrizations in CMAM to modify wave breaking because many aspects of the model climatology would be impacted with no way of separating the effects (Ray et al., 2016). Therefore, a suite of model runs were computed with the TLP to test a range of mean circulation strengths and mixing efficiencies. The TLP runs began with the CMAM30HR best fit to the TLP model and then the best combination was selected to match the ACE-FTS measurement profiles of CFC-11 and CFC-12, as well as an age-of-air estimate derived from balloon-measurements. In section 6, the TLP simulations are used to investigate the behaviour of the tropical pipe in CMAM30. While the TLP-based analysis does not identify a specific mechanism, it can separate the contributions of model biases in the BDC and in mixing to biases in the resulting distribution of species.

3 Comparison methods and sampling considerations

3.1 Measurement-model comparison techniques

Two of the comparison techniques used in this study are described here; the first is the comparison of zonal means, and the second is the computation of joint probability density functions.

3.1.1 Zonal mean comparison technique

To assess the transport and chemistry in CMAM30HR, measurements of N_2O, CFC-12, and CFC-11 are compared with simulated concentrations in latitude-pressure coordinates. A common method of visualizing the distribution of long-lived trace gases is the zonal mean cross section. In this work, data from the ACE-FTS profiles, sampled CMAM30HR profiles, and the relative difference profiles were averaged in 5° latitude bins and over 18 pressure levels (equally distributed in the log of pressure from 450 hPa to 10 hPa), corresponding to altitude ranges from approximately 5 km to 30 km. In these plots, color contours indicate the VMR of the species. The comparison of the ACE-FTS measurements and the subsampled CMAM30HR output is shown as the average of the differences, defined as CMAM30HR minus ACE-FTS divided by the mean of the model output and measurementsACE-FTS measurement. The altitude of the average thermal tropopauses is typically indicated by a blue black line. All measurements and subsampled model output between June 2004 and May 2010 have been included, representing an average of six years of observations and simulations.

3.1.2 Joint probability density functions

Tracer-tracer correlations have been used in a number of studies to identify transport and mixing characteristics in the stratosphere and to derive climatologies from sparse data (e.g., Plumb and Ko, 1992; Plumb, 1996; Toon et al., 1999; Sankey and Shepherd, 2003; Hegglin and Shepherd, 2007). Plumb and Ko (1992) showed that long-lived species exhibit compact correlations even with varying meteorological conditions, minimizing discrepancies resulting from sampling and daily variations; thus, sparse measurements, such as those from aircraft, can be useful in model assessment studies (e.g., Sankey and Shepherd, 2003). The correlations used here are N_2O/CFC-11 and include only the stratospheric data available (data located at altitudes above 2 km above the thermal tropopause). The correlations produced from both the ACE-FTS measurements and CMAM30HR simulations exhibit compact relationships that tend to be densely populated. Determining whether the model can capture the clustering in addition to the overall shape is important to understanding whether the stratosphere is well-reproduced by the model. Understanding the density distribution of a data set is particularly useful for tracer-tracer relationships with compact correlations. Following the methods of Sparling (2000) and Hegglin and Shepherd (2007), normalized joint probability distribution functions (JPDFs) have been calculated for the ACE-FTS and CMAM30HR correlations described above. JPDFs are two-dimensional histograms that reveal the clustering of data (Hegglin and Shepherd, 2007) and can be used to test how well a model captures the behaviour of trace gases in the stratosphere. Hegglin and Shepherd (2007) have shown the impact of ACE-FTS measurement uncertainties in joint PDFs by comparing the full model output, subsampled model output, and ACE-FTS measurements. They found that there was larger variability in the ACE-FTS joint PDFs compared to those of the subsampled CMAM output.

3.2 The influence of beta angle

ACE-FTS records a series of spectra along a slanted path line-of-sight during each occultation. The length of this slanted path is different for each occultation. Each ACE-FTS spectrum is assigned a latitude and longitude at the 30-km tangent point, geometrically calculated (Boone et al., 2005, 2013). A sample year of the geometric 30-km tangent point latitudes is provided in Fig. 3 (black circles), showing the annual repeating latitudinal coverage. The beta angle parameter, a measure of the angle between the solar vector and the satellite orbit plane, has also been included in Fig. 3 (red circles). The beta angle is an important parameter to consider because as it changes, so does the geographic distance between each spectrum acquired through the profile of any given occultation. The distance is greatest at high beta angles (both positive and negative), which occur when ACE is in view of the Sun for longer periods. Since the FTS instrument measurement frequency is held at a constant two second interval, more measurements per profile and longer ground-paths of the retrieved profile occur at high beta angles.

Considering the impact of observation sampling is a critical step when comparing measurements with model output. The work of Toohey and von Clarmann (2013), Toohey et al. (2013), and Millán et al. (2016) illustrate the necessity for considering the sampling patterns resulting from different measurement techniques and satellite orbits. The ground path length of a profile is considered because a single profile can be representative of more than one geographic region, typically varying more over latitude than longitude. A refraction model is used to determine the geographic locations along the slant path of the ACE-FTS profiles (Boone et al., 2005, 2013). At the 30-km tangent altitude, it has been found that for 98% of the ACE-FTS occultations, the difference between the geometric latitude and the refraction calculation is less than 0.2°. A useful marker of a nominal occultation length is at a beta angle of 53°, corresponding to an occultation duration of three minutes. Occultations longer than three minutes, at beta angles larger than 53°, measure across large spatial distances and represent approximately 12% of the ACE-FTS data used in this work. Both horizontal and vertical variations within the CMAM30HR output will impact the comparison to ACE-FTS measurements. The CMAM30HR fields are output on a grid with a spatial resolution of approximately 400 km. While most ACE-FTS occultations have a shorter horizontal extent than the CMAM30HR grid point footprints, there are some occultations that fall outside of a single grid point range in the upper troposphere and lower stratosphere. For example, between 5 km and 30 km, 15% of occultations extend across more than one CMAM30HR grid point footprint, of which 82% are at beta angles greater than 53° and 80% are at latitudes poleward of 30°. Various model sampling techniques have been investigated since occultations span multiple grid point footprints in both latitude and longitude, as well as vertically.

3.3 Comparison of sampling techniques

To determine the impact of sampling the model output at varying levels of detail, three methods were tested ~~and compared to~~ by sampling the full model output (the ~~CHAM30HR~~ CMAM30HR output at all latitudes and longitudes for each 5° latitude bin) between June 2004 and May 2010. All three methods began with identifying the temporally coincident three-dimensional CMAM30HR output (latitude, longitude, pressure) for each ACE-FTS profile; the output within 3 hours of the occultation was selected with no temporal interpolation. The three-dimensional output was interpolated in the vertical dimension to the ACE-FTS profile pressure grid, which is different for each occultation since the retrievals are ~~performed~~ provided on an altitude scale. The 'basic' sampling method involved selecting a vertical column based on the nearest neighbour grid point to the 30 km tangent point location with no interpolation and no consideration of the vertical extent of the profile. The 'intermediate' level of sampling extracted the vertical column based on a bilinear interpolation of the four closest grid points to the 30 km geometric tangent point but with no consideration of the variation in geographical location of the tangent points ~~; therefore, there was horizontal interpolation but no vertical interpolation~~above or below 30 km. The 'advanced' sampling method improves on the intermediate level of sampling by performing the bilinear interpolation at each level of the ACE-FTS profile using the distinct geographic locations, derived from the refraction model (Boone et al., 2005, 2013), for the respective level. Therefore, at each vertical level in the ACE-FTS profile, a spatial bilinear interpolation including the four geographically closest grid points was computed to determine the comparable CMAM30HR VMR. To illustrate the sampling effect between 450 hPa and 10 hPa (~~5-30~~ 5–30 km), the ~~relative differences~~differences, relative to the advanced technique, in the zonal mean of N_2O over the observation period ~~(June 2004 – May 2010)~~ (June 2004 – May 2010) are compared in Fig. 4. The advanced method is compared to the full output of the model (Fig. 4a), the basic model sampling (Fig. 4b) and the intermediate sampling (Fig. 4c).

The comparison of the advanced sampling and full model output in the stratosphere is dominated by the influence of the polar vortex in both hemispheres. Generally, there is good agreement throughout the troposphere, with less than 5~~%~~ % differences. For the long-lived tracers investigated in this work, the free troposphere is well mixed such that there is minimal influence of the ACE sampling pattern. In the stratosphere, however, there are pronounced differences on the order of 20%. Air in the polar vortex is typically ~~representative~~ composed of older air brought down from higher altitudes. Therefore, tracer concentrations within the vortex and vortex edge tend to be significantly different from those in the mid-latitude surf zone during the winter in each hemisphere. The differences seen in Fig. 4a occur because comparing the full output to ~~measurements~~ 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 average that includes the model simulation at all longitudes and time periods. At the edge of the vortex, tracer concentration gradients are strong, so comparing measurements to the full output of the model will tend to smear the influence of the vortex on tracer concentrations. The differences are not symmetric latitudinally due to different dynamical conditions in each hemisphere. For example, in the Antarctic stratosphere in September there is a strong decrease in the geographic extent of the polar vortex with height such that the vortex is much wider geographically at 100 hPa than at 10 hPa. A similar phenomenon occurs in the Arctic but it is much more variable both spatially and vertically.

The advanced sampling technique is compared to the basic sampling in Fig. 4b. The distinction between Fig. 4a and 4b is that rather than using the full model output, the nearest grid point is selected based on the geographic location of the 30 km tangent point of the ACE-FTS measurements. Even this basic level of sampling improves the comparison in the stratosphere substantially, bringing the range of differences down to $\pm 10\%$. It is worth noting that the stratospheric differences in the midlatitudes are on the same order of magnitude with a similar latitudinal pattern but of opposite sign. The differences of 5-10% are primarily negative in the Southern Hemisphere stratosphere and positive in the Northern Hemisphere. This pattern occurs because each ACE-FTS profile is tilted such that the top of the profile is always further north than the bottom of the profile, leading to a directional bias. In the Northern Hemisphere, profiles tend to ‘point’ toward the North Pole, therefore measurements in this hemisphere are subject to a poleward bias. In the Southern Hemisphere, profiles point toward the equator, leading to an equatorward bias in sampling. The choice of ‘closest’ grid point likely biases the comparisons, leading to the differences in Fig. 4b.

Figure 5 illustrates the average latitudinal extent of occultations ~~,~~ between the 5 km and 30 km tangent altitudes, showing the two directionalities for each 5° latitude bin included in Fig. 4b, with error bars indicating one standard deviation from the mean latitudinal extent. A poleward bias implies that the 30 km tangent point is located poleward of the 5 km tangent point and an equatorward bias reflects when the 30 km tangent point is located equatorward of the 5 km tangent point.

In the midlatitude region of the Northern Hemisphere, the average latitudinal extent of occultations exhibits a primarily poleward bias while the occultations in the Southern Hemisphere midlatitudes exhibit a primarily equatorward bias. The Northern Hemisphere poleward bias in Fig. 5 corresponds to the positive differences in the Northern Hemisphere midlatitude stratosphere in Fig. 4b and the Southern Hemisphere equatorward bias corresponds to the negative differences in the Southern Hemisphere midlatitude stratosphere. Since the basic sampling is limited to the 30 km tangent altitude, the comparison to the advanced technique in Fig. 4b reflects the influence of the geographical extent of the ACE-FTS profiles. In the Northern Hemisphere, contributions from sampling the model at lower latitudes (which tend to have a higher concentration) lead to positive differences between the two sampling techniques; while in the Southern Hemisphere, contributions from sampling the model at higher latitudes (lower concentrations) lead to negative differences between the advanced and basic sampling techniques. This nuance in the sampling pattern highlights the importance of considering the sampling pattern of the ACE-FTS occultations when comparing measurements to model output.

Figure 4c compares the advanced and the intermediate sampling. With approximately $\pm 2\%$ differences between the two techniques, it is clear that the intermediate sampling technique can account for much of the geographic extent of the ACE-FTS profiles at this model resolution. However, if comparing to a model with a finer resolution or if a larger vertical extent is considered, accounting for the full geographic extent of the profile will become more important. The more detail that is included in the sampling of the model, the more comparable the output is to the observations. The advanced method of sampling provides the most appropriate model profiles for direct comparison between ACE-FTS and CMAM30HR. Therefore, all comparison results shown in this study utilize CMAM30HR output that has been sampled using the advanced technique.

4 Zonally-averaged tracer morphologies

~~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.~~

4.1 General features of tracer morphology comparisons

The zonally-averaged annual-mean distribution of N_2O is presented in Fig. 6. The N_2O simulated by CMAM30HR is shown in Fig. 6a, the ACE-FTS measurements are shown in Fig 6b, and the average of the profile differences within each 5° latitude bin is shown in Fig. 6c. Both the ACE-FTS measurements and the CMAM30HR distribution of N_2O in Fig. 6 show many of the features that are expected of a long-lived tracer with a tropospheric source and chemical losses that occur primarily in the stratosphere. 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. There is a hemispheric asymmetry in the decrease with altitude beyond the tropical region. The southern extratropical and Antarctic concentrations of N_2O tend to decrease with altitude more rapidly ~~in the Southern Hemisphere than~~ than those in the Northern Hemisphere. This ~~is likely caused by significant~~ asymmetry is likely driven by differences in the ~~conditions of the influence of downwelling within the polar vortex between the two hemispheres~~isolation of the polar vortex in each hemisphere and the large-scale downwelling that is largely dependent on this isolation. By visual comparison, the lowest concentrations observed and simulated appear to be in the Antarctic region between 30 hPa and 10 hPa and the Arctic stratosphere above 20 hPa. The quantitative comparison between the ACE-FTS and CMAM30HR zonal mean N_2O distributions in the bottom panel of Fig. 6 reveals significant differences throughout the lower stratosphere, with the largest differences in the northern polar region. CMAM30HR simulates larger concentrations of N_2O in the lower stratosphere. Upwelling in the tropics, descent in the extratropics, and mixing in the surf zone define the transport controls on the distributions in the stratosphere. The differences observed in Fig. 6c are influenced by the combined effects of these features on the measured and the simulated concentrations of N_2O. Therefore, if there were no issues in the simulated stratospheric transport, the differences would be of similar magnitude to the upper troposphere comparisons (less than ±5~~%~~%), unless there was a significant flaw in the chemical losses in the model.

Investigating measurement-model comparisons using more than one trace gas leverages the varying lifetimes of and chemical processes of each gas. The comparisons of CMAM30HR and ACE-FTS, equivalent to the bottom panel of Fig. 6, are shown in Fig. 7a-c for N_2O, CFC-12, and CFC-11, respectively. Each of the panels shows the differences as a percentage, but note that the scale for Fig. 7c is different. This is because the CFC-11 relative differences become large in the stratosphere where concentrations tend toward zero. All three species show good measurement-model agreement (within approximately 5%) in the well mixed troposphere. In the tropics, the VMRs of these three species remain relatively constant up into the lower stratosphere where chemical loss processes begin to break down the compounds. Above 70 hPa in the tropics, the CFC-12 and N_2O comparisons show similar agreement (on the order of 5%). However, above 50 hPa in the tropics, CFC-11 exhibits both positive and negative differences between the measurements and model simulations. These differences in CFC-11 are also observed outside the tropics above 70 hPa and are much higher (on the order of 50%). In the Northern Hemisphere extratropics, the differences are primarily positive but become more variable closer to the northern polar region. Very small concentrations above 70 hPa, which occur because of the significant photolytic losses in the tropical lower stratosphere, lead to the large magnitude of the differences in CFC-11. The irregular pattern in the CFC-11 differences is driven by the variability in the measurements as ACE-FTS reaches its detection limit.

4.2 Seasonality of the tracer morphology comparisons

The structure and intensity of the BDC varies seasonally. In general, the BDC is strongest in the Northern Hemisphere winter ~~due to wave driven~~ because of wave-driven enhancements initiated by topography, and because, during the this time of year, ,climatological westerlies facilitate wave propagation into the stratosphere (e.g., Rosenlof, 1995; Plumb and Eluszkiewicz, 1999). It is well known that tropical upwelling is stronger in the summer hemisphere; therefore during the December-January-February (DJF) season, the upwelling is strongest in the Southern Hemisphere (e.g., Yulaeva et al., 1994). Investigating the comparisons between the CMAM30HR simulations and the ACE-FTS observations in a seasonal context helps to determine whether the differences observed earlier are related to the behaviour of the BDC. If the differences observed in Fig. 7 are driven by the simulation of the BDC in the model, it would be expected that the morphology of the seasonal differences would appear to follow the behaviour of the BDC.

For each season, Fig. 8 identifies the differences between the simulation and the measurements. The seasonal composites shown here do not fully represent the seasons because of the sampling pattern of the ACE-FTS (recall Fig. 3). However, the comparisons are relevant since the CMAM30HR output has been subsampled, as previously described. The most obvious features across all seasons in Fig. 8 are the same as those of the differences shown in Fig. 7a. There is good agreement in the lower stratosphere at all latitudes and in the tropics up to about 50 hPa. In the mid-stratosphere, CMAM30HR simulates higher concentrations of N_2O than those measured by ACE-FTS. Some of the largest differences occur at the high northern latitudes during boreal winter and spring, presumably in the region of downwelling within the polar vortex.

The large disagreements in the north polar region during winter and spring indicate that the ~~upwelling~~ downwelling portion of the BDC across the different seasons is not well characterized by CMAM30HR. ~~However~~Meanwhile, the shifting of the agreement in the tropical region through the seasons indicates ~~a robust simulation~~ that the simulation is consistent with the spatial distribution of the observations in this region. For example, the difference in the southern tropical latitudes appears small (close to 0~~%~~%) up to 50 hPa and to approximately 40°S in DJF, but the agreement diminishes in this region in the March-April-May (MAM) season, presumably when the tropical upwelling begins to decline in strength and shift toward the equator. A similar pattern is observed in the Northern Hemisphere during Austral winter where the differences in the northern tropical latitudes appear to be small up to 50 hPa and to approximately 40° N in the ~~June-July-August~~ June-July-August (JJA) and ~~September-October-November~~ September-October-November (SON) seasons. These results support the understanding that the most rapid tropical upwelling is occurs in the summer hemisphere as first reported by Yulaeva et al. (1994).

~~The polar regions of each hemisphere in the comparisons of~~ In the comparisons shown in Fig. 8, the Northern high latitudes measurement-model differences are significantly different compared to those in the Southern high latitudes. There are negative differences at high southern latitudes in MAM, JJA, and part of SON. The differences seem quite asymmetric when compared with the results for the Northern Hemisphere. The negative differences at high southern latitudes appear to descend between MAM and JJA and begin to weaken in SON and the vortex break-down. There is also some asymmetry in the differences between 30 and 10 hPa between the Northern and Southern Hemispheres, ~~particular~~ particularly in winter for each hemisphere. These differences are likely due to the behaviour of the polar vortex in each hemisphere. In particular, the differences may be related to the model's (either CMAM30HR or the ERA-Interim reanalysis 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.

Since the model run compared here has been nudged to the ERA-Interim meteorology, it cannot be simply concluded that the differences are due to the variable nature of the vortex. The vertical migration of the negative differences in the southern polar region across MAM, JJA, and SON suggests the vortex variability physical or chemical mechanism as the cause. The negative differences mean that there is an underprediction in the CMAM30HR simulation, which could happen if air that is too old is brought down into the vortex. As the vortex forms in fall, the negative differences appear and descend through the winter, reaching a maximum latitudinal extent. The appearance of the negative differences in the comparison of N_2O between the observations and the model is conspicuous because elsewhere N_2O is higher in CMAM30HR throughout the lower and middle stratosphere.

In Fig. 8b and Fig. 8d, the large positive differences in the Northern Hemisphere stratosphere may be caused by too much topographic wave driving in CMAM30HR (and CMAM30). This would lead to the movement of air from the tropical region into the extratropics and polar regions, too quickly and thereby simulating higher than expected concentrations of N_2O in the Northern Hemisphere stratosphere.

5 Comparison of mixing barriers

It is well understood that quasi-horizontal mixing flattens tracer isopleths in mixing regions and sharpens gradients at mixing barriers (e.g., Plumb, 2002). However, it can be very difficult to separate the effects of mixing barriers from the residual circulation when looking at zonal mean comparisons between measurements and models. Therefore, it is necessary to scrutinize mixing barriers individually.

5.1 The polar vortex

Consideration of the behaviour of the polar vortex in both hemispheres is necessary as they have atmospheric processes that affect their behaviour differently over time. For this purpose, the monthly mean differences between ACE-FTS and CMAM30HR over the time period of the study have been determined for the stratospheric abundances of N_2O and CFC-12. CFC-11 has been excluded from this comparison because the limited vertical extent of the sensitivity of the measurement results in too few data in the stratosphere. All comparisons shown here are profiles located poleward of 60°, and show the mean of the difference between the ACE-FTS and CMAM30HR profiles, ~~expressed as a percent of the average of the two profiles~~relative to the ACE-FTS profile. The comparisons here extend the work of Pendlebury et al. (2015), who discussed the polar region simulations in CMAM30 extensively by comparing temperature, ozone, methane, and water vapour up to 0.001 hPa with a variety of satellite instruments including ACE-FTS. All the species investigated in Pendlebury et al. (2015) have much shorter lifetimes than those of N_2O and CFC-12. The advantage of using species with long lifetimes is that at least some of the parcels of air that are sampled have been through the deep branch of the BDC. By restricting comparisons to the polar stratosphere, it is primarily air from the deep branch of the BDC that is being investigated. Tracers ~~,~~ with a stratospheric sink ~~,~~ are most depleted from the deep branch because they have had the most time for chemical loss to occur since they entered the stratosphere.

The comparison of the N_2O and CFC-12 difference time series (Fig. 9) demonstrates that there is inter-annual variability that is consistent between the two gases in the Arctic. While the two species shown follow similar patterns over time, there appear to be larger differences in the CFC-12 comparison than in the N_2O comparison. There are two possible (and related) reasons for this difference: the range in the concentrations of CFC-12 is much larger than that of N_2O, and there are differences in their respective chemical losses. For example, the photolysis loss of CFC-12 is faster than that of N_2O throughout much of the stratosphere, as evidenced by the differences in their lifetimes (102 years for CFC-12 and 123 years for N_2O). Generally, it appears that the model simulates higher concentrations of both species compared to ACE-FTS measurements through much of the stratosphere, with the largest differences occurring above 30 hPa. When the concentration of either tracer becomes very small (typically air that has descended from the upper stratosphere or mesosphere), the relative differences between ACE-FTS and CMAM30HR can be enhanced. These differences are most clear in the N_2O comparisons during the autumns of 2004 and 2009, and the springs of 2007 and 2010; while in the CFC-12 comparisons, the springs of 2005, 2006, and 2008 exhibit additional occurrences of this feature.

During each of these periods, ACE-FTS observed much lower concentrations compared to CMAM30HR. Both the speed and structure of the residual circulation within the CMAM30HR run can contribute to the observed differences. It is possible that the BDC in the CMAM30HR simulation is drawing air through the deep branch of the BDC too rapidly. The vertical structure of the differences observed in Fig. 9, particularly between 70 hPa and 30 hPa, may be caused by ACE-FTS measuring a descent in the air mass that CMAM30HR does not simulate. It is more likely that the model circulation is not moving enough air through the loss region of these tracers and through to the polar vortex. For photolytic tracers, the structure of the circulation is more important than the speed of the residual circulation because photolysis rates are so fast. Above a certain level in the stratosphere, the tracer is completely destroyed when air passes through the region. The distribution of photolytic species is a mixture of air that ~~was~~ passed through the region of rapid loss and the air that by-passed the loss processes. This result is consistent with the work of Pendlebury et al. (2015), where they found large differences in temperature and ozone between satellite observations and CMAM30.

There is less interannual variability in the Southern Hemisphere comparisons (Fig. 10) than that in the Northern Hemisphere (Fig. 9). This is expected since the variability of the southern polar stratospheric dynamics is much less than that in the northern polar stratosphere. However, the magnitude of the differences between the measurements and simulations is larger in the Antarctic stratosphere than in the Arctic stratosphere. Moreover, while the patterns of the differences in N_2O and CFC-12 are quite similar in Fig. 10a and 10b, the magnitude is more pronounced in the CFC-12 comparisons. The largest differences occur above 30 hPa where the concentrations of CFC-12 are extremely low. The peak in the magnitude of the CFC-12 differences appears to increase in vertical extent through the Austral springtime. The largest differences tend to occur during summer (December) at around 40 hPa for both tracers. The differences in CFC-12 are established at the top of the vortex in July and propagate down until vortex break up in December. However, this propagation doesn't occur to the same extent in the N_2O comparisons, which may be a reflection of the differences in the chemistry of the two tracers. For example, a source of N_2O in the lower thermosphere has recently been identified in ACE-FTS measurements by Sheese et al. (2016). The N_2O source descends into the mesosphere and stratosphere, thereby influencing air that is circulated in the BDC (Sheese et al., 2016). The transport of enhanced N_2O downwards from the upper atmosphere has also been detected by Funke et al. (2008a, b). The CMAM30HR does not include this source of N_2O. The results presented here are consistent with the methane comparisons discussed in Pendlebury et al. (2015). Generally, the descent of the model's high bias is observed in both hemispheres for all three trace gas species. The results of Pendlebury et al. (2015) indicate that the high bias is consistent with a fast BDC and that the downward propagation of the bias is a problem with the parameterizations in the model above 10 hPa.

5.2 The extratropical tropopause

The transport barrier at the extratropical tropopause can be permeable, allowing the exchange of air between the troposphere and the stratosphere. Understanding how CMAM30HR simulates this exchange assists in the interpretation of mixing effectiveness in the model and the impact of its ~~vertical~~ finite resolution on the comparisons. Since ACE-FTS predominantly samples the polar regions and has fewer samples of extratropical latitudes, interpreting latitudinal or seasonal dependence of the exchange of air across the tropopause in the full atmosphere using these measurements must be considered from a tropopause coordinate perspective (Hegglin et al., 2009). In this study, a diagnostic of the tropopause barrier has been developed for comparison between the simulations and the measurements. The tropopause height, used in this analysis to define tropospheric air and stratospheric air, is the thermally defined tropopause based on the derived meteorological products for ACE-FTS (Manney et al., 2007) and based on sampled temperature profiles from CMAM30 output for the CMAM30HR simulations.

Since CFC-11 has a strong vertical gradient in concentration in the stratosphere, it can be used as a proxy for determining the exchange of air across the tropopause. The diagnostic developed for this analysis is the frequency of intrusions, which signifies the frequency of stratospheric (tropospheric) air penetrating into the troposphere (stratosphere). A data point is defined as an intrusion based on two criteria: the physical location of data point and the concentration relative to a tropospheric concentration threshold and only data below the ~~420 K~~ 420 K potential temperature layer of the atmosphere have been considered. The tropospheric threshold was defined separately for each hemisphere as the tropospheric mean minus 1.5 times the tropospheric standard deviation. An intrusion frequency metric has been developed for comparison between the simulated and observed concentrations. A tropospheric intrusion is identified when a measurement is physically located in the stratosphere but its concentration is larger than the tropospheric threshold. A stratospheric intrusion is identified when a measurement is physically located in the troposphere but its concentration is smaller than the tropospheric threshold. For each five degree latitude bin between 20° and 60° latitude (N and S), the frequency of tropospheric and stratospheric intrusions have been determined. This technique was used to calculate frequencies for both the ACE-FTS measurements and CMAM30HR profiles.

The comparison of ACE-FTS and CMAM30HR tropospheric and stratospheric intrusions within the southern and northern extratropical latitudes is shown in Fig. 11a and 11b, respectively. There appears to be better agreement for the stratospheric intrusion comparisons and at some latitudes there is very good agreement between ACE-FTS (red) and CMAM30HR (black). However, there are some latitudes, such as ~~30-40° S and 20-25° N~~30−40° S and 20−25° N, where the stratospheric intrusions identified in the simulation are a factor of two fewer than those from the ACE-FTS measurements. In general, isentropes that are in the extratropical lowermost stratosphere are in the troposphere in the tropics (e.g., Holton et al., 1995). Fast isentropic transport occurs because wave motions cause air to rapidly change latitude, leading to the transport of stratospheric air into the troposphere. Since the disagreement in measured stratospheric ~~intrusions~~ exchanges is largest at latitudes where this isentropic transport tends to occur implies that CMAM30HR is not capturing this mechanism well or simply lacks the resolution to fully resolve the stratospheric intrusions features.

Based on Fig. 11, tropospheric intrusions occur more frequently in the ACE-FTS data than stratospheric intrusions and vary more significantly in number across the latitudes in both hemispheres. Additionally, the differences between the measurements and simulations are larger than for the stratospheric intrusions. It is possible that there is more tropospheric air found in the stratosphere with this method than stratospheric air found in the troposphere because tropospheric CFC-11 is more easily distinguishable. The manner in which the intrusions have been defined here does not rule out stratospheric air being identified as tropospheric if there is rapid, poleward transport out of the lower tropical stratosphere. Schwartz et al. (2015) suggest that it is easier to distinguish tropospheric intrusions into the stratosphere using tropospheric tracers, and stratospheric intrusions into the troposphere using stratospheric tracers. Tropospheric CFC-11 concentrations found in the stratosphere at extratropical latitudes indicate air that likely has not cycled through the BDC because the air parcels have not experienced any loss processes that can only happen in the stratosphere; while stratospheric air in the troposphere could be representative of a range of aged air. For example, this air could have cycled through the BDC very quickly by re-entering the troposphere in the subtropics or it could have gone through the deep branch, allowing it to be more depleted and therefore more identifiable as stratospheric-like air.

A mechanism for extratropical tropospheric air to be uplifted into the stratosphere has been identified recently. Building on the work of Pan et al. (2009), Peevey et al. (2014) showed that the occurrence of double tropopauses is associated with the strength of the tropopause inversion layer (TIL), as well as Rossby wave breaking. They also showed that as the strength of the TIL increases, cyclonic circulation in the upper troposphere switches to anticyclonic circulation, thereby driving an increase in the upward motion. Based on the double tropopause calculations of Schwartz et al. (2015), it is likely that ACE-FTS observes this phenomenon of upward motion, leading to a higher frequency of tropospheric intrusions across the extratropical regions. CMAM30HR appears to inadequately simulate this mechanism since it doesn't have a sharply defined tropopause compared to reality (Birner and Bönisch, 2011). It is possible that both the spatial and vertical resolutions of the simulations performed limit the model's ability to capture this synoptic-scale activity.

Seasonal averages of the tropospheric and stratospheric intrusions are shown in Fig. 12, where Fig. 12a and Fig. 12b are the tropospheric intrusion frequencies for the southern and northern extratropics, respectively and Fig. 12c and Fig. 12d are the stratospheric intrusion frequencies for each hemisphere. For this case, the data shown in Fig. 11 were averaged seasonally in each hemisphere and the error bars represent the latitudinal variability defined as one standard deviation of the mean. Recall that since the intrusion frequencies in Fig. 12 are affected by the sampling pattern of ACE-FTS, the seasonality shown may not be representative of the actual seasonality of the atmosphere. The stratospheric intrusion comparisons are remarkably good; there does not appear to be a significant difference between the simulations and the measurements. Therefore, ~~it is unlikely that there is a physical mechanism or deficiencies in the model leading to the differences observed in Fig. 11 and the differences~~ the differences are primarily due to the ~~model resolution~~ finite horizontal and vertical resolution of the model which leads to differences in the representation of stratosphere-troposphere intrusion events. The comparisons of tropospheric intrusions exhibit similar behaviour between the two hemispheres. The measurement-model differences appear to be largest during SON and smallest during MAM. The consistency of the increase during SON in tropospheric intrusion frequency between the two hemispheres may be the result of convective over-spill into the stratosphere across the subtropical tropopause barrier if the majority of the tropospheric intrusions are driven by convection in the tropics. This convective influence could extend across the extratropical latitudes via isentropic transport from the tropics. The strength of the convection during JJA and SON and its influence on mixing across the tropopause barrier is observed by the largest variability in the stratospheric intrusions. ~~For example, if the model is does not have a strong enough Asian monsoon circulation, then there would be fewer tropospheric intrusions simulated, particularly in JJA.~~

5.3 The tropical pipe

In this section, JPDFs are used to investigate the tropical pipe barrier. Figure 13 illustrates the ~~N_2O/CFC~~ N_2O/CFC JPDFs of the entire Northern Hemisphere stratosphere, beginning at 2 km above the tropopause to 30 km (up to 10 hPa), for ACE-FTS (Fig. 13a,c) and CMAM30HR (Fig. 13b,d) for CFC-12 (Fig. 13a,b) and CFC-11 (Fig. 13c,d). Hegglin and Shepherd (2007) have highlighted the use of this technique for comparing CMAM and ACE-FTS measurements. The ~~N_2O/CFC-12~~ N_2O/CFC-12 JPDFs (Fig. 13a and 13b) exhibit a quasi-linear relationship in both the measurements and the simulations. The loss rates of CFC-12 and N_2O are very similar in the upper troposphere and lower stratosphere, leading to a linear relationship in the JPDFs. The model JPDFs tend to peak at the higher concentrations of N_2O and CFC-12 but are more evenly distributed throughout the range of concentrations. ~~This difference~~ These differences between the measurements and simulations ~~is~~ are likely due to the overly rapid BDC in the model simulation, ~~as previously discussed .~~ leading to higher concentrations of the trace gases in the simulation, which is consistent with the zonal mean comparisons discussed in Section 4. The ACE-FTS data in Fig. 13a are much more scattered than the CMAM30HR in Fig. 13b, where the simulated tracers are highly correlated and the JPDF is highly compact. The differences in the spread of the correlations are related to differences in mixing and chemistry, but are also influenced by the precision of the ACE-FTS measurements and the constraint of the boundary conditions in the simulations. The CMAM30HR JPDF is very compact because of the similarity of the chemical losses of the two tracers in the model. Additionally, the surface boundary conditions applied do not represent the variability observed in the atmosphere. It is the atmospheric variability that contributes to the variability observed in the ACE-FTS JPDF around 150-200 ppbv 150 – 200 ppbv of N_2O (Hegglin and Shepherd, 2007).

The N_2O/CFC-11 N_2O/CFC-11 JPDFs show two segments of linear correlations in both the measurements and model results (shown in Fig. 13c and 13d, respectively) that would have otherwise been overlooked in data dense tracer-tracer correlation plots (also see Hegglin and Shepherd (2007)). The presence of a bimodal correlation between N_2O and CFC-11 has been previously observed by Plumb (1996). The separation is caused by differences in local chemistry in the tropical pipe region. The tropics are somewhat isolated from the midlatitudes so that the steeper slope is a signature of the local chemistry, or the relative loss rates of CFC-11 and N_2O in the tropical lower stratosphere. This relationship is observed because the photochemical lifetime of CFC-11 is shorter than the time scale for mixing by horizontal eddy transport timescale (Plumb, 2002) . According to Plumb and Ko (1992), slope equilibrium conditions that define the linear relationship seen in the N_2O/CFC-12 N_2O/CFC-12 JPDF are only satisfied if the photochemical lifetime of a species is much greater than the horizontal eddy transport time scale. When this condition is satisfied, the slopes of the isopleths are only a function of atmospheric circulation.

Isolating the measurements and simulations in the tropical region allows for the characteristics of the tropical pipe to be investigated. A JPDF comparison is provided in Fig. 14 for both the ACE-FTS measurements and the CMAM30HR simulations in the tropics during the DJF, MAM, JJA, and SON seasons. 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. Bimodal behaviour is observed in each season and in both the measured and simulated JPDFs. In general, the maximum of the JPDF appears to be positioned towards higher concentrations in the simulation compared to the measurements, where the maximum in the probability tends to extend throughout the shallower segment. As was observed in the N_2O/CFC-12 N_2O/CFC-12 JPDFs in Fig. 13, the probabilities tend to be weighted towards the higher concentrations implying that there is younger air (of tropospheric origin) in the simulated stratosphere than the measured stratosphere. Of note is that the length and the width of the shallower segment are consistently larger in the measurements. The longer length – extending to low concentrations – indicates the presence of older air not found in the model simulations. The larger width also coincides with the degree of separation between the primary and secondary segment in each season, where the simulated JPDFs appear to have a much greater separation than the measurements. These features are likely dependent on the amount of quasi-horizontal mixing influencing the JPDFs, implying that there is not enough quasi-horizontal mixing occurring in the simulation. 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 an indication of insufficient mixing into the tropics, rather than a product of a too rapid tropical ascent. If it was just too fast ascent, both tracers would be affected in the same way. But because the mid-latitudes N_2O/CFC-11 N_2O/CFC-11 relationship is less steep, it is the mixing in of this air that makes the tropics have shallower slope. Then it becomes a question of whether mixing is underestimated because the ascent is too fast or if the model doesn't simulate the structure of the pathways of the BDC correctly.

~~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.~~ Fig. 14. During DJF, there is minimal separation between the two segments. The position of the maximum in the JPDF (the red region) in Fig. 14a does not extend below 200 ppbv of N_2O, while the location of the maximum of the simulated JPDF in Fig. 14b is limited to above 250 ppbv of N_2O with the maximum primarily located above the separation of the two segments. During MAM, the maximum probability of the ACE-FTS JPDF in Fig. 14c extends throughout the shallow and steep segments. This implies that there is significant mixing occurring during this time period. In the simulated MAM, Fig. 14d, the maximum of the probability is restricted to the higher concentrations of N_2O, prior to the separation of the two segments. This indicates that while mixing is occurring, it does not occur frequently enough to simulate the atmosphere well.

During JJA, the two segments begin to separate in both the measured and simulated JPDFs. The shallower segment in the simulated JPDF is the shortest during JJA of all the seasons with the maximum of the probability residing at the highest concentrations. This indicates that the upwelling during this season is too strong and it may also be too isolated since there appears to be more mixing in the measured JPDF. During SON, the separation between the segments is most prominent, indicating that this time of year exhibits the least mixing. Based on these comparisons, it is still difficult to discern the relative contributions of the tropical upwelling and quasi-horizontal mixing to the differences in the JPDFs of the model and measurements. To interpret these differences, the results of the TLP model simulations described previously are used. These simulations provide a basis for determining how much the residual circulation needs to slow down within CMAM30HR and what impact that may have on mixing of air between the tropical and extratropical regions.

6 Using a tropical leaky pipe model to interpret CMAM30HR

~~The TLP model is used here to identify the changes to~~ An analysis of the effects of the strength of the BDC and the mixing in CMAM30HR was achieved through a suite of simulations computed with the TLP to test a range of mean circulation strengths and mixing efficiencies. The TLP runs began with the CMAM30HR best fit to the ~~CMAM30HR tropical upwelling~~TLP model and then the best combination was selected to match the ACE-FTS measurement profiles of CFC-11 and ~~effective mixing that may improve the simulations. It is difficult to adjust model parametrizations to modify wave breaking because many aspects of the model would be impacted with no way of separating the effects (Ray et al., 2016). This includes the complexity of wave activity that contributes to stratospheric~~ CFC-12, as well as an age-of-air estimate derived from balloon-borne measurements (Ray et al., 2016). There were 480 simulations initialized with different combinations of $w*$ (velocity of tropical upwelling) and ϵ (the mixing efficiency) settings. The fraction of the CMAM30 $w*$used to initialize the TLP model ranged from 0.20 to 1.24 and the ϵ ranged from 0.18 to 1.50. In each TLP simulation run, the relationship between mean circulation and mixing ~~.~~ ~~Therefore, without a simplified model , diagnosing the causes of discrepancies between measurements and model output solely~~

from the global model is problematic. Prior to running the experiments used here, is constrained by the vertically-averaged mixing efficiency (Ray et al., 2016). The mixing efficiency in the TLP model is defined as $\epsilon = \alpha / \lambda\tau$, where α is the ratio of tropical to extratropical mass, λ is the rate of the mean circulation influence, specifically the mass flux out of the TLP model was tuned to be representative of CMAM30HR by fitting estimates of the mean tropical upwelling ($w*$) and mean mixing levels (ϵ). The tuning procedure is described in Ray et al. (2016)tropics due to mean circulation, and τ is the mixing time or time scale for mass flux between the tropics and extratropics (Ray et al., 2014; Garny et al., 2014). Therefore, the mixing efficiency is dependent on both the mean circulation as well as the horizontal mixing mass flux (Garny et al., 2014). The same set suite of TLP model experiments Ray et al. (2016) describe are also used here but the analysis has been modified to investigate the behaviour of the tropical pipe in CMAM30HR in further detail.

A reduced mean circulation would likely correspond to less mixing and longer mixing times. Mixing levels are defined in the TLP by the ratio of horizontal mixing mass flux to horizontal mean mass flux scaled by the width of the tropical pipe region (Garny et al., 2014). Ray et al. (2016) found that the CMAM30HR simulations best match the ACE-FTS measurements when the $w*$$w^*$ is between 0.27 mm/s and 0.32 mm/s 0.27 mm/s and 0.32 mm/s (a reduction from the fitting estimate of 0.4 mm/s) 0.4 mm/s) and ϵ ranges from 0.7 to 1.2 (an increase from the fitting estimate of 0.55). Based on the in-mixing time profiles shown by Ray et al. (2016), it is apparent that Ray et al. (2016)found that since ϵ is inversely proportional to both λ and τ, there is a compensating effect with changes in w^* (λ) or τ. For the CMAM30HR changes derived, w^* needed to be slowed down significantly below 20 km, and above 24 km. For constant ϵ that would result in larger τ (less mixing). However, Ray et al. (2016) found that ϵ also needed to be increased so there needed to be more mixing than would result from slower w^* and constant τ, but not enough of an increase in ϵ so that the mixing times were less (more mixing) than CMAM30HR has currently. Ray et al. (2016) concluded that with the increase in ϵ, mixing times need to increase to slow down the mixing at all levels, although the differences are only significant in the lower part of the stratosphere, below 20 km, and above 24 km. These are physically consistent changes since the mean circulation is driven by wave breaking, which also causes mixing between the tropics and extratropicstimes are reduced but still longer than the current CMAM30HR mixing times.

In this sectionstudy, the TLP simulations are used in a spatial and seasonal context to determine the changes required in the CMAM30HR simulations to match the ACE-FTS observations. The three regions investigated (the tropics and the northern and southern extratropics) were defined by the turn-around latitude of the tropical upwelling and by exclusion of the polar vortex in each hemisphere using a 1.2 x 10^{-4} s^{-1} 1.2 x 10^{-4} s^{-1} scaled potential vorticity threshold (e.g., Manney et al., 2007). The vortex is excluded in these comparisons because the TLP model does not simulate its complexity. For each of the 480 simulations of the TLP model and the ACE-FTS measurements, profiles for CFC-11 and CFC-12 were averaged over 16-27 km 16−27 km in the extratropics and 16-29 km 16−29 km in tropics. To capture the monthly coverage of the ACE-FTS measurements, a weighting function for each region and season was applied based on the relative contribution of the occultations observed in each month of the particular season and region (the tropics and northern/southern extratropics). The vertically averaged and monthly weighted measurements were compared to the individual simulations of the TLP model that represent varying levels of tropical upwelling and mixing efficiencies. For each of the comparisons, the absolute value of the differences is used and scaled to the maximum of the range of all differences so that each of the scales range from zero to one.

Figure 15 shows four examples of these comparisons plotted by the seasonally averaged, regionally specific modelled values of mixing, ϵ, and upwelling, $w*$, for each simulation. The x-axis and y-axis values do not represent the setting used to run the simulation - — they are the values simulated by the TLP model and lead to the curvature seen in the plots. The shading of each plot indicates the level of agreement, where the darker regions indicate the minimum differences between ACE-FTS and the TLP simulations. The white contours illustrate the agreement isolineswhite-to-black shading is reinforced by white contours of the same quantity to illustrate the comparison more clearly. The relative CMAM30HR position is shown by the red marker and the error bars represent the estimated range of uncertainty based on the optimization exercise with the TLP model. These comparisons all indicate that the CMAM30HR values of ϵ and $w*$ would require some change to bring the simulations closer to agreement with the ACE-FTS measurements (i.e. bringing the red marker towards the dark shaded region). The dependence of $w*$ on ϵ and vice versa is not the same across regions and seasons for either CFC-11 or CFC-12. There are often limited ranges in which agreement between the TLP simulations and ACE-FTS can be assessed but it is clear that changes could lead to improvements in the agreement (such as in Fig. 15b). There are some scenarios where the mixing is required to change much more than the tropical upwelling (see Fig. 15c for an example) or where there can be a range of mixing levels efficiency and upwelling values that lead to better agreement (as in Fig. 15d).

To quantify the The changes that would bring CMAM30HR into agreement with ACE-FTS based on the TLP simulation comparisons , an agreement matrix has for CFC-11 and CFC-12 have been calculated. This was done by finding where the differences between ACE-FTS and the TLP simulation are below a certain threshold (0.65 for tropics and 0.2 for extratropicsall regions and seasons) and determining the ranges of $w*$ and ϵ over which this occurs. The average values (square markerscyan 'x' markers for CFC-11 and blue square markers for CFC-12) and ranges (error bars) of the changes to $w*$ and ϵ calculated are shown in Fig. 16. The changes in $w*$ were calculated based on the absolute value of $w*$ so that the interpretation of the differences calculated was not dependent on the sign of $w*$. This means that a positive change is always an acceleration of the BDC, even when $w*$ is typically negative, such as in the extratropics. Figure 16a shows these changes for each region and across the four seasons. The sign of $w*$ values is indicated above Fig. 16a. The tropical region changes required are typically centered around zero for all seasons and have the largest range of changes of any region in any season, where the green symbols are for the Northern Hemisphere extratropics, the black symbols are for the tropics, and the red symbols are for the Southern Hemisphere extratropics. In some cases, the agreement between ACE-FTS and the TLP simulations did not meet the 0.2 threshold so there is no result available for Fig. 16.

The changes required to $w*$ in the tropical region are significant in the CFC-11 comparisons in two seasons: DJF and JJA, indicating that the $w*$ strength should be reduced by 0.05 mm/s to 0.10 mm/s in DJF and by 0.03 mm/s to 0.12 mm/s in JJA. The lack of a significant change during MAM and SON indicates that there is no argument for changes to the tropical upwelling in CMAM30HR. The Southern Hemisphere $w*$ changes are not significantly different from zero during any season, suggesting that changes in the $w*$ in this region would not improve the comparisons. However, there are significant changes identified in three of the four seasons in the extratropical region of the Northern Hemisphere, where the range in the change does not include zero. In DJF, an increase in $w*$ is required in the Northern Hemisphere by the CFC-12 comparisons, implying that strength of the downwelling during this season should increase by up to 0.08 mm/s. There is a similar requirement

~~during~~ 0.08 mm/s. However, the CFC-11 comparisons indicate that $w*$ should decrease by approximately 0.05 mm/s. This contradiction necessitates the conclusion that no change is recommended in $w*$ in the Northern Hemisphere extratropical region during DJF. During MAM in the Northern Hemisphere~~with a minimal range, requiring~~ , CFC-11 and CFC-12 comparisons suggest an increase in the downwelling by approximately ~~0.05 mm/s~~0.05 mm/s. By summer in the Northern Hemisphere (JJA), the tropical pipe has shifted poleward such that the Northern Hemisphere extratropical $w*$ values are positive and require a decrease in value as compared to the measurements. ~~This~~ While this season tends to have the most active tropical upwelling~~but~~ , these results suggest that the upwelling in CMAM30HR maybe too rapid or, since the southern extratropics appear to have increased downwelling during JJA, may be displaced in latitude.

The calculations of changes to the mixing parameter, ϵ, ~~provide more clarification, particularly for the tropical region. Figure 16bshows that tropical mixing needs to increase in every season. There appears to be a seasonal cycle in this result, where the maximum occurs during JJA. The preceding season remains at approximately the same average change required as DJF but the range is much more narrow leading up to the large increase in JJAand subsequent drop off in SON. This result quantifies and supports the idea that the difference in the separation of the JPDF segments between CMAM30HR and ACE-FTS is due to insufficient quasi-horizontal mixing between the tropics and extratropics~~shown in Fig. 16b, indicate that changes could be made in the tropics and extratropics differently across the seasons. The results indicate, based on CFC-11 and CFC-12 comparisons, that mixing efficiency in the tropics and in the Northern Hemisphere extratropics needs to increase during JJA. Based on the CFC-12 comparison, mixing efficiency in the Southern Hemisphere extratropics should also be increased during JJA. Significant changes are indicated from the CFC-11 comparisons for the remaining seasons in the extratropical regions. In the Northern Hemisphere and Southern Hemisphere during DJF, decreased mixing efficiency is suggested. During MAM, CFC-12 indicates there may be an increase suggested but the result is not supported by CFC-11 comparisons. In the ~~Northern Hemisphere, there are significant changes proposed in both MAM and JJA. 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 lowerstratosphere 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~~Southern Hemisphere, CFC-11 comparisons show that a decrease in mixing efficiency is necessary. During SON, CFC-11 comparisons also indicate a decrease in mixing efficiency in the Northern Hemisphere only. Generally, in all seasons except JJA, mixing efficiency in the extratropics is suggested to decrease.

Combining the ϵ parameter results with the decrease in $w*$ suggested in Fig. 16a, an increase in mixing time scales between the tropics and Northern Hemisphere extratropics is required during JJA. It follows that if $w*$ needs to be reduced in the model then a reduction in wave activity is required. The 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.

7 Conclusions

In this work, ACE-FTS measurements of CFC-11, CFC-12, and N_2O have been used to assess the CMAM30HR simulations of these tracers and, thereby, indirectly the transport processes in the lower stratosphere in the model. By treating each tracer in the specified dynamics simulation explicitly, the CMAM30HR run allows for the direct comparison of the measurements to model output. The advanced sampling technique employed here allows for detailed interpretation of the comparisons. Of the species investigated, it was found that CMAM30HR consistently overpredicts tracer concentrations in the lower to mid-stratosphere. The largest and most widespread overpredictions occur in the Northern Hemisphere winter and spring, when the BDC is most active in that hemisphere.

The investigation of simulated mixing barriers identified a number of issues in the CMAM30HR simulations. 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~~ BDC in CMAM30 simulations observed in the zonal mean comparisons. The extratropical tropopause barrier in the model appears to represent stratospheric ~~intrusions~~ intrusion events well, as evidenced by Fig. 12c and 12d. However, tropospheric intrusions are poorly simulated in most seasons (Fig. 12a and 12b), with the largest discrepancies occurring during JJA in both hemispheres. The tropical pipe mixing barrier analysis suggests that while the strength of the simulated BDC (i.e. upwelling in the tropics and downwelling in the extratropics) may partially explain the too young air found in the mid-stratosphere, mixing efficiency may play at least as prominent a role and seems to be underestimated particularly, in the JJA season in all regions.

~~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~~ The analysis presented here highlights the importance of scrutinizing the mixing efficiency in CCMs and GCMs since it ~~appears to~~ 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.

Acknowledgements. This project was funded by grants from the Canadian Space Agency (CSA) and the Natural Sciences and Engineering Research Council of Canada (NSERC). The Atmospheric Chemistry Experiment (ACE), also known as SCISAT, is a Canadian-led mission mainly supported by the CSA and the NSERC. The development of the CMAM30 data set was funded by the CSA. We thank Peter Bernath for his leadership of the ACE mission. We also thank Ted Shepherd, Dylan Jones, and John Scinocca for their leadership and support of the CMAM30 Project. The HATS measurements were funded in part by the Atmospheric Chemistry Project of the National Oceanographic and Atmospheric Administration (NOAA) Climate and Global Change Program. The authors thank Geoff Dutton, Debra Mondeel, and Steve Montzka for their work on measurements of these important trace gases in the HATS program.

References

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

Abrams, M. C., Manney, G. L., Gunson, M. R., Abbas, M. M., Chang, A. Y., Goldman, A., Irion, F. W., Michelsen, H. A., Newchurch, M. J., Rinsland, C. P., Salawitch, R. J., Stiller, G. P., and Zander, R.: Trace gas transport in the Arctic Vortex inferred from ATMOS ATLAS-2 observations during April 1993, Geophys. Res. Lett., 23, 2345–2348, doi:10.1029/96GL00704, 1996a.

Abrams, M. C., Manney, G. L., Gunson, M. R., Abbas, M. M., Chang, A. Y., Goldman, A., Irion, F. W., Michelsen, H. A., Newchurch, M. J., Rinsland, C. P., Salawitch, R. J., Stiller, G. P., and Zander, R.: ATMOS/ATLAS3 observations of long-lived tracers and descent in the Antarctic Vortex in November 1994, Geophys. Res. Lett., 23, 2341–2344, doi:10.1029/96GL00705, 1996b.

Allen, D. R., Stanford, J. L., Nakamura, N., López Valverde, M. A., López-Puertas, M., Taylor, F. W., and Remedios, J. J.: Antarctic polar descent and planetary wave activity observed in ISAMS CO from April to July 1992, Geophys. Res. Lett., 27, 665–668, doi:10.1029/1999GL010888, 2000.

Allen, N. D. C., Bernath, P. F., Boone, C. D., Chipperfield, M. P., Fu, D., Manney, G. L., Oram, D. E., Toon, G. C., and Weisenstein, D. K.: Global carbon tetrachloride distributions obtained from the Atmospheric Chemistry Experiment (ACE), Atmos. Chem. Phys., 9, 7449–7459, doi:10.5194/acp-9-7449-2009, 2009.

Austin, J., Shindell, D., Beagley, S. R., Brühl, C., Dameris, M., Manzini, E., Nagashima, T., Newman, P., Pawson, S., Pitari, G., Rozanov, E., Schnadt, C., and Shepherd, T. G.: Uncertainties and assessments of chemistry-climate models of the stratosphere, Atmos. Chem. Phys., 3, 1–27, doi:10.5194/acp-3-1-2003, 2003.

Beagley, S. R., de Grandpré, J., Koshyk, J. N., McFarlane, N. A., and Shepherd, T. G.: Radiative-dynamical climatology of the first-generation Canadian middle atmosphere model, Atmosphere-Ocean, 35, 293–331, doi:10.1080/07055900.1997.9649595, 1997.

Bernath, P. F.: The Atmospheric Chemistry Experiment (ACE), J. Quant. Spectrosc. Radiat. Transf., 186, 3 – 16, doi:10.1016/j.jqsrt.2016.04.006, 2017.

Bernath, P. F., McElroy, C. T., Abrams, M. C., Boone, C. D., Butler, M., Camy-Peyret, C., Carleer, M., Clerbaux, C., Coheur, P. F., Colin, R., DeCola, P., DeMazière, M., Drummond, J. R., Dufour, D., Evans, W. F. J., Fast, H., Fussen, D., Gilbert, K., Jennings, D. E., Llewellyn, E. J., Lowe, R. P., Mahieu, E., McConnell, J. C., McHugh, M., McLeod, S. D., Michaud, R., Midwinter, C., Nassar, R., Nichitiu, F., Nowlan, C., Rinsland, C. P., Rochon, Y. J., Rowlands, N., Semeniuk, K., Simon, P., Skelton, R., Sloan, J. J., Soucy, M. A., Strong, K., Tremblay, P., Turnbull, D., Walker, K. A., Walkty, I., Wardle, D. A., Wehrle, V., Zander, R., and Zou, J.: Atmospheric Chemistry Experiment (ACE): Mission overview, Geophys. Res. Lett., 32, doi:10.1029/2005GL022386, 2005.

Birner, T. and Bönisch, H.: Residual circulation trajectories and transit times into the extratropical lowermost stratosphere, Atmos. Chem. Phys., 11, 817–827, doi:10.5194/acp-11-817-2011, 2011.

Boone, C. D., Nassar, R., Walker, K. A., Rochon, Y., McLeod, S. D., Rinsland, C. P., and Bernath, P. F.: Retrievals for the atmospheric chemistry experiment Fourier-transform spectrometer, Appl. Opt., 44, 7218–7231, doi:10.1364/AO.44.007218, 2005.

Boone, C. D., Walker, K. A., and Bernath, P. F.: Version 3 retrievals for the atmospheric chemistry experiment Fourier transform spectrometer (ACE-FTS), The Atmospheric Chemistry Experiment ACE at 10: A Solar Occultation Anthology, A. Deepak Publishing, Hampton, Va., 2013.

Brewer, A. W.: Evidence for a world circulation provided by the measurements of helium and water vapour distribution in the stratosphere, Q. J. Roy. Meteor. Soc., 75, 351–363, doi:10.1002/qj.49707532603, 1949.

Brown, A. T., Chipperfield, M. P., Boone, C., Wilson, C., Walker, K. A., and Bernath, P. F.: Trends in atmospheric halogen containing gases since 2004, J. Quant. Spectrosc. Radiat. Transf., 112, 2552–2566, doi:10.1016/j.jqsrt.2011.07.005, 2011.

Brown, A. T., Volk, C. M., Schoeberl, M. R., Boone, C. D., and Bernath, P. F.: Stratospheric lifetimes of CFC-12, CCl_4, CH_4, CH_3Cl and N_2O from measurements made by the Atmospheric Chemistry Experiment-Fourier Transform Spectrometer (ACE-FTS), Atmos. Chem. Phys., 13, 6921–6950, doi:10.5194/acp-13-6921-2013, 2013.

Brown, A. T., Chipperfield, M. P., Richards, N. A. D., Boone, C., and Bernath, P. F.: Global stratospheric fluorine inventory for 2004–2009 from Atmospheric Chemistry Experiment Fourier Transform Spectrometer (ACE-FTS) measurements and SLIMCAT model simulations, Atmos. Chem. Phys., 14, 267–282, doi:10.5194/acp-14-267-2014, 2014.

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

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge Sanz, B. M., Morcrette, J. J., Park, B. K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J. N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. Roy. Meteorol. Soc., 137, 553–597, doi:10.1002/qj.828, 2011.

deGrandpré, J., Beagley, S. R., Fomichev, V. I., Griffioen, E., McConnell, J. C., Medvedev, A. S., and Shepherd, T. G.: Ozone climatology using interactive chemistry: Results from the Canadian Middle Atmosphere Model, J. Geophys. Res., 105, 26 475–26 491, doi:10.1029/2000JD900427, 2000.

Diallo, M., Legras, B., and Chédin, A.: Age of stratospheric air in the ERA-Interim, Atmos. Chem. Phys., 12, 12 133–12 154, doi:10.5194/acp-12-12133-2012, 2012.

Dickinson, R. E.: Vertical propagation of planetary rossby waves through an atmosphere with Newtonian cooling, J. Geophys. Res., 74, 929–938, doi:10.1029/JB074i004p00929, 1969.

Dobson, G. M. B.: Origin and Distribution of the Polyatomic Molecules in the Atmosphere, Proc. Roy. Soc., 236, 187–193, doi:10.1098/rspa.1956.0127, 1956.

Dobson, G. M. B., Harrison, D. N., and Lawrence, J.: Measurements of the Amount of Ozone in the Earth's Atmosphere and Its Relation to Other Geophysical Conditions. Part III, Proc. Roy. Soc., 122, 456–486, doi:10.1098/rspa.1929.0034, 1929.

Elkins, J. W., Thompson, T. M., Swanson, T. H., Butler, J. H., Hall, B. D., Cummings, S. O., Fisher, D. A., and Raffo, A. G.: Decrease in the growth rates of atmospheric chlorofluorocarbons 11 and 12, Nature, 364, doi:10.1038/364780a0, 1993.

Engel, A., Möbius, T., Bönisch, H., Schmidt, U., Heinz, R., Levin, I., Atlas, E., Aoki, S., Nakazawa, T., Sugawara, S., Moore, F., Hurst, D., Elkins, J., Schauffler, S., Andrews, A., and Boering, K.: Age of stratospheric air unchanged within uncertainties over the past 30 years, Nat. Geosci., 2, 28–31, doi:10.1038/NGEO388, 2009.

Eyring, V., Harris, N. R. P., Rex, M., Shepherd, T. G., Fahey, D. W., Amanatidis, G. T., Austin, J., Chipperfield, M. P., Dameris, M., Forster, P. M. D. F., Gettelman, A., Graf, H. F., Nagashima, T., Newman, P. A., Pawson, S., Prather, M. J., Pyle, J. A., Salawitch, R. J., Santer, B. D., and Waugh, D. W.: A Strategy for Process-Oriented Validation of Coupled Chemistry-Climate Models, Bull. Amer. Meteor. Soc., 86, 1117–1133, doi:10.1175/BAMS-86-8-1117, 2005.

Funke, B., García-Comas, M., López-Puertas, M., Glatthor, N., Stiller, G. P., von Clarmann, T., Semeniuk, K., and McConnell, J. C.: Enhancement of N_2O during the October–November 2003 solar proton events, Atmos. Chem. Phys., 8, 3805–3815, doi:10.5194/acp-8-3805-2008, 2008a.

Funke, B., López-Puertas, M., Garcia-Comas, M., Stiller, G. P., von Clarmann, T., and Glatthor, N.: Mesospheric N_2O enhancements as observed by MIPAS on Envisat during the polar winters in 2002–2004, Atmos. Chem. Phys., 8, 5787–5800, doi:10.5194/acp-8-5787-2008, 2008b.

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

Gerber, E. P.: Stratospheric versus Tropospheric Control of the Strength and Structure of the Brewer-Dobson Circulation, J. Atmos. Sci., 69, 2857–2877, doi:10.1175/JAS-D-11-0341.1, 2012.

Greenblatt, J. B.: Correction to "Tracer-based determination of vortex descent in the 1999/2000 Arctic winter" by J. B. Greenblatt et al., J. Geophys. Res., 107, doi:10.1029/2002JD001597, 2003.

Greenblatt, J. B., Jost, H. J., Loewenstein, M., Podolske, J. R., Hurst, D. F., Elkins, J. W., Schauffler, S. M., Atlas, E. L., Herman, R. L., Webster, C. R., Bui, T. P., Moore, F. L., Ray, E. A., Oltmans, S., Vömel, H., Blavier, J. F., Sen, B., Stachnik, R. A., Toon, G. C., Engel, A., Müller, M., Schmidt, U., Bremer, H., Pierce, R. B., Sinnhuber, B. M., Chipperfield, M., and Lefèvre, F.: Tracer-based determination of vortex descent in the 1999/2000 Arctic winter, J. Geophys. Res., 107, doi:10.1029/2001JD000937, 2002.

Haenel, F. J., Stiller, G. P., von Clarmann, T., Funke, B., Eckert, E., Glatthor, N., Grabowski, U., Kellmann, S., Kiefer, M., Linden, A., and Reddmann, T.: Reassessment of MIPAS age of air trends and variability, Atmos. Chem. Phys., 15, 13 161–13 176, doi:10.5194/acp-15-13161-2015, 2015.

Hall, T. M. and Plumb, R. A.: Age as a diagnostic of stratospheric transport, J. Geophys. Res., 99, 1059–1070, doi:10.1029/93JD03192, 1994.

Hall, T. M. and Waugh, D. W.: Stratospheric residence time and its relationship to mean age, J. Geophys. Res., 105, 6773–6782, doi:10.1029/1999JD901096, 2000.

Hardiman, S. C., Lin, P., Scaife, A. A., Dunstone, N. J., and Ren, H.-L.: The influence of dynamical variability on the observed Brewer-Dobson circulation trend, Geophys. Res. Lett., 44, 2885–2892, doi:10.1002/2017GL072706, 2017.

Haynes, P. and Shuckburgh, E.: Effective diffusivity as a diagnostic of atmospheric transport: 2. Troposphere and lower stratosphere, J. Geophys. Res., 105, 22 795–22 810, doi:10.1029/2000JD900092, 2000.

Hegglin, M. I. and Shepherd, T. G.: O_3-N_2O correlations from the Atmospheric Chemistry Experiment: Revisiting a diagnostic of transport and chemistry in the stratosphere, J. Geophys. Res., 112, doi:10.1029/2006JD008281, 2007.

Hegglin, M. I., Boone, C. D., Manney, G. L., Shepherd, T. G., Walker, K. A., Bernath, P. F., Daffer, W. H., Hoor, P., and Schiller, C.: Validation of ACE-FTS satellite data in the upper troposphere/lower stratosphere (UTLS) using non-coincident measurements, Atmos. Chem. Phys., 8, 1483–1499, doi:10.5194/acp-8-1483-2008, 2008.

Hegglin, M. I., Boone, C. D., Manney, G. L., and Walker, K. A.: A global view of the extratropical tropopause transition layer from Atmospheric Chemistry Experiment Fourier Transform Spectrometer O_3, H_2O, and CO, J. Geophys. Res., 114, D00B11, doi:10.1029/2008JD009984, 2009.

Hegglin, M. I., Tegtmeier, S., Anderson, J., Froidevaux, L., Fuller, R., Funke, B., Jones, A., Lingenfelser, G., Lumpe, J., Pendlebury, D., Remsberg, E., Rozanov, A., Toohey, M., Urban, J., von Clarmann, T., Walker, K. A., Wang, R., and Weigel, K.: SPARC Data Initiative: Comparison of water vapor climatologies from international satellite limb sounders, J. Geophys. Res., 118, 11,824–11,846, doi:10.1002/jgrd.50752, 2013.

Hegglin, M. I., Plummer, D. A., Shepherd, T. G., Scinocca, J. F., Anderson, J., Froidevaux, L., Funke, B., Hurst, D., Rozanov, A., Urban, J., von Clarmann, T., Walker, K. A., Wang, H. J., Tegtmeier, S., and Weigel, K.: Vertical structure of stratospheric water vapour trends derived from merged satellite data, Nat. Geosci., 7, 768–776, doi:10.1038/NGEO2236, 2014.

Hoffmann, L., Hoppe, C. M., Müller, R., Dutton, G. S., Gille, J. C., Griessbach, S., Jones, A., Meyer, C. I., Spang, R., Volk, C. M., and Walker, K. A.: Stratospheric lifetime ratio of CFC-11 and CFC-12 from satellite and model climatologies, Atmos. Chem. Phys., 14, 12 479–12 497, doi:10.5194/acp-14-12479-2014, 2014.

Holton, J. R., Haynes, P. H., McIntyre, M. E., Douglass, A. R., Rood, R. B., and Pfister, L.: Stratosphere-troposphere exchange, Rev. Geophys., 33, 403–439, doi:10.1029/95RG02097, 1995.

Jin, J. J., Semeniuk, K., Jonsson, A. I., Beagley, S. R., McConnell, J. C., Boone, C. D., Walker, K. A., Bernath, P. F., Rinsland, C. P., Dupuy, E., Ricaud, P., De La Noë, J., Urban, J., and Murtagh, D.: Co-located ACE-FTS and Odin/SMR stratospheric-mesospheric CO 2004 measurements and comparison with a GCM, Geophys. Res. Lett., 32, doi:10.1029/2005GL022433, 2005.

Jones, A., Walker, K. A., Jin, J. J., Taylor, J. R., Boone, C. D., Bernath, P. F., Brohede, S., Manney, G. L., McLeod, S., Hughes, R., and Daffer, W. H.: Technical Note: A trace gas climatology derived from the Atmospheric Chemistry Experiment Fourier Transform Spectrometer (ACE-FTS) data set, Atmos. Chem. Phys., 12, 5207–5220, 2012.

Jonsson, A. I., de Grandpré, J., Fomichev, V. I., McConnell, J. C., and Beagley, S. R.: Doubled CO_2-induced cooling in the middle atmosphere: Photochemical analysis of the ozone radiative feedback, J. Geophys. Res., 109, D24 103, doi:10.1029/2004JD005093, 2004.

Kawamoto, N. and Shiotani, M.: Interannual variability of the vertical descent rate in the Antarctic polar vortex, J. Geophys. Res., 105, 11 935–11 946, doi:10.1029/2000JD900076, 2000.

Koo, J.-H., Walker, K. A., Jones, A., Sheese, P. E., Boone, C. D., Bernath, P. F., and Manney, G. L.: Global climatology based on the ACE-FTS version 3.5 dataset: Addition of mesospheric levels and carbon-containing species in the UTLS, J. Quant. Spectrosc. Radiat. Transf., 186, 52 – 62, doi:10.1016/j.jqsrt.2016.07.003, 2017.

Lin, P. and Fu, Q.: Changes in various branches of the Brewer-Dobson circulation from an ensemble of chemistry climate models, J. Geophys. Res., 118, 73–84, doi:10.1029/2012JD018813, 2013.

Mahieu, E., Duchatelet, P., Demoulin, P., Walker, K. A., Dupuy, E., Froidevaux, L., Randall, C., Catoire, V., Strong, K., Boone, C. D., Bernath, P. F., Blavier, J.-F., Blumenstock, T., Coffey, M., De Mazière, M., Griffith, D., Hannigan, J., Hase, F., Jones, N., Jucks, K. W., Kagawa, A., Kasai, Y., Mebarki, Y., Mikuteit, S., Nassar, R., Notholt, J., Rinsland, C. P., Robert, C., Schrems, O., Senten, C., Smale, D., Taylor, J., Tétard, C., Toon, G. C., Warneke, T., Wood, S. W., Zander, R., and Servais, C.: Validation of ACE-FTS v2.2 measurements of HCl, HF, CCl_3F and CCl_2F_2 using space-, balloon- and ground-based instrument observations, Atmos. Chem. Phys., 8, 6199–6221, doi:10.5194/acp-8-6199-2008, 2008.

Mahieu, E., Chipperfield, M. P., Notholt, J., Reddmann, T., Anderson, J., Bernath, P. F., Blumenstock, T., Coffey, M. T., Dhomse, S. S., Feng, W., Franco, B., Froidevaux, L., Griffith, D. W. T., Hannigan, J. W., Hase, F., Hossaini, R., Jones, N. B., Morino, I., Murata, I., Nakajima, H., Palm, M., Paton-Walsh, C., Russell, III, J. M., Schneider, M., Servais, C., Smale, D., and Walker, K. A.: Recent Northern Hemisphere stratospheric HCl increase due to atmospheric circulation changes, Nature, 515, 104–107, doi:10.1038/nature13857, 2014.

Manney, G. L., Daffer, W. H., Zawodny, J. M., Bernath, P. F., Hoppel, K. W., Walker, K. A., Knosp, B. W., Boone, C., Remsberg, E. E., Santee, M. L., Harvey, V. L., Pawson, S., Jackson, D. R., Deaver, L., McElroy, C. T., McLinden, C. A., Drummond, J. R., Pumphrey, H. C., Lambert, A., Schwartz, M. J., Froidevaux, L., McLeod, S., Takacs, L. L., Suarez, M. J., Trepte, C. R., Cuddy, D. C., Livesey, N. J., Harwood, R. S., and Waters, J. W.: Solar occultation satellite data and derived meteorological products: Sampling issues and comparisons with Aura Microwave Limb Sounder, J. Geophys. Res., 112, doi:10.1029/2007JD008709, 2007.

Manney, G. L., Harwood, R. S., MacKenzie, I. A., Minschwaner, K., Allen, D. R., Santee, M. L., Walker, K. A., Hegglin, M. I., Lambert, A., Pumphrey, H. C., Bernath, P. F., Boone, C. D., Schwartz, M. J., Livesey, N. J., Daffer, W. H., and Fuller, R. A.: Satellite observations and modeling of transport in the upper troposphere through the lower mesosphere during the 2006 major stratospheric sudden warming, Atmos. Chem. Phys., 9, 4775–4795, doi:10.5194/acp-9-4775-2009, 2009.

McIntyre, M. E. and Palmer, T. N.: Breaking planetary waves in the stratosphere, Nature, 305, doi:10.1038/305593a0, 1983.

McIntyre, M. E. and Palmer, T. N.: The 'surf zone' in the stratosphere, J. Atmos. Terr. Phys., 46, 825–849, doi:10.1016/0021-9169(84)90063-1, 1984.

McLandress, C., Shepherd, T. G., Scinocca, J. F., Plummer, D. A., Sigmond, M., Jonsson, A. I., and Reader, M. C.: Separating the Dynamical Effects of Climate Change and Ozone Depletion. Part II: Southern Hemisphere Troposphere, J. Climate, 24, 1850–1868, doi:10.1175/2010JCLI3958.1, 2011.

McLandress, C., Shepherd, T. G., Polavarapu, S., and Beagley, S. R.: Is Missing Orographic Gravity Wave Drag near 60°S the Cause of the Stratospheric Zonal Wind Biases in Chemistry-Climate Models?, J. Atmos. Sci., 69, 802–818, 2012.

McLandress, C., Scinocca, J. F., Shepherd, T. G., Reader, M. C., and Manney, G. L.: Dynamical Control of the Mesosphere by Orographic and Nonorographic Gravity Wave Drag during the Extended Northern Winters of 2006 and 2009, J. of Atmos. Sci., 70, doi:10.1175/JAS-D-12-0297.1, 2013.

McLandress, C., Shepherd, T. G., Reader, M. C., Plummer, D. A., and Shine, K. P.: The Climate Impact of Past Changes in Halocarbons and CO_2 in the Tropical UTLS Region, J. Climate, 27, 8646–8660, doi:10.1175/JCLI-D-14-00232.1, 2014.

Meinshausen, M., Vogel, E., Nauels, A., Lorbacher, K., Meinshausen, N., Etheridge, D. M., Fraser, P. J., Montzka, S. A., Rayner, P. J., Trudinger, C. M., Krummel, P. B., Beyerle, U., Canadell, J. G., Daniel, J. S., Enting, I. G., Law, R. M., Lunder, C. R., O'Doherty, S., Prinn, R. G., Reimann, S., Rubino, M., Velders, G. J. M., Vollmer, M. K., Wang, R. H. J., and Weiss, R.: Historical greenhouse gas concentrations for climate modelling (CMIP6), Geosci. Model Develop., 10, 2057–2116, doi:10.5194/gmd-10-2057-2017, 2017.

Merrifield, M. A., Thompson, P. R., and Lander, M.: Multidecadal sea level anomalies and trends in the western tropical Pacific, Geophys. Res. Lett., 39, doi:10.1029/2012GL052032, 2012.

Millán, L. F., Livesey, N. J., Santee, M. L., Neu, J. L., Manney, G. L., and Fuller, R. A.: Case studies of the impact of orbital sampling on stratospheric trend detection and derivation of tropical vertical velocities: solar occultation vs. limb emission sounding, Atmos. Chem. Phys., 16, 11 521–11 534, doi:10.5194/acp-16-11521-2016, 2016.

Montzka, S. A., Butler, J. H., Myers, R. C., Thompson, T. M., Swanson, T. H., Clarke, A. D., Lock, L. T., and Elkins, J. W.: Decline in the Tropospheric Abundance of Halogen from Halocarbons: Implications for Stratospheric Ozone Depletion, Science, 272, 1318–1322, doi:10.1126/science.272.5266.1318, 1996.

Mote, P. W., Rosenlof, K. H., McIntyre, M. E., Carr, E. S., Gille, J. C., Holton, J. R., Kinnersley, J. S., Pumphrey, H. C., Russell, III, J. M., and Waters, J. W.: An atmospheric tape recorder: The imprint of tropical tropopause temperatures on stratospheric water vapor, J. Geophys. Res., 101, 3989–4006, doi:10.1029/95JD03422, 1996.

Neu, J. L. and Plumb, R. A.: Age of air in a "leaky pipe" model of stratospheric transport, J. Geophys. Res., 104, 19 243–19 255, doi:10.1029/1999JD900251, 1999.

Neu, J. L., Hegglin, M. I., Tegtmeier, S., Bourassa, A., Degenstein, D., Froidevaux, L., Fuller, R., Funke, B., Gille, J., Jones, A., Rozanov, A., Toohey, M., von Clarmann, T., Walker, K. A., and Worden, J. R.: The SPARC Data Initiative: Comparison of upper troposphere/lower stratosphere ozone climatologies from limb-viewing instruments and the nadir-viewing Tropospheric Emission Spectrometer, J. Geophys. Res., 119, 6971–6990, doi:10.1002/2013JD020822, 2014.

Niwano, M., Yamazaki, K., and Shiotani, M.: Seasonal and QBO variations of ascent rate in the tropical lower stratosphere as inferred from UARS HALOE trace gas data, J. Geophys. Res., 108, doi:10.1029/2003JD003871, 2003.

Oberländer-Hayn, S., Gerber, E., Abalichin, J., Akiyoshi, H., Kerschbaumer, A., Kubin, A., Kunze1, M., Langematz, U., Meul, S., Michou, M., Morgenstern, O., and Oman, L. D.: Is the Brewer-Dobson circulation increasing or moving upward?, Geophys. Res. Lett., 43, 1772–1779, doi:10.1002/2015GL067545, 2016.

Pan, L. L., Randel, W. J., Gille, J. C., Hall, W. D., Nardi, B., Massie, S., Yudin, V., Khosravi, R., Konopka, P., and Tarasick, D.: Tropospheric intrusions associated with the secondary tropopause, J. of Geophys. Res., 114, D10 302, doi:10.1029/2008JD011374, 2009.

Park, M., Randel, W. J., Kinnison, D. E., Emmons, L. K., Bernath, P. F., Walker, K. A., Boone, C. D., and Livesey, N. J.: Hydrocarbons in the upper troposphere and lower stratosphere observed from ACE-FTS and comparisons with WACCM, J. of Geophys. Res., 118, 1964–1980, doi:10.1029/2012JD018327, 2013.

Pawson, S., Kodera, K., Hamilton, K., Shepherd, T. G., Beagley, S. R., Boville, B. A., Farrara, J. D., Fairlie, T. D. A., Kitoh, A., Lahoz, W. A., Langematz, U., Manzini, E., Rind, D. H., Scaife, A. A., Shibata, K., Simon, P., Swinbank, R., Takacs, L., Wilson, R. J., Al-Saadi, J. A., Amodei, M., Chiba, M., Coy, L., de Grandpré, J., Eckman, R. S., Fiorino, M., Grose, W. L., Koide, H., Koshyk, J. N., Li, D., Lerner, J., Mahlman, J. D., McFarlane, N. A., Mechoso, C. R., Molod, A., O'Neill, A., Pierce, R. B., Randel, W. J., Rood, R. B., and Wu, F.: The GCM-reality intercomparison project for SPARC (GRIPS): Scientific issues and initial results, Bull. Amer. Meteor. Soc., 81, 781–796, doi:10.1175/1520-0477(2000)081<0781:TGIPFS>2.3.CO;2, 2000.

Peevey, T. R., Gille, J. C., Homeyer, C. R., and Manney, G. L.: The double tropopause and its dynamical relationship to the tropopause inversion layer in storm track regions, J. Geophys. Res., 119, 10,194–10,212, doi:10.1002/2014JD021808, 2014.

Pendlebury, D., Plummer, D., Scinocca, J., Sheese, P., Strong, K., Walker, K., and Degenstein, D.: Comparison of the CMAM30 data set with ACE-FTS and OSIRIS: polar regions, Atmos. Chem. Phys., 15, 12 465–12 485, doi:10.5194/acp-15-12465-2015, 2015.

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

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

Plumb, R. A.: A "tropical pipe" model of stratospheric transport, J. Geophys. Res., 101, 3957–3972, doi:10.1029/95JD03002, 1996.

Plumb, R. A.: Stratospheric Transport, J. Meteorol. Soc. Japan, 80, 793–809, doi:10.2151/jmsj.80.793, 2002.

Plumb, R. A. and Eluszkiewicz, J.: The Brewer-Dobson Circulation: Dynamics of the Tropical Upwelling, J. Atmos. Sci., 56, 868–890, doi:10.1175/1520-0469(1999)056<0868:TBDCDO>2.0.CO;2, 1999.

Plumb, R. A. and Ko, M. K. W.: Interrelationships between mixing ratios of long-lived stratospheric constituents, J. Geophys. Res., 97, 10 145–10 156, doi:10.1029/92JD00450, 1992.

Plummer, D. A., Scinocca, J. F., Shepherd, T. G., Reader, M. C., and Jonsson, A. I.: Quantifying the contributions to stratospheric ozone changes from ozone depleting substances and greenhouse gases, Atmos. Chem. Phys., 10, 8803–8820, doi:10.5194/acp-10-8803-2010, 2010.

Ray, E. A., Moore, F. L., Elkins, J. W., Hurst, D. F., Romashkin, P. A., Dutton, G. S., and Fahey, D. W.: Descent and mixing in the 1999–2000 northern polar vortex inferred from in situ tracer measurements, J. Geophys. Res., 107, doi:10.1029/2001JD000961, 2002.

Ray, E. A., Moore, F. L., Rosenlof, K. H., Davis, S. M., Sweeney, C., Tans, P., Wang, T., Elkins, J. W., Bönisch, H., Engel, A., Sugawara, S., Nakazawa, T., and Aoki, S.: Improving stratospheric transport trend analysis based on SF_6 and CO_2 measurements, J. of Geophys. Res., 119, doi:10.1002/2014JD021802, 2014.

Ray, E. A., Moore, F. L., Rosenlof, K. H., Plummer, D. A., Kolonjari, F., and Walker, K. A.: An idealized stratospheric model useful for understanding differences between long-lived trace gas measurements and global chemistry-climate model output, J. Geophys. Res., 121, 5356–5367, doi:10.1002/2015JD024447, 2016.

Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., Kent, E. C., and Kaplan, A.: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, J. Geophys. Res., 108, doi:10.1029/2002JD002670, 2003.

Rosenlof, K. H.: Seasonal cycle of the residual mean meridional circulation in the stratosphere, J. Geophys. Res., 100, 5173–5191, doi:10.1029/94JD03122, 1995.

Sander, S. P., Friedl, R. R., Abbatt, J. P. D., Barker, J. R., Burkholder, J. B., Golden, D. M., Kolb, C. E., Kurylo, M. J., Moortgat, G. K., Wine, P. H., Huie, R. E., and Orkin, V. L.: Chemical Kinetics and Photochemical Data for Use in Atmospheric Modeling, Evaluation Number 17, JPL Publication 10-6, Jet Propulsion Laboratory, Pasadena, USA, 2011.

Sankey, D. and Shepherd, T. G.: Correlations of long-lived chemical species in a middle atmosphere general circulation model, J. Geophys. Res., 108, doi:10.1029/2002JD002799, 2003.

Schauffler, S. M., Atlas, E. L., Donnelly, S. G., Andrews, A., Montzka, S. A., Elkins, J. W., Hurst, D. F., Romashkin, P. A., Dutton, G. S., and Stroud, V.: Chlorine budget and partitioning during the Stratospheric Aerosol and Gas Experiment (SAGE) III Ozone Loss and Validation Experiment (SOLVE), J. Geophys. Res., 108, doi:10.1029/2001JD002040, 2003.

Schwartz, M. J., Manney, G. L., Hegglin, M. I., Livesey, N. J., Santee, M. L., and Daffer, W. H.: Climatology and variability of trace gases in extratropical double-tropopause regions from MLS, HIRDLS, and ACE-FTS measurements, J. Geophys. Res., 120, 843–867, doi:10.1002/2014JD021964, 2015.

Scinocca, J. F., McFarlane, N. A., Lazare, M., Li, J., and Plummer, D.: Technical Note: The CCCma third generation AGCM and its extension into the middle atmosphere, Atmos. Chem. Phys., 8, 7055–7074, doi:10.5194/acp-8-7055-2008, 2008.

Seviour, W. J. M., Butchart, N., and Hardiman, S. C.: The Brewer-Dobson circulation inferred from ERA-Interim, Q. J. Roy. Meteorol. Soc., 138, 878–888, 2012.

Sheese, P. E., Boone, C. D., and Walker, K. A.: Detecting physically unrealistic outliers in ACE-FTS atmospheric measurements, Atmos. Meas. Tech., 8, 741–750, doi:10.5194/amt-8-741-2015, 2015.

Sheese, P. E., Walker, K. A., Boone, C. D., Bernath, P. F., and Funke, B.: Nitrous oxide in the atmosphere: First measurements of a lower thermospheric source, Geophys. Res. Lett., 43, 2866–2872, doi:10.1002/2015GL067353, 2016.

Shepherd, T. G.: Transport in the Middle Atmosphere, J. Meteorol. Soc. Japan, 85B, 165–191, doi:10.2151/jmsj.85B.165, 2007.

Shepherd, T. G. and McLandress, C.: A Robust Mechanism for Strengthening of the Brewer-Dobson Circulation in Response to Climate Change: Critical-Layer Control of Subtropical Wave Breaking, J. Atmos. Sci., 68, 784–797, doi:10.1175/2010JAS3608.1, 2011.

Shepherd, T. G., Plummer, D. A., Scinocca, J. F., Hegglin, M. I., Fioletov, V. E., Reader, M. C., Remsberg, E., von Clarmann, T., and Wang, H. J.: Reconciliation of halogen-induced ozone loss with the total-column ozone record, Nat. Geosci., 7, 443–449, doi:10.1038/NGEO2155, 2014.

SPARC-CCMVal: (Stratospheric Processes And their Role in Climate): Report on the Evaluation of Chemistry-Climate Models, SPARC report No. 5, WCRP-132, WMO/TD-No. 1526, 2010.

Sparling, L. C.: Statistical perspectives on stratospheric transport, Rev. Geophys., 38, 417–436, doi:10.1029/1999RG000070, 2000.

Stiller, G. P., von Clarmann, T., Höpfner, M., Glatthor, N., Grabowski, U., Kellmann, S., Kleinert, A., Linden, A., Milz, M., Reddmann, T., Steck, T., Fischer, H., Funke, B., López-Puertas, M., and Engel, A.: Global distribution of mean age of stratospheric air from MIPAS SF_6 measurements, Atmos. Chem. Phys., 8, 677–695, doi:10.5194/acp-8-677-2008, 2008.

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

Strahan, S. E., Douglass, A. R., Stolarski, R. S., Akiyoshi, H., Bekki, S., Braesicke, P., Butchart, N., Chipperfield, M. P., Cugnet, D., Dhomse, S., Frith, S. M., Gettelman, A., Hardiman, S. C., Kinnison, D. E., Lamarque, J. F., Mancini, E., Marchand, M., Michou, M., Morgenstern, O., Nakamura, T., Olivié, D., Pawson, S., Pitari, G., Plummer, D. A., Pyle, J. A., Scinocca, J. F., Shepherd, T. G., Shibata, K., Smale, D., Teyssèdre, H., Tian, W., and Yamashita, Y.: Using transport diagnostics to understand chemistry climate model ozone simulations, J. Geophys. Res., 116, doi:10.1029/2010JD015360, 2011.

Strong, K., Wolff, M. A., Kerzenmacher, T. E., Walker, K. A., Bernath, P. F., Blumenstock, T., Boone, C., Catoire, V., Coffey, M., De Mazière, M., Demoulin, P., Duchatelet, P., Dupuy, E., Hannigan, J., Höpfner, M., Glatthor, N., Griffith, D. W. T., Jin, J. J., Jones, N., Jucks, K., Kuellmann, H., Kuttippurath, J., Lambert, A., Mahieu, E., McConnell, J. C., Mellqvist, J., Mikuteit, S., Murtagh, D. P., Notholt, J., Piccolo, C., Raspollini, P., Ridolfi, M., Robert, C., Schneider, M., Schrems, O., Semeniuk, K., Senten, C., Stiller, G. P., Strandberg, A., Taylor, J., Tétard, C., Toohey, M., Urban, J., Warneke, T., and Wood, S.: Validation of ACE-FTS N_2O measurements, Atmos. Chem. Phys., 8, 4759–4786, doi:10.5194/acp-8-4759-2008, 2008.

Tegtmeier, S., Hegglin, M. I., Anderson, J., Bourassa, A., Brohede, S., Degenstein, D., Froidevaux, L., Fuller, R., Funke, B., Gille, J., Jones, A., Kasai, Y., Krüger, K., Kyrölä, E., Lingenfelser, G., Lumpe, J., Nardi, B., Neu, J., Pendlebury, D., Remsberg, E., Rozanov, A., Smith, L., Toohey, M., Urban, J., von Clarmann, T., Walker, K. A., and Wang, R. H. J.: SPARC Data Initiative: A comparison of ozone climatologies from international satellite limb sounders, J. Geophys. Res., 118, 12,229–12,247, doi:10.1002/2013JD019877, 2013.

Tegtmeier, S., Hegglin, M. I., Anderson, J., Funke, B., Gille, J., Jones, A., Smith, L., von Clarmann, T., and Walker, K. A.: The SPARC Data Initiative: comparisons of CFC-11, CFC-12, HF and SF_6 climatologies from international satellite limb sounders, Earth Sys. Sci. Data, 8, 61–78, doi:10.5194/essd-8-61-2016, 2016.

Toohey, M. and von Clarmann, T.: Climatologies from satellite measurements: the impact of orbital sampling on the standard error of the mean, Atmos. Meas. Tech., 6, 937–948, doi:10.5194/amt-6-937-2013, 2013.

Toohey, M., Hegglin, M. I., Tegtmeier, S., Anderson, J., Añel, J. A., Bourassa, A., Brohede, S., Degenstein, D., Froidevaux, L., Fuller, R., Funke, B., Gille, J., Jones, A., Kasai, Y., Krüger, K., Kyrölä, E., Neu, J. L., Rozanov, A., Smith, L., Urban, J., von Clarmann, T., Walker, K. A., and Wang, R. H. J.: Characterizing sampling biases in the trace gas climatologies of the SPARC Data Initiative, J. Geophys. Res., 118, 11,847–11,862, doi:10.1002/jgrd.50874, 2013.

Toon, G. C., Blavier, J. F., Sen, B., Margitan, J. J., Webster, C. R., May, R. D., Fahey, D., Gao, R., Del Negro, L., Proffitt, M., Elkins, J., Romashkin, P. A., Hurst, D. F., Oltmans, S., Atlas, E., Schauffler, S., Flocke, F., Bui, T. P., Stimpfle, R. M., Bonne, G. P., Voss, P. B., and Cohen, R. C.: Comparison of MkIV balloon and ER-2 aircraft measurements of atmospheric trace gases, J. Geophys. Res., 104, 26 779–26 790, doi:10.1029/1999JD900379, 1999.

Velazco, V. A., Toon, G. C., Blavier, J.-F. L., Kleinböhl, A., Manney, G. L., Daffer, W. H., Bernath, P. F., Walker, K. A., and Boone, C.: Validation of the Atmospheric Chemistry Experiment by noncoincident MkIV balloon profiles, J. Geophys. Res., 116, D06 306, doi:10.1029/2010JD014928, 2011.

Vyushin, D. I., Fioletov, V. E., and Shepherd, T. G.: Impact of long-range correlations on trend detection in total ozone, J. Geophys. Res., 112, doi:10.1029/2006JD008168, 2007.

Waugh, D. and Hall, T.: Age of Stratospheric Air: Theory, Observations, and Models, Rev. Geophys., 40, doi:10.1029/2000RG000101, 2002.

Waymark, C., Walker, K. A., Boone, C. D., and Bernath, P. F.: ACE-FTS version 3.0 data set: validation and data processing update, Ann. Geophys., 56, Fast Track–1, doi:10.4401/ag-6339, 2013.

Yulaeva, E., Holton, J. R., and Wallace, J. M.: On the Cause of the Annual Cycle in Tropical Lower-Stratospheric Temperatures, J. Atmos. Sci., 51, 169–174, doi:10.1175/1520-0469(1994)051<0169:OTCOTA>2.0.CO;2, 1994.

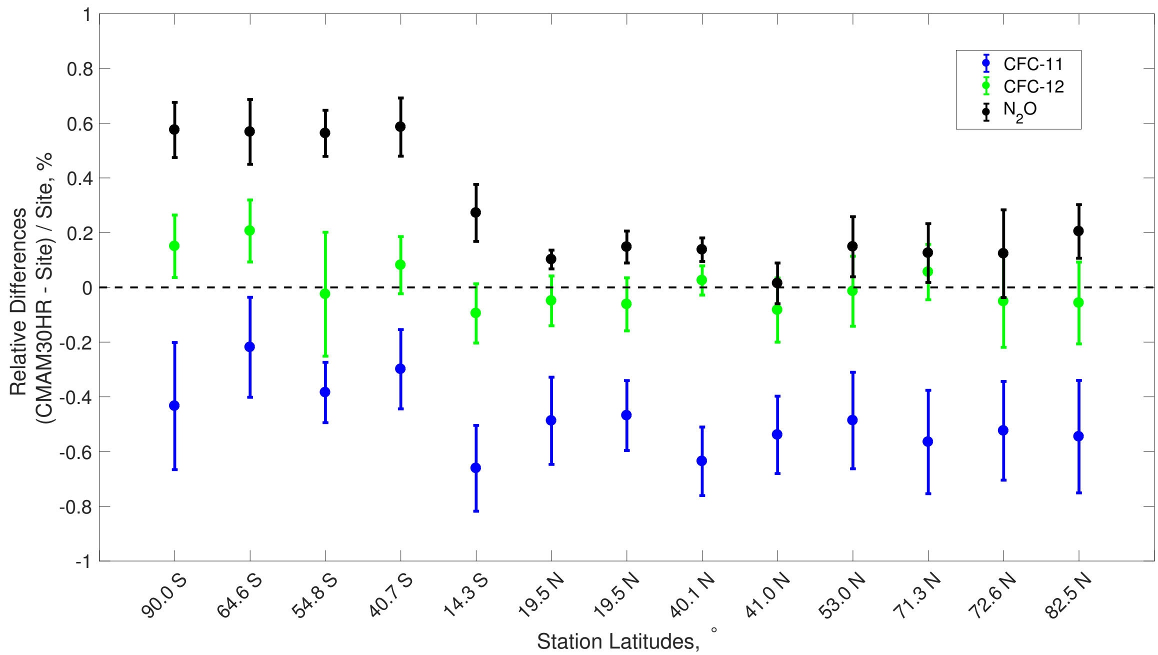


Figure 1. Comparison of CMAM30HR ~~run~~ simulations of CFC-11 (blue x), CFC-12 (green diamond), and N_2O (black circle) to the HATS surface flask network of measurements ~~, relative~~ at various locations around the world. Locations of measurement sites are indicated by latitude. Relative differences are calculated as the difference between the concentration at the surface site ~~subtracted from~~ and the lowest model layer of the nearest neighbor gridbox to the site in the CMAM30HR ~~simulation~~output, divided by the ~~average of the two, as described in the text~~measured concentrations. The relative differences ~~and~~ were calculated based on the ~~uncertainties included~~ monthly averaged observations and simulations. Shown here are the mean of the relative differences between May 2004 and June 2010 and the error bars indicate one standard deviation of the mean of the relative differences over the time ~~series~~period.

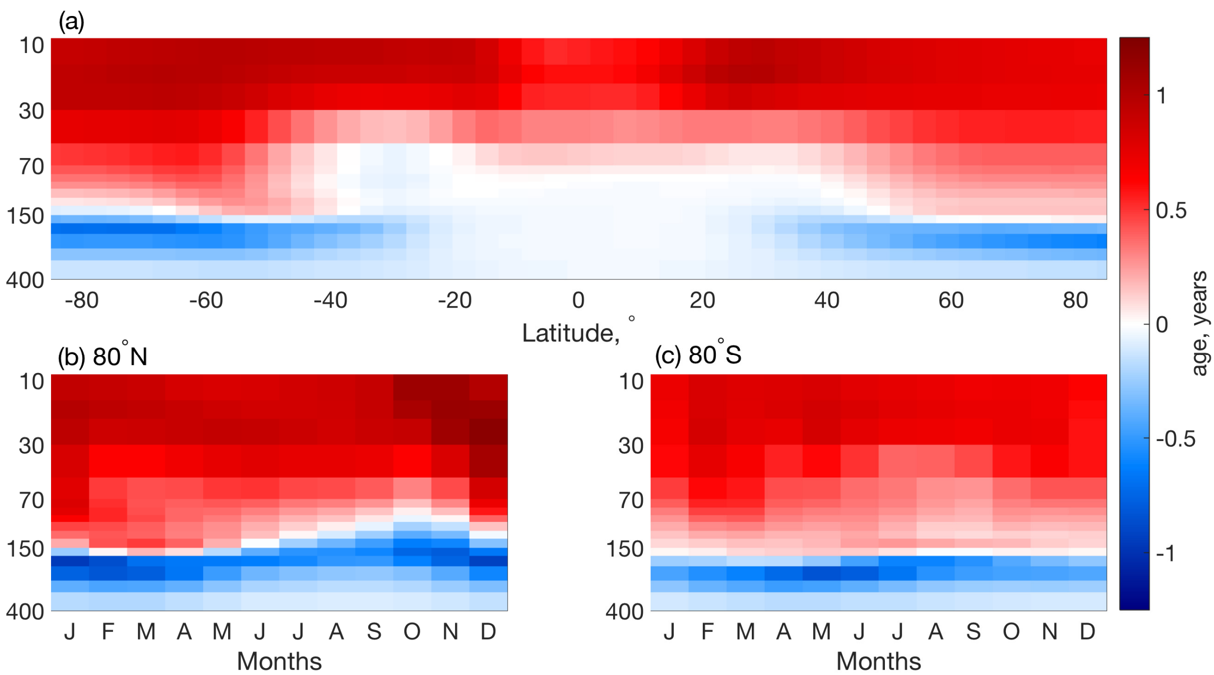


Figure 2. A comparison of the age of stratospheric air in the free-running CMAM and ~~the nudged CMAM,~~ CMAM30 ~~,~~ averaged between 2004 and 2010. (a) The zonal mean difference of the CMAM30 mean age subtracted from the free running mean age (years), and the monthly time series difference (b) at ~~80° S~~80° S, and (c) at ~~80° N~~80° N.

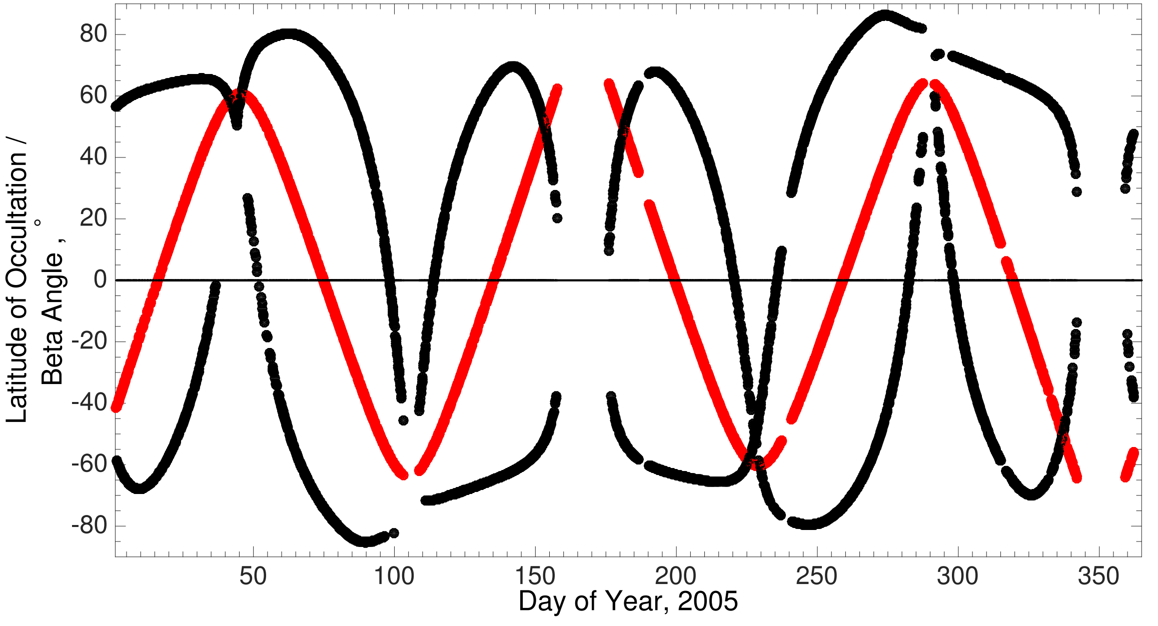


Figure 3. The ACE-FTS sampling pattern for the year 2005. Each black circle is the latitude of the ~~30 km~~ 30 km tangent height of an occultation and each red circle is the corresponding beta angle of the occultation.

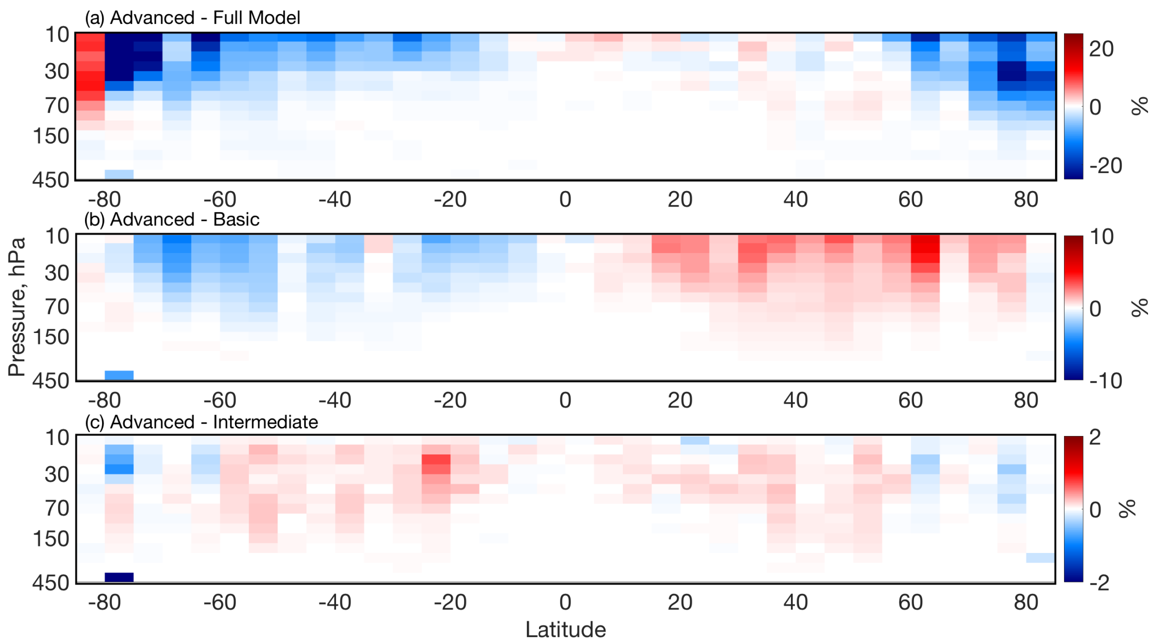


Figure 4. The comparisons of the three sampling methods described in the text using N_2O simulations in CMAM30HR. The relative differences (%) defined as the difference between the advanced sampling and (a) the full model output, (b) basic sampling, (c) intermediate sampling, divided by the advanced sampling. Note the different color scales in each panel.

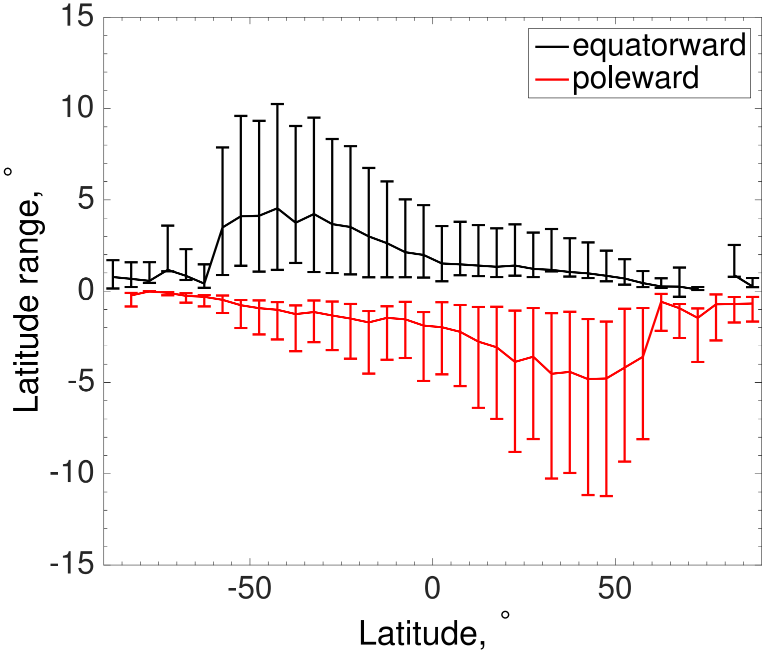


Figure 5. Average latitude ranges covered by ACE-FTS occultations in a given 5° latitude bin, separated by an equatorward bias (black) and a poleward bias (red) as defined in the text. The error bars indicate one standard deviation from the mean latitudinal extent for the given 5° latitude bin.

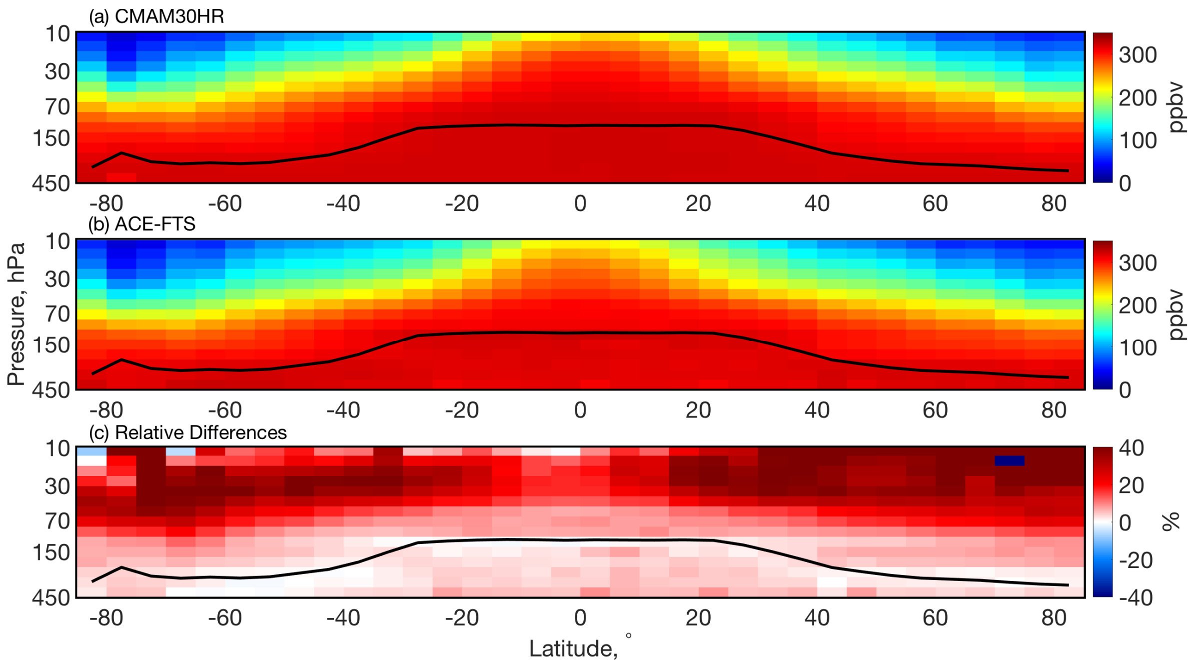


Figure 6. ~~Zonally averaged~~ Zonally-averaged annual-mean latitude-altitude distributions of N_2O; (a) CMAM30HR (ppbv), (b) ACE-FTS (ppbv), (c) the mean relative difference (in %) between sampled model and ACE-FTS profiles, divided by the ~~mean of their respective values (100*(CMAM30HR−~~ACE-FTS ~~/((CMAM30HR+ACE-FTS)/2)))~~profiles (100*(CMAM30HR−ACE-FTS)/ACE-FTS). The ~~blue~~ black line indicates the location of the thermally-defined tropopause.

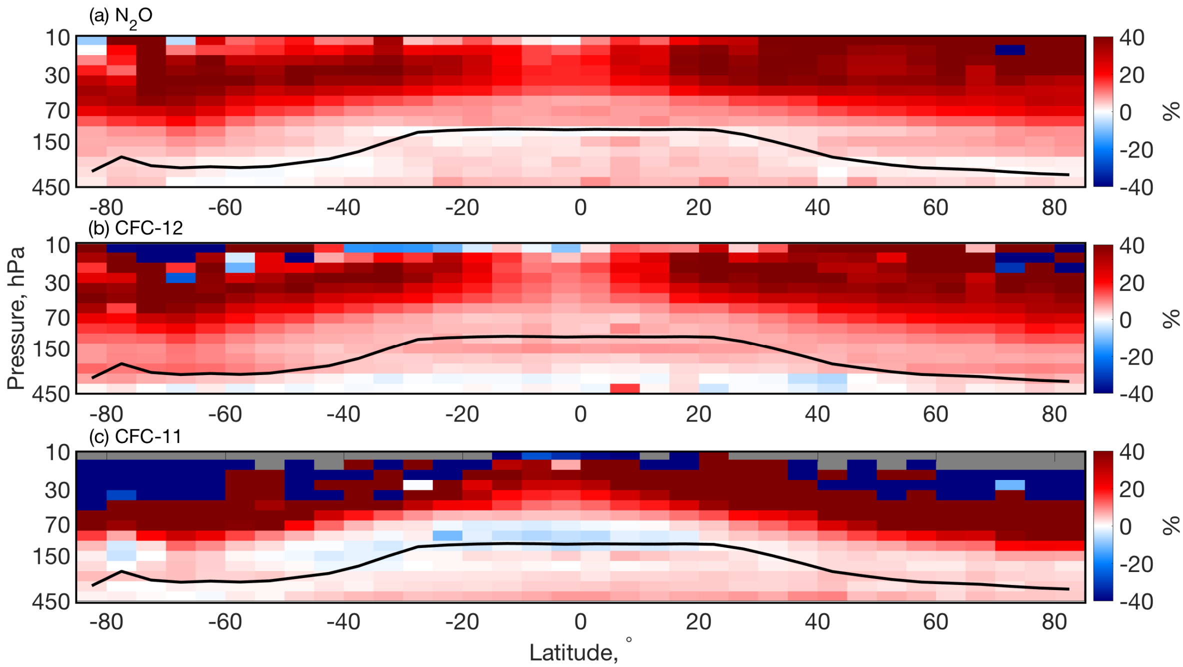


Figure 7. Same as Fig. 6c for (a) N_2O, (b) CFC-12, and (c) CFC-11. ~~Note the different scale used on panel (c)~~The grey regions indicate where no data is available.

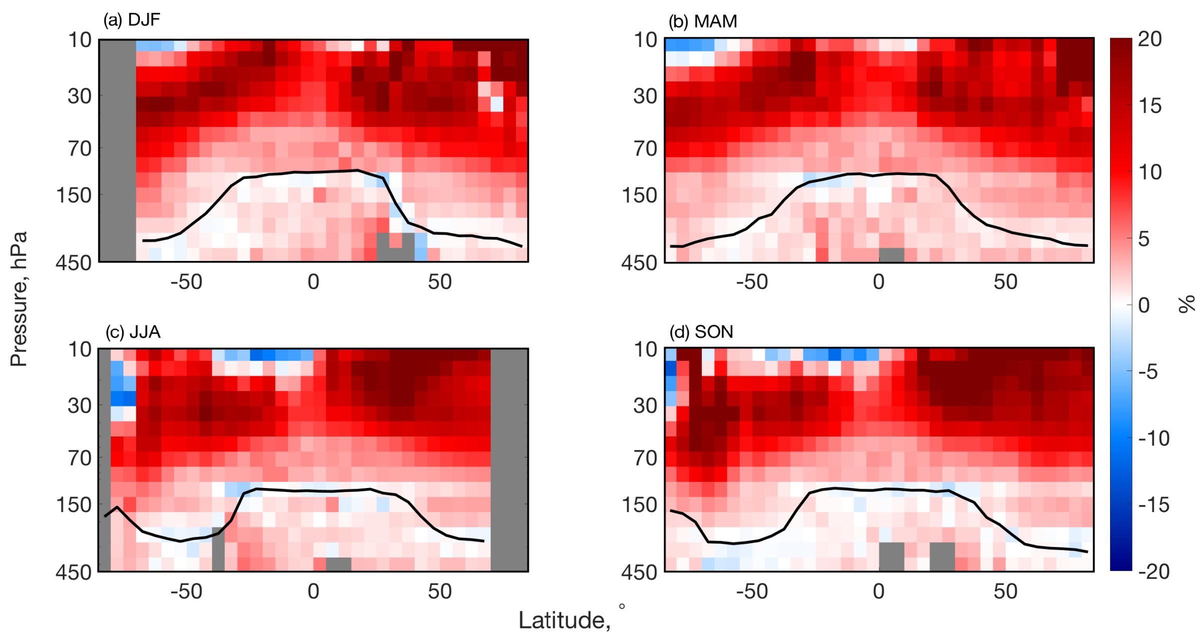


Figure 8. The relative mean of individual ACE-FTS profiles subtracted from CMAM30HR profiles of N_2O, divided by the ~~mean of~~ ACE-FTS ~~and CMAM30HR~~ measurements for each season (a) ~~December-January-February (DJF)~~December-January-February (DJF), (b) ~~March-April-May (MAM)~~March-April-May (MAM), (c) ~~June-July-August (JJA)~~June-July-August (JJA), and (d) ~~September-October-November (SON)~~September-October-November (SON). The grey regions indicate where no data is available.

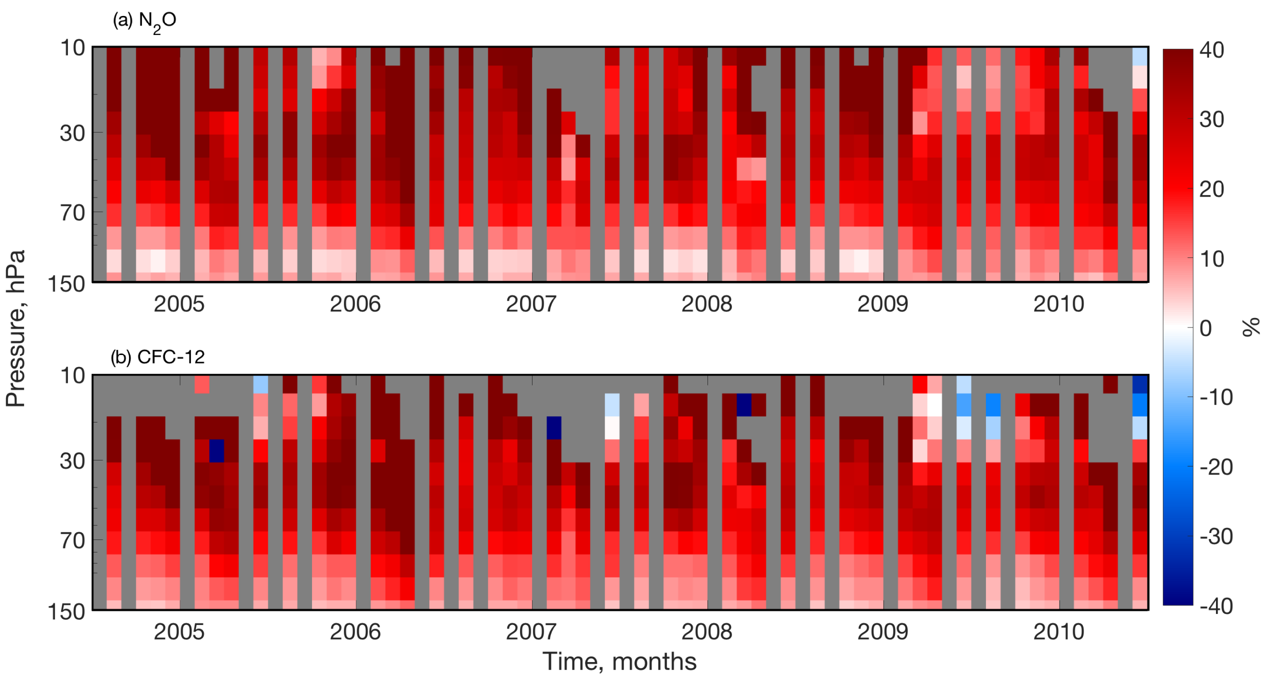


Figure 9. A monthly time series of the average relative differences between the model and the measurements (CMAM30HR minus ACE-FTS, divided by ~~the mean~~ACE-FTS), as in Fig. 6c for (a) N_2O and (b) CFC-12 in the Northern Hemisphere ~~(60° - 90° N)~~ (60° - 90° N) between June 2004 and May 2010. The grey regions indicate where no data is available.

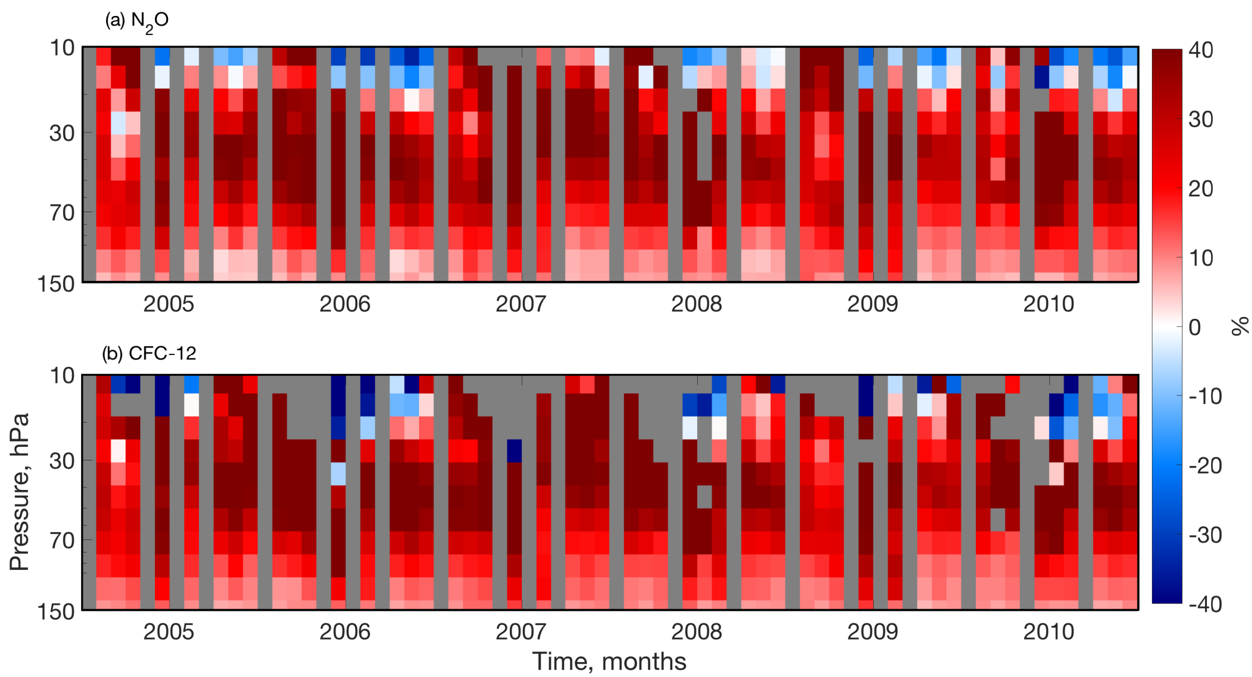


Figure 10. A monthly time series of the average relative differences between measurements and the model (CMAM30HR minus ACE-FTS, divided by ~~the mean~~ACE-FTS), as in Fig. 6c for (a) N_2O and (b) CFC-12 in the Southern Hemisphere ~~(60° S - 90° S)~~ (60° S - 90° S) between June 2004 and May 2010. The grey regions indicate where no data is available.

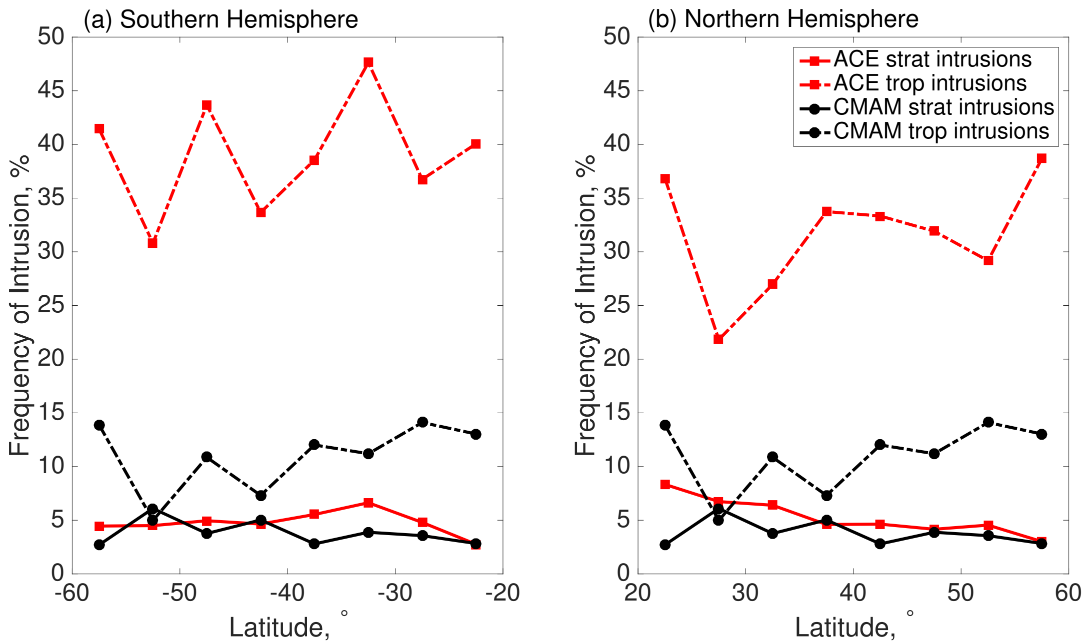


Figure 11. The frequency of ~~intrusions~~ intrusion across the tropopause for both CMAM30HR (black circles) and ACE-FTS (red squares) below the 420 K potential temperature isotherm in (a) the Southern Hemisphere and (b) the Northern Hemisphere extratropical region between ~~20° - 60° N/S~~20° - 60° N/S, respectively. ~~The~~ For each 5° latitude bin, the frequency of intrusion ~~frequencies~~ events are separated by stratospheric intrusions (solid lines), where stratospheric-like air is found in the troposphere, and tropospheric intrusions (dashed lines),~~as defined~~ where tropospheric-like air is found in the ~~text, for each 5° latitude bin~~stratosphere.

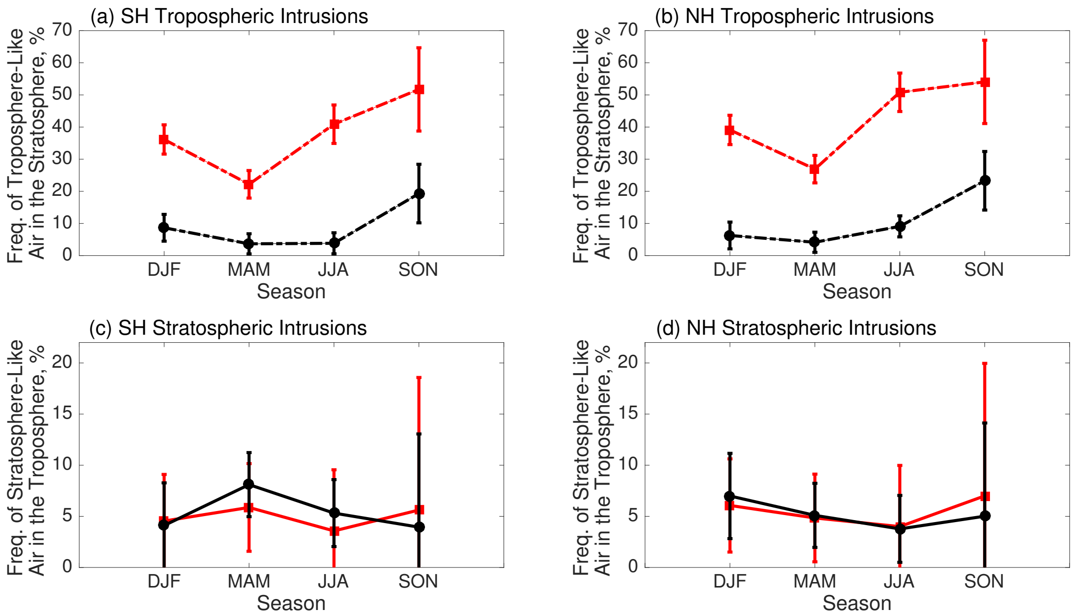


Figure 12. A seasonal representation of the intrusions depicted in Figure 11: (a) tropospheric intrusions in the Southern Hemisphere, (b) tropospheric intrusions in the Northern Hemisphere, (c) stratospheric intrusions in the Southern Hemisphere, and (d) stratospheric intrusions in the Northern Hemisphere. Each season is an average of the extratropical intrusion frequency with error bars indicating one standard deviation of the seasonal mean. ACE-FTS intrusions are in red squares and CMAM30HR intrusions are in black circles. Note the seasonality represented may be impacted by the sampling of ACE-FTS and may not representative for the full atmosphere

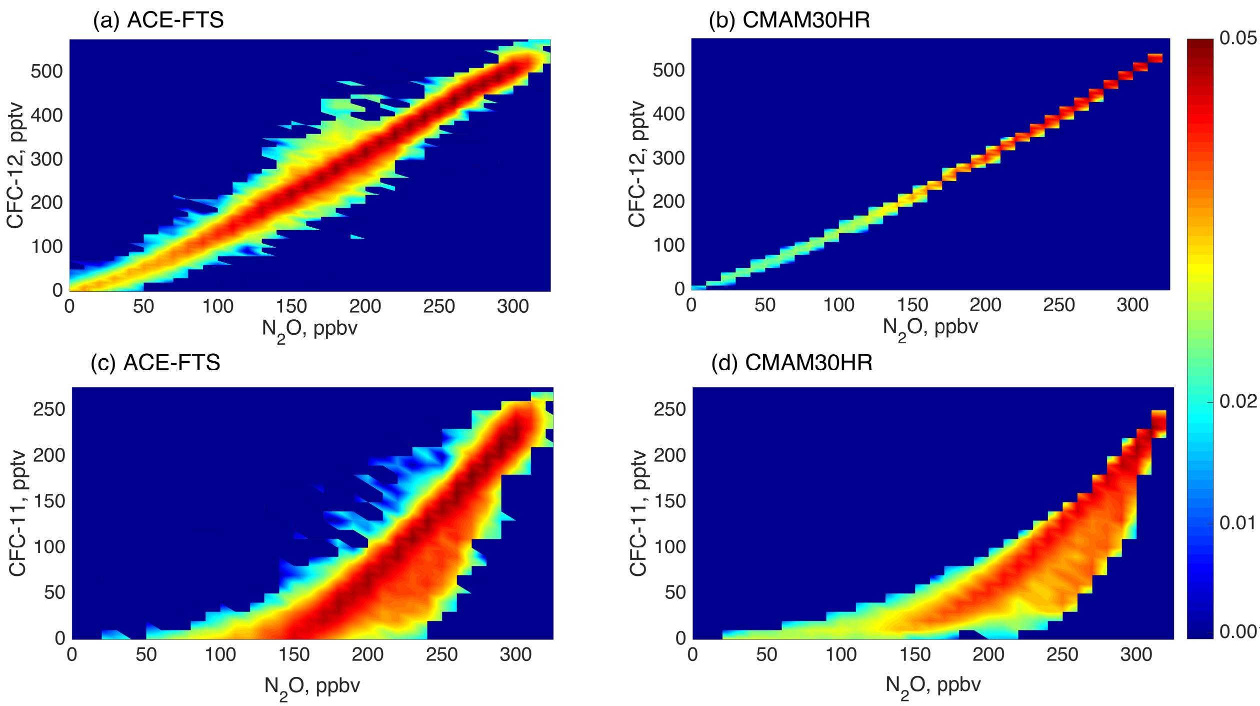


Figure 13. JPDFs, as described in the text, of N_2O/CFC-12 for (a) ACE-FTS and (b) CMAM30HR, and N_2O/CFC-11 for (c) ACE-FTS and (d) CMAM30HR. All stratospheric ACE-FTS observations and subsampled model output in the Northern Hemisphere are included.

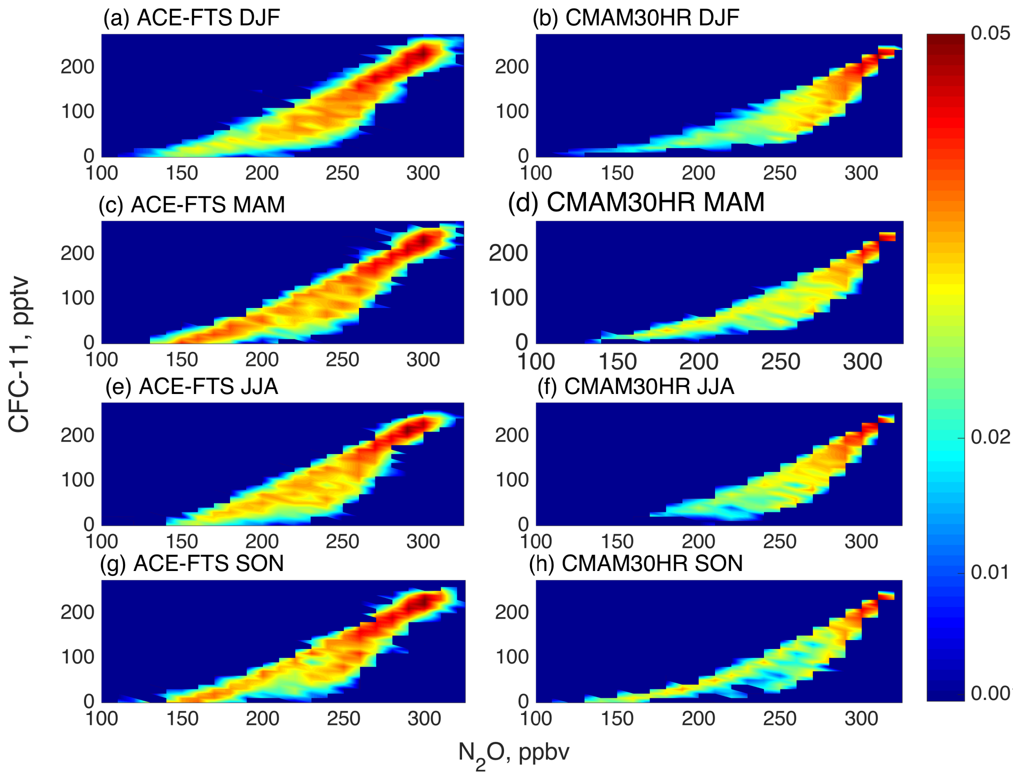


Figure 14. Tropical JPDFs of N_2O/CFC-11, as described in the text, separated by season. Only stratospheric ACE-FTS observations (left column) and CMAM30HR simulations (right column) within the height-dependent tropical turn-around-latitudes have been included.

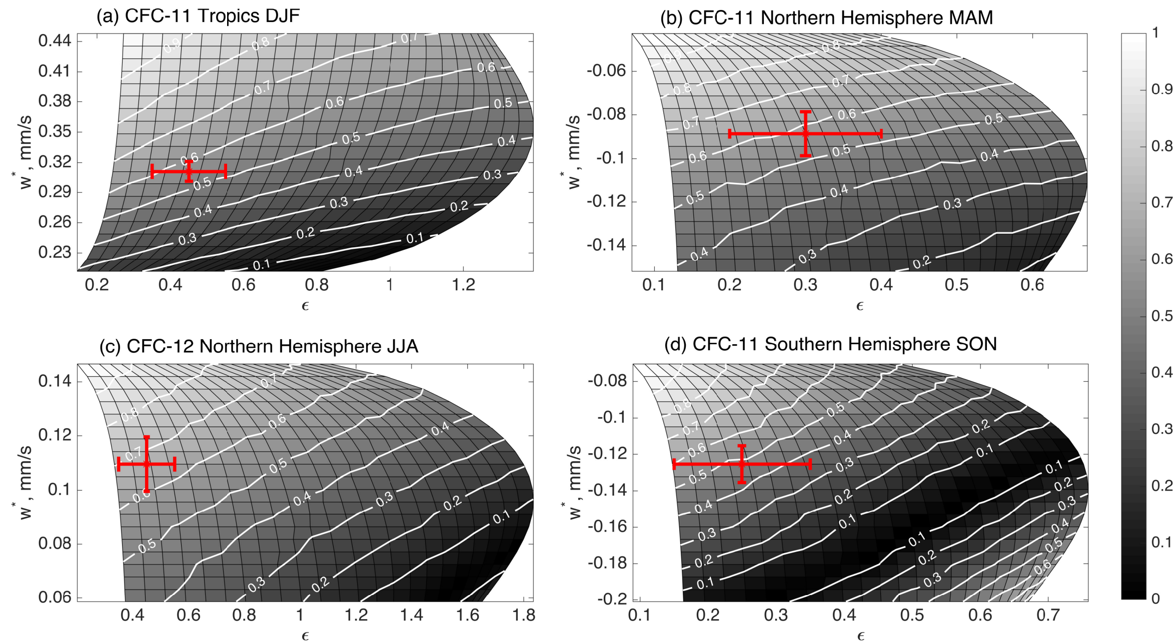


Figure 15. Examples of the TLP comparisons described in the text. The grey scale on each panel indicates the comparison between ACE-FTS and the TLP simulations over various $w^{\star}$ (y-axes) and ϵ cases (x-axes). The white lines identify the agreement with ACE-FTS (gray colour scale). The red marker indicates the CMAM30HR estimated values of $w^{\star}$ and ϵ. The examples shown are (a) CFC-11 tropics during the DJF season, (b) CFC-11 Northern Hemisphere during the MAM season, (c) CFC-12 Northern Hemisphere during the JJA season, and (d) CFC-11 Southern Hemisphere during the SON season.

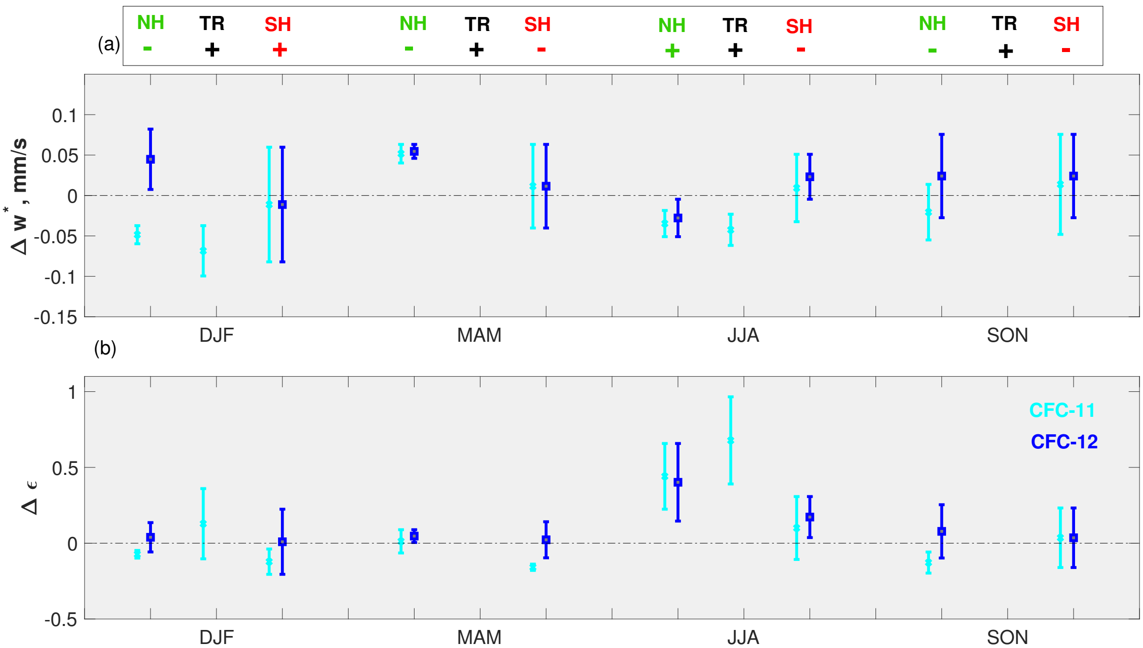


Figure 16. A summary of the TLP comparisons by season and region: (a) indicates the changes required in tropical upwelling, $w^{\star}$, while (b) indicates the changes in the mixing, ϵ, required for the CMAM30HR simulation to agree with the ACE-FTS observations. The '-' and '+' symbols at the top of the figure indicate the direction of mass transport (downwelling or upwelling, respectively) for each region and season as indicated by color and location region (green for Northern Hemisphere, black for tropics, and red for the Southern Hemisphere). The changes required based on CFC-11 are shown in cyan and the figurechanges required based on CFC-12 are shown in blue.