Using transport diagnostics to understand Chemistry Climate Model ozone simulations

S. E. Strahan[1], A. R. Douglass[2], R. S. Stolarski[2], H. Akiyoshi[3], S. Bekki[4], P. Braesicke[5], N. Butchart[6], M.P. Chipperfield[7], D. Cugnet[4], S. Dhomse[7], S.M. Frith[8], A. Gettelman[9], S.C. Hardiman[6], D.E. Kinnison[9], J.-F. Lamarque[9], E. Mancini[10], M. Marchand[4], M. Michou[11], O. Morgenstern[12], T. Nakamura[3], D. Olivié[13], S. Pawson[2], G. Pitari[10], D.A. Plummer[14], J.A. Pyle[5], E. Rozanov[15,16], J.F. Scinocca[14], T.G. Shepherd[17], K. Shibata[18], D. Smale[12], H. Teyssèdre[11], W. Tian[7], and Y. Yamashita[3,19]

[1]University of Maryland Baltimore County, Baltimore, Maryland, USA

[2]NASA Goddard Space Flight Center, Greenbelt, Maryland, USA

[3]National Institute of Environmental Studies, Tsukuba, Japan

[4]LATMOS, IPSL, UVSQ, UPMC, CNRS, INSU, Paris, France

[5]NCAS Climate-Chemistry, Centre for Atmospheric Science, Department of Chemistry, Cambridge University, Cambridge, UK

[6]Hadley Centre, Met Office, Exeter, UK

[7]School of Earth and Environment, University of Leeds, Leeds, UK

[8]Science Systems and Applications Inc., Lanham, Maryland, USA

[9]National Center for Atmospheric Research, Boulder, Colorado, USA

[10]Dipartimento di Fisica, Università degli Studi dell' Aquila, L'Aquila, Italy

[11]GAME/CNRM, Météo-France, CNRS, Toulouse, France

[12]National Institute of Water and Atmospheric Research, Lauder, New Zealand

[13]Department of Geosciences, University of Oslo, Oslo, Norway

[14]Canadian Centre for Climate Modeling and Analysis, Environment Canada, Victoria, British Columbia, Canada

[15]Physical-Meteorological Observatory/World Radiation Center, Davos, Switzerland

[16]Institüt für Atmosphäre und Klima, ETH Zurich, Zurich, Switzerland

[17]Department of Physics, University of Toronto, Toronto, Ontario, Canada

[18]Meteorological Research Institute, Japan Meteorological Agency, Tsukuba, Japan

[19]NOAA Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey, USA.

Abstract. We demonstrate how observations of N_2O and mean age in the tropical and midlatitude lower stratosphere (LS) can be used to identify realistic transport in models. The results are applied to 15 Chemistry Climate Models (CCMs) participating in the 2010 WMO assessment. Comparison of the observed and simulated N_2O/mean age relationship identifies models with fast or slow circulations and reveals details of model ascent and tropical isolation. The use of this process-oriented N_2O/mean age diagnostic identifies models with compensating transport deficiencies that produce fortuitous agreement with mean age.

We compare the diagnosed model transport behavior with a model's ability to produce realistic LS O_3 profiles in the tropics and midlatitudes. Models with the greatest tropical transport problems show the poorest agreement with observations. Models with the most realistic LS transport agree more closely with LS observations and each other. We incorporate the results of the chemistry evaluations in the SPARC CCMVal Report (2010) to explain the range of CCM predictions for the return-to-1980 dates for global (60°S-60°N) and Antarctic column ozone. Later (earlier) Antarctic return dates are generally correlated to higher (lower) vortex Cl_y levels in the LS, and vortex Cl_y is generally correlated with the model's circulation although model Cl_y chemistry or Cl_y conservation can have a significant effect. In both regions, models that have good LS transport produce a smaller range of predictions for the return-to-1980 ozone values. This study suggests that the current range of predicted return dates is unnecessarily large due to identifiable model transport deficiencies.

1. Introduction

Chemistry climate models (CCMs) are the current state-of-the-art tools used to assess stratospheric ozone and make predictions of its future evolution (World Meteorological Organization, in press; WMO, 2007). Ozone distributions are controlled by transport, chemistry, and temperature (i.e., dynamics and radiation). In the stratosphere, the processes that control ozone are expressed by the ozone tendency equation,

$$dO_3/dt = \text{Transport} + P - L(O_x) + L(NO_x) + L(Cl_x) + L(Br_x) + L(HO_x),$$

where P is O_3 production and the L-terms are loss processes due to different radical families. In the lower stratosphere O_3 chemistry is slow and distributions are controlled primarily by transport. In the middle and upper stratosphere, photochemistry is fast but transport still plays an important role because it controls the distributions of long-lived families that produce radicals involved in O_3 loss processes (Perliski et al., 1989; Douglass et al., 2004). Photochemistry and temperature control the steady state balance between radicals and their precursors, e.g., NO_x/NO_y and ClO_x/Cl_y. Transport and chemistry in a model must both be physically realistic to produce a credible simulation of O_3.

The use of O_3 as a measure of realism in a simulation is fraught with problems. In some cases, compensating deficiencies in the processes affecting O_3 produce a realistic result. In other cases, an O_3 profile or column may be insensitive to some of the terms in the tendency equation. For example, Douglass et al. (1997) noted that while their model's O_3 profiles agreed with observations from several UARS instruments, their simulated long-lived tracer profiles did not. The tracer profiles were poor due to the dependence on horizontal transport, which was excessive in the model. The good agreement of the simulated ozone with observations indicated that the sum of O_3 loss processes was reasonable, but as the long-lived tracers were too high, the relative fractional losses from different cycles (i.e., NO_x, HO_x, ClO_x, and O_x) were probably incorrect (Douglass et al., 2004). Comparisons with total column O_3 observations of the recent past have been used in previous ozone assessments as an indication of model performance (Stratospheric Processes and their Role in Climate (SPARC) Chemistry Climate Model Validation (CCMVal), 2010; WMO, 2007; WMO, 2003). Andersen et al. (2006) examined 10 2-D and 4 3-D O_3 simulations used in the 2007 WMO assessment and found no correlation between the past and future column ozone trends simulated by individual models. They noted that overestimates of O_3 trends in one latitude region compensated for underestimates in other regions giving a false agreement between the model 60°S-60°N column O_3 trend and the trend from observations. The use of column O_3

as a measure of model performance goes back at least 20 years (e.g., Jackman et al., 1991) and is still used today (e.g., Austin et al., 2010).

Douglass et al. (1999) proposed setting standards for model evaluation that were based on objective comparisons with observations, including quantitative scoring. This approach became feasible in the 1990's with the availability of multi-year, near-global stratospheric trace gas data sets from satellite-based instruments such as those on the Upper Atmosphere Research Satellite (UARS). This approach is physically based and therefore may identify areas where model improvement is needed. Objective, observationally-based evaluations are also advantageous because they provide a way to quantitatively reassess a model after improvements have been made. Today, more than 10 years since Douglass et al. (1999), many satellite data sets are available for the development of model diagnostics. Balloon and aircraft data sets and ground-based networks also provide valuable observations for the development of model evaluation tools.

Recently, Eyring et al. (2006) and Waugh and Eyring (2008) adopted a physically-based approach for evaluating CCM simulations by assessing the representation of processes that affect stratospheric ozone. They posit that diagnosing transport and dynamical processes in CCMs, rather than O_3, is a meaningful way to evaluate a model's ability to make reliable projections of future composition. This approach was applied to the CCMs in the 2007 WMO assessment (WMO, 2006) to help interpret model predictions of future O_3 levels. The Stratospheric Processes and their Role in Climate (SPARC) Chemistry Climate Model Validation (CCMVal) Report (2010, hereinafter referred to as SCR)) builds on this approach and represents the most comprehensive effort to-date toward evaluating model processes by developing observationally based diagnostics for radiation, dynamics, transport, and chemistry. The 18 CCMs participating in the most recent WMO assessment (WMO, in press) were evaluated in this report and are listed in Table 1. Details of the chemistry, transport, dynamics, and radiation representations in these models, along with details of the reference simulations performed, can be found in Morgenstern et al. (2010).

The transport evaluation in the SCR concluded that tropical ascent and subtropical (horizontal) mixing were two fundamental processes that strongly affect the distributions of ozone and ozone-depleting substances. By mixing we mean irreversible transport of extratropical air into the tropics and not the export of tropical air to the midlatitudes, which is required by mass conservation because pressure (mass) decreases with altitude. The SCR concluded that at least half of the participating CCMs had significant issues with lower stratosphere (LS) transport, that is, circulation, mixing, or both. A useful

transport metric developed in the SCR was the average mean age (AMA). The AMA grade is a measure of how well a model agrees with mean age observations at 7 different altitudes and latitudes. Mean age depends on both circulation (i.e., tropical ascent rate) and mixing into the tropics, and on the balance between them as a function of altitude. Requiring a model to have realistic mean age at many different altitudes and latitudes was shown to be the most stringent transport diagnostic yet developed. Unfortunately, mean age at pressures other than 50 hPa (~20 km) has only been measured at a few latitudes, so the mean age data set necessary to fully constrain model LS transport does not yet exist.

In the lower stratosphere, mean age is generally less than 4 yrs and is strongly correlated with N_2O. The correlation between simulated N_2O and mean age was described in Hall et al. (1999) in an evaluation of more than a dozen models. Although there was a wide variation in model mean ages, most differing greatly from observations, they found that the models' N_2O/mean age relationships were qualitatively similar. Mean age and N_2O are correlated because the longer air remains in the stratosphere, the older it gets and the more N_2O is photochemically destroyed. There is a limit to this relationship because as transport time in the stratosphere increases, the maximum altitude attained increases and all N_2O is eventually destroyed, ending the correlation. For most of the lower stratosphere, N_2O is strongly correlated with mean age because its mixing ratios reflect a mixture of young (high N_2O) air that has not travelled to the loss region (~5 hPa and above) and older (low N_2O) air that has experienced photochemical loss.

In this paper, we show how global observations of N_2O can be combined with existing mean age observations from 150-30 hPa to extend the usefulness of the AMA concept by using N_2O as a proxy for age in the LS. The N_2O/mean age analysis is applied to 15 CCMs and is interpreted in terms of the models' representation of tropical ascent and isolation. Comparison of these results with O_3 profile data investigates the physical link between this transport diagnosis and model O_3 simulations in the tropics and midlatitudes below 30 hPa. Since much of the ozone column resides in the lower stratosphere, realistic representation of lower stratospheric transport is an essential component of simulating realistic total column ozone. We examine predictions of the return-to-1980 ozone columns for the 15 CCMs that simulated the 21st century. Using the transport diagnostics presented here along with some of the CCMVal chemistry evaluations, we can explain at least half of the range of model-predicted return dates for global (60°S-60°N) and October Antarctic ozone columns. This study suggests that the current range of predicted return dates is unnecessarily large due to identifiable modeling deficiencies. While there are significant uncertainties in the return dates due to the unknown levels of future ODS and GHG

emissions, the use of transport and chemistry diagnostics to identify models with credible LS transport and photochemistry may reduce the uncertainty in return dates caused by unrealistic representation of important processes controlling stratospheric ozone.

2. N_2O and Mean Age Observations used in the Transport Diagnosis

In this section we review the observations of N_2O and mean age that are used to develop empirical constraints on lower stratospheric transport. Satellite observations of N_2O from the Atmospheric Chemistry Experiment (ACE) satellite instrument onboard SCISAT-1 (2004-2008) (Jones et al., in preparation) and from the Microwave Limb Sounder (MLS) satellite instrument onboard Aura (2004-2009) (Livesey et al., 2007; Lambert et al., 2007) are used to create an annual mean climatology for each 5-year data set. The use of a 5-year mean reduces the effect of the quasi-biennial oscillation on LS N_2O distributions. The MLS v2.2 N_2O observations have daily 82°S-82°N coverage in the stratosphere extending to 100 hPa. The ACE v2.2 N_2O has latitudinal coverage that varies during the year; as a result the annual mean climatology is restricted to 68°S-68°N. Levels of 150 hPa and above are used here. We also use mean age derived from CO_2 observations from balloon (Engel et al., 2009; Andrews et al., ACE 2001) and aircraft (Andrews et al., 2001) campaigns. Figure 1 shows zonal annual mean MLS N_2O, ACE N_2O, and their percentage difference. Where these data sets have spatial overlap, 68°S-68°N, and at the pressure levels of interest in this study, 100-20 hPa, ACE and MLS differences are almost always less than 10%; for most of the LS they are less than 5%. In the ACE N_2O validation study, Strong et al. (2008) report that the mean profile differences between ACE and MLS is ±5% from 1-100 hPa, with MLS showing low bias at pressures greater than 32 hPa.

Figure 2 displays the mean age data used in this analysis. The tropical profiles come from three OMS balloon CO_2 profiles measured in February and November, 1997 (Andrews et al., 2001) and from two balloon CO_2 profiles measured in June, 2005 (A. Engel, pers. comm.). The average profile from these two data sets (red) and its 1σ uncertainty are used in the analysis. The mean midlatitude mean age and uncertainty profiles were derived from CO_2 and SF_6 data from 27 balloon flights, 32°N-51°N, spanning 30 years (Engel et al., 2009). The bottom panel of Figure 2 shows the 50 hPa (~20 km) mean age as a function of latitude. It was derived from numerous aircraft flights in different months but does not represent an annual mean as well as the midlatitude profile does. Nevertheless, these data sets represent our best estimate of the annual-averaged mean age.

Although mean age observations do not have the spatial coverage of satellite N_2O measurements, they do span a broad latitude and altitude range. Figure 3 shows ACE annual mean N_2O from the same latitudes and altitudes as the mean age observations. Mean age of less than 4.5 years and N_2O greater than 150 ppb form a compact and nearly linear relationship. There is a slight deviation from linearity at mean age <0.5 yr but the relationship is still compact. Where mean age is >4.5 years, age and N_2O are uncorrelated. Hall et al. (1999) and Schoeberl et al. (2000) have explained the loss of correlation as a result of increasing N_2O loss (photochemical exposure) as air ages. This is analogous to the compact relationship observed in the LS for N_2O and Cl_y (Schauffler et al., 2003), which also arises due to photochemical aging of air. The uncorrelated region of Figure 3 comes only from measurements in the NH midlatitudes, 20-7 hPa, where mean age has a constant value of ~4.8 yrs. The large variation of N_2O while mean age is constant suggests that numerous stratospheric pathways with the same mean age but with different photochemical exposure histories bring parcels to this region. For mean ages ≤4.5 years the same linear relationship exists for both tropical (red) and extratropical (blue and green) points considered separately, suggesting that for the extrapolar lower stratosphere where N_2O >150 ppb (i.e., for regions where very old air does not make up a large fraction of the age spectrum), N_2O can be used as a proxy for mean age. This relationship makes N_2O an especially useful transport tracer because of the availability of global, multi-year data sets.

We have tested the fit of the N_2O/mean age relationship separately for ACE and MLS N_2O observations and found only very small differences in the fitted slope and offset. This is not surprising given the close agreement seen in Fig. 1 from 20-90 hPa. The use of ACE measurements is advantageous because it extends the analysis of the N_2O/mean age relationship into the lowermost stratosphere (150 hPa) and its mean absolute difference with other satellite N_2O measurements is ±10 ppb in the LS (Strong et al., 2008). The use of MLS observations has some disadvantages in that there are no retrievals below 100 hPa and the systematic uncertainty at 100 hPa is large, about 25% (Livesey et al., 2007). In addition, MLS N_2O has an unrealistic inverted profile in the tropical LS (a local maxima at 70 hPa) that is not seen by ACE or in aircraft data. This causes the scatterplot of MLS N_2O and mean age to deviate sharply from the fitted line below mean age of 0.6 yrs. ACE and MLS N_2O have the greatest disagreement at 100 hPa (~10 %), where MLS is low biased compared to ACE (Strong et al., 2008). For these reasons we use the ACE N_2O/mean age relationship for all analyses.

3. The physical basis for the mean age/N_2O relationship

The relationship between N_2O, the age spectrum, and mean age can be explored using a well-evaluated chemistry transport model with credible lower stratospheric transport (Strahan et al., 2007). The Global Modeling Initiative (GMI) chemistry and transport model (CTM) was integrated for 5 years using GEOS4-GCM meteorological fields ('G4GCM') with boundary conditions and sea surface temperatures (SSTs) for 1994-1998; this run is referred to as GMI-G4GCM. The G4GCM meteorological fields were produced by the GEOSCCM (Pawson et al., 2008), one of the models evaluated in the SCR. An age tracer experiment was also run with the GMI CTM by recycling the 5-year set of G4GCM meteorological fields to complete a 20-yr simulation. Lower stratospheric transport in this simulation was evaluated by comparisons with MLS N_2O and O_3 in Strahan et al. (2007). They showed that this CTM simulation was able to closely match means, variability, and annual cycles in both species in the LS. Figure 4 compares mean age from the GMI-G4GCM with mean age observations shown in Figure 2. Model mean ages are realistic in the LS at most latitudes but slightly young in the SH midlatitude (upper left panel) and at altitudes above 50 hPa in the NH midlatitudes (upper right panel). The tropical-midlatitude age gradient (lower right panel), which is a measure of tropical ascent rate (Neu and Plumb, 1999), is within the uncertainty range of the observations suggesting that the ascent rate is about right from 100-10 hPa. Tropical mean ages (lower left panel) above 50 hPa may be slightly low, which could result from too much tropical isolation or slightly fast ascent rates. Overall, the GMI-G4GCM mean age comparisons shown in Figure 4 suggest that this simulation has generally realistic circulation and mixing in the tropics.

Because of the demonstrated realism of LS transport in the GMI-G4GCM simulation, we use its age spectra to show how N_2O and mean age are physically linked. Figure 5 shows the model's age spectra from 68°S-68°N, 100-10 hPa. Age spectra in the tropics show a large pulse that propagates upwards, reflecting the recent arrival of tropospheric air (i.e., young age elements). Moving upwards in the tropics, the pulse's amplitude decreases and the size of the tail increases as older air mixes into the tropics during ascent. In the extratropics at 52 hPa and above, a larger percentage of the age elements are older than 2 yrs, indicating that air may not have arrived directly from the tropics but instead through the downward branch of the Brewer-Dobson circulation. In each panel, the area to the left of the vertical dashed line shows the youngest 1.8 yrs of the spectrum; this fraction of the spectrum was found to have a nearly linear relationship with mean age for levels 37 hPa and below. The mean age of each spectrum (black) and the percentage of age elements ≤1.8 years (red) are also shown in each panel.

Figure 6 shows that mean age and N_2O are related through the percentage of young air in a parcel in the GMI-G4GCM simulation. In each panel, the black points show the correlation over the entire model domain (surface to mesosphere, pole to pole) while the red points represent only the domain of the observations (150-30 hPa, 68°S-68°N). Figure 6a shows that if we consider only points in the observed domain (red), we see a line that closely resembles the correlated part of the observed N_2O/mean age relationship shown in Figure 3. Figure 6b shows the relationship between model mean age and the percentage of young age elements. The monotonic relationship found in the domain of the observations (red) indicates that LS mean age is directly proportional to the contribution from young air. The link between young air and N_2O is demonstrated in Figure 6c. While the model NH and SH have slightly offset red curves, both curves show a compact and nearly 1:1 relationship between N_2O and young air in the domain of the observations. Above 30 hPa in the extratropics, where N_2O <100 ppb, descent of older air that has experienced photochemical loss contributes significantly to the age spectrum, breaking down the mean age/N_2O relationship. Age spectra from this region (e.g., the top row of Fig. 5) have a long tail, suggesting sufficient transport time for air to have passed through photochemical loss regions.

In the model LS, N_2O shows a monotonic relationship with the percentage of young age elements. In these regions, air parcels contain at least 5-10% young air. These results explain why LS N_2O can be used as a proxy for mean age up to ages of ~4-4.5 yrs. When there is less than 5-10% young air in the spectrum, the tail has a controlling influence on the mean age, and because old air in the tail has little or no N_2O, the correlation is lost.

4. Interpretation of Mean Age and N_2O in Chemistry Climate Models

We now examine N_2O and mean age in the CCMs that participated in the 2010 WMO ozone assessment in order to assess their LS transport characteristics. Table 1 lists the CCMs analyzed here; details of the participating models and additional references can be found in the SCR and in Morgenstern et al. (2010). The analysis presented in this section uses zonal monthly mean output of N_2O and mean age from the last decade of the REF-B1 simulation of the recent past (~1960-2006). Three of the participating CCMS did not provide mean age output and thus cannot be evaluated here (CCSRNIES, E39CA, and EMAC).

Figure 7 compares the observed global N_2O/mean age relationship with 15 CCMs for the LS (150-30 hPa, 68S-68N); this is the same data and uncertainty shown in Figure 3. Poleward of 60°S, model output is restricted to 50-150 hPa. The interpretation is fairly simple when the model points (red) fall on a curve

that is flatter or steeper than the observations (black). For example, most of the top two rows of Figure 7 (AMTRAC3, CNRM-ACM, Niwa-SOCOL, SOCOL, and UMETRAC) show maximum mean age less than observed while showing a minimum N_2O lower than observed. This relationship can only be achieved by a circulation that is too fast. The fast circulation quickly transports N_2O upward in the stratosphere where it is photochemically destroyed, resulting in a low mean age for a given value of N_2O. The bottom row (LMDZ and the UMUKCA models) shows model curves that are steeper than observed, have maximum mean age that is 1-3 years older than observed, and N_2O that is lower than observed. High mean age may be the result of a slow circulation, but may also be caused by too much recirculation (mixing) between the tropics and midlatitudes.

While the observations show a nearly linear relationship throughout this LS domain (Fig. 3), not all models do. AMTRAC3, CNRM-ACM, and LMDZ show a large change in slope, suggesting that the N_2O/mean age relationship is breaking down. The observations show that this occurs when mean age is greater than 4.5 yrs and N_2O is less than 150 ppb (Fig. 3) - ~30 hPa and above in the midlatitudes. The loss of correlation in the models suggests that parcels found near the top of this domain (30 hPa) have already spent significant time at high altitudes where N_2O is destroyed. Alternatively, low N_2O could be caused by a photochemical error.

UMSLIMCAT and the models in rows 3 and 4 have curves that agree better with the observed slope, however, they do not all show the same endpoint as the observations (i.e., the same maximum age and minimum N_2O). Further insight into these models requires a look at tropical transport behavior.

Figure 8 assesses ascent rate and horizontal mixing by examining the N_2O and mean age profiles in the tropics. The rightmost column shows the tropical subset of N_2O/mean age points shown in Figure 7. The first three columns of Figure 8 show three different profile comparisons with observations that diagnose the effects of 1) ascent and mixing together (tropical mean age), 2) tropical ascent rate (horizontal mean age gradient), and 3) horizontal mixing (tropical N_2O). Tropical mean age (column 1) increases with height as a function of both the ascent rate and the horizontal mixing strength. The agreement of model and observed tropical mean age only shows that the combined effects of ascent and mixing produce a realistic mean age in the model. Columns 2 and 3 identify how ascent and horizontal mixing each contributes to the overall tropical transport seen in column 1. These comparisons diagnose whether a problem lies with circulation or mixing, or both.

The horizontal age gradient profile (column 2) is an empirical measure of ascent rate because tropical and midlatitude mean ages are affected equally by horizontal mixing across the subtropics (Neu and Plumb, 1999; SCR). When the horizontal age gradient is less (greater) than observed, the ascent rate is too fast (slow). The comparison of modeled and observed tropical N_2O profiles (column 3) is a measure of the cumulative degree of horizontal mixing into the tropics as air ascends. Below 30 hPa, tropical N_2O has essentially no loss so the decreasing mixing ratio is due only to horizontal mixing. Model N_2O that is higher (lower) than observed can be interpreted as too little (much) mixing of older air into the tropics. The tropical N_2O differences are normalized to zero at 100 hPa and the yellow shading indicates the range of uncertainty in the LS based on validation against several satellite instruments (Strong et al., 2008). If midlatitude N_2O is biased high or low in a model, this diagnostic will have to be considered in light of the bias. For example, if the observed-model N_2O profile were good but the model's midlatitude N_2O was known to be low-biased, this would imply that mixing was actually too weak.

All models in Figure 8a are too young in the tropics, particularly above 70 hPa. CNRM-ACM, Niwa-SOCOL, SOCOL, and AMTRAC3 have fast ascent rates. They all have about the right amount of horizontal mixing except SOCOL, which has borderline too much mixing above 50 hPa. CNRM-ACM has the youngest mean age and fastest ascent of any of the CCMs. UMETRAC has young tropical mean age yet has a good ascent rate and good horizontal mixing. Neu and Plumb (1999) note that the horizontal mean age gradient is a good measure of ascent rate only in the case of small vertical diffusion. Model vertical diffusion that is too strong may explain why mean age is uniformly young throughout the UMETRAC stratosphere. Its midlatitude N_2O is not biased and so should not affect the interpretation of the mixing diagnostic.

The models in Figure 8b show the best agreement overall with all tropical mean age and N_2O diagnostic quantities, indicating they have the best ascent rates and horizontal mixing of the CCMs. CMAM has slightly rapid ascent near 50 hPa, but otherwise shows very good tropical transport behavior. UMSLIMCAT has variable ascent rates, a bit slow below 50 hPa and a bit fast above, but the overall tropical transport is good. Both CMAM and UMSLIMCAT are slightly young at 30 hPa. The CAM3.5 age gradient indicates fast ascent, but tropical mean age and mixing appear correct. This is explained by the high-biased CAM3.5 midlatitude N_2O. While the mixing diagnostic (column 3) suggests mixing is about right, it must actually be too strong because its midlatitude N_2O is too high, therefore more of it must mix in to sufficiently lower the tropical mixing ratio. In this case, tropical mean age comes out right because the fast ascent (which decreases age) is compensated by excessive mixing (which increases

age). This is consistent with the WACCM results. WACCM has very good agreement with observed midlatitude N_2O, and WACCM and CAM3.5 use the same atmospheric GCM. WACCM shows borderline fast ascent and too much mixing at 50 hPa and above – these effects compensate to give WACCM a good tropical mean age profile very much like CAM3.5. GEOSCCM is a little slow in ascent below 50 hPa but otherwise agrees very closely with the diagnostics.

The models in Figure 8c show a mixture of tropical transport problems. MRI produces a good tropical mean age profile at all levels but the ascent and mixing diagnostics indicate that this is fortuitous. MRI ascent and mixing look good from 100-50 hPa, but above 50 hPa mixing is too strong while ascent is too fast. The results for ULAQ are similar, but ULAQ has larger variations in ascent rate: much slower 100-70 hPa and much faster above 50 hPa. Like MRI, the fast ascent combined with strong mixing results in a good tropical mean age profile due to compensating effects from two physically unrealistic processes. LMDZ has very slow ascent from 100-50 hPa. Above 50 hPa, the ascent rate appears correct and mixing is good at all levels. The UMUKCA models show very slow ascent at all levels, allowing more time for horizontal mixing as the air ascends. However, the simple interpretation of the mixing diagnostic is incorrect due to the very low-biased midlatitude N_2O found in both models. When low N_2O mixes into the tropics, less mixing is required to produce the 'correct' tropical N_2O. We can only conclude that slow ascent leads to much older than observed age. The results presented here are consistent with the transport conclusions drawn in the SCR.

5. Discussion

5.1 Transport Effects on O_3 in the Lower Stratosphere

In this section we examine models' LS O_3 profiles to see if their agreement with observations is related to their transport characteristics. We include all models that submitted simulations for the recent past ('REF-B1') and future ('REF-B2') scenarios. At all latitudes above 30 hPa and in the polar LS, O_3 is influenced by processes in addition to transport, e.g., gas and condensed phase chemistry and temperature. Thus, the best place to look for a link between ozone and transport is in the extrapolar LS. Restricting the comparison to pressures of 150-30 hPa reduces but does not eliminate the influence of photochemistry. Also, model representation of polar loss processes or the polar vortex transport barrier may affect midlatitude profiles.

Figure 9 shows the difference between annual mean O_3 profiles from Aura MLS and 16 CCMs for the SH and NH midlatitudes and the tropics. The model annual means are a 10-yr average calculated from the

last 10 years of the REF-B1 run, usually 1997-2006. Although the CCSRNIES model did not submit the necessary mean age output for the transport diagnostic, the SCR transport evaluation characterized CCRSNIES as having slightly fast ascent rate and too much subtropical mixing. Annual zonal mean was calculated from 4 years of Aura MLS v2.2 O_3 (Livesey et al., 2007); precision error is negligible because of the large number of profiles in the average. The uncertainty of MLS O_3 as a function of pressure ranges from up to 30% in the lowermost stratosphere to 5-8% at 50 hPa and above. MLS systematic uncertainties are indicated by the shaded area about the zero line. The top panels show the differences in mixing ratio while the bottom panels show the differences as a percentage of observed O_3; the use of percentage makes it easier to examine O_3 differences where mixing ratios are small (below ~70 hPa). The top panels extend to higher levels than the bottom panels to illustrate some of the large O_3 differences found in the photochemically active region. No model agrees with MLS within the uncertainties at all locations shown. The five models with the best representation of circulation and mixing, identified by their tropical N_2O/mean age relationship in Fig. 8b, are plotted in red. Overall, these models show closer agreement with MLS O_3 than most models, and spread among them is small. The five models shown in blue or green have the greatest LS transport problems, either very fast ascent with weak mixing (CNRM-ACM), very slow ascent (UMUKCA-METO and UMUKCA-UCAM), or combined ascent and mixing problems (MRI and ULAQ). Their performances are consistent with their LS transport diagnoses. Below the O_3 maximum, slow ascent and excessive mixing each act to increase mean age and hence O_3: the UMUKCA models, MRI, and ULAQ all have higher than observed O_3. (For ULAQ this is only true in the tropics). Fast ascent produces younger mean age and lower ozone, and CNRM-ACM consistently has too low O_3.

The six models shown in black (AMTRAC3, CCSRNIES, LMDZrepro, Niwa-SOCOL, SOCOL, and UMETRAC) have identifiable transport deficiencies yet they compare as well with observations from 100-50 hPa as do the models with realistic transport. Clearly good agreement with observed LS O_3 by itself is not evidence of good transport. The top center panel shows that three of the six models shown in black have much higher than observed tropical O_3 near 30 hPa (CCRSNIES, Niwa-SOCOL, and SOCOL), suggesting a possible chemistry problem. Figure 9 demonstrates that the LS transport diagnostics are able to physically link poor LS O_3 profiles with transport behavior only in cases where the tropical transport is quite poor.

5.2 Impact of Transport on the Ozone Column

Roughly half of the ozone column resides in the LS, but to what degree can transport diagnostics explain the entire model column? Figure 10 shows the annual global mean (60°S-60°N) of the column O_3 anomaly (with respect to 1980) from the REF-B2 (future) simulation for 15 CCMs and observed annual mean column O_3 anomaly from the Merged Ozone Data Set (Stolarski and Frith, 2006). Model O_3 columns may be biased because most CCMs do not explicitly calculate tropospheric ozone; most comparisons in this paper will be made with column anomalies. Differences in CCMs approaches to tropospheric O_3, as well as differences between the CCMs in the implementation of the REF-B2 forcings, such as SST projections, will have some impact on model O_3 predictions; details of these differences can be found in Morgenstern et al. (2010). E39CA, EMAC, and UMETRAC did not a submit REF-B2 simulation and thus are not included in the remainder of the discussion.

Figure 10 shows several interesting results. First, there are two models with very similar transport diagnostics (AMTRAC3 and Niwa-SOCOL, Fig. 8a) that showed 'reasonable' agreement with present day O_3 profiles yet make widely different predictions of the return to 1980 ozone values. These models, shown in black, are the first and the last to cross the zero line (2004 and 2060). This large difference in return date in spite of their similar LS O_3 profiles and their similar transport diagnostics shows that these models must have a different balance of processes controlling O_3. That is, transport and chemistry (and implicitly temperature) have different-sized terms in their ozone tendency equations. Time-varying SSTs and source gas forcings in the future scenario cause changes to a model's transport circulation and chemistry. When a model responds to a perturbation, e.g., a change in SST or source gases, its O_3 change depends on the sensitivity of each term in the O_3 tendency equation to that change. Two models having the same O_3 profile but different chemistry and transport terms are unlikely to have the same response to a perturbation. Differences in sensitivities will result in different size effects on transport and chemistry terms for each model, and thus there is no reason to expect that models with the same O_3 profile today would predict the same response to a perturbation, i.e., an O_3 profile 50 years from now. By evaluating model transport processes in the LS we are effectively trying determine whether a model's O_3 transport tendency terms (horizontal and vertical) are realistic and how those terms affect a model's prediction of future O_3.

This figure also shows that the ozone 60°S-60°N return dates for the set of CCMs do not sort out by transport characteristics alone. For example, the four models with fast circulations (AMTRAC3, CCSRNIES, Niwa-SOCOL, and SOCOL, all in black) have return dates spanning the full range of the CCM

predictions, 2004-2058, yet the two models (UMUKCA) with the slowest circulations have return dates near 2025, earlier than the median (2030) or mean (2032) of the 15 CCMs.

Finally, this figure shows that the five CCMs with the best LS transport characteristics (red) have a narrow range of predicted return dates (2026-2040). The random selection of 5 return dates from this figure would result in a range of 14 years or less only 7% of the time, so this reduction is unlikely to be due to simply reducing the number of results compared. This suggests two things. One, that similar, credible LS transport leads to similar O_3 predictions, and two, that much of the large range of predicted O_3 return dates may be the result of problems with model transport. However, the O_3 tendency equation has a number of chemical terms and thus far the role of chemistry has not been considered. We expect that physically realistic transport processes are essential for realistic O_3 simulations, but we cannot conclude that the narrow range of predicted return dates is the result of good transport unless the models also have realistic chemistry.

5.3 Role of Chemistry

Chapter 6 of the SCR evaluated chemistry in CCMs by looking at radicals and radical precursors, tracer-tracer correlations, and photolysis rates. This section briefly summarizes major issues affecting O_3 that were identified in the chemistry chapter. It is beyond the scope of this paper to quantify how each model's column O_3 levels are influenced by each contributing process, e.g., transport, temperature, losses due to various families such as HO_x, NO_x, and ClO_x).

The SCR photochemical intercomparison ('PhotoComp') evaluated the accuracy of J-values, but only 9 CCMs participated. AMTRAC3, CCSRNIES, and Niwa-SOCOL showed the greatest inaccuracies, while UMSLIMCAT, WACCM, GEOSCCM, and LMDZ had highly accurate J-values (SCR, 2010). Photochemical steady state (PSS) model comparisons were made for both radicals and radical precursors. Niwa-SOCOL did not participate in the PSS comparisons, but it uses photolysis and chemistry schemes that are very similar to SOCOL (SCR, 2010). CCSRNIES, CNRM-ACM, and SOCOL had the most disagreements with observations for the precursors, and MRI and SOCOL had poor agreement with some radicals (e.g. O^1D, HO_x, NO_x/NO_y, or ClO/Cl_y). AMTRACS, CAM3.5, MRI, SOCOL, and ULAQ have a problem with the ClO/Cl_y ratio, and a problem with the Cl_y vs N_2O relationship was found for AMTRAC3, CNRM-ACM, and SOCOL. CCSRNIES, Niwa-SOCOL, SOCOL, and UMUKCA-METO have total chlorine in the upper stratosphere that exceeded total Cl emitted at the surface, indicating a lack of conservation of Cl_y. (The UMUKCA-METO model has excessive Cl_y due to mistreatment of tropospheric HCl removal.) The PSS comparison

identified no major issues with radicals or precursors for CMAM, EMAC, GEOSCCM, LMDZ, UMSLIMCAT, or WACCM.

Models with both transport and chemistry problems have the potential to produce compensating errors that result in fortuitous agreement with O_3 columns. Of the10 CCMs that agree to within 5% of observed annual mean 60°S-60°N ozone columns from 1979-2009 (the green lines in Figure 13), five have transport and chemistry problems. Niwa-SOCOL has a fast circulation, lack of conservation of Cl_y, and inaccuracies in its J-values. CNRM-ACM and SOCOL have fast circulations and disagreements with radical precursors, and SOCOL also lacks Cl_y conservation. ULAQ has a mixture of circulation and mixing problems and a problem with ClO/Cl_y. AMTRAC3 has a fast circulation and mixed accuracy for J-values. Problems were found in its ClO/Cl_y and Cl_y/N_2O relationships, which are probably due to the parameterization used to calculate Cl_y.

Of the five models that demonstrated very good LS transport, none had any major chemistry problems, although CAM3.5 has minor problems with ClO/Cl_y and NO_x/NO_y. All five except for UMSLIMCAT also agree closely with column O_3 observations shown in Figure 13. It is worth nothing that the LMDZ model performed very well in the photochemical, radical, and precursor evaluations and its only identified transport deficiency is slow ascent in the tropics below 50 hPa. As tropical LS ozone is typically very low (<2 ppm), it is possible that slow ascent below 50 hPa has only a small impact on its 60°S-60°N column mean. The LMDZ return date is within range predicted by the five models with the best LS transport.

5.4 Return to 1980 Column O_3 in the Antarctic

In addition to transport and gas phase chemistry, the temporal and spatial distribution of PSC-producing temperatures, heterogeneous chemistry, and the presence of mixing barriers also affect polar ozone columns. Figure 11 shows October Antarctic column O_3 and column O_3 anomalies relative to 1980 for 15 CCMs. The red lines represent models with good LS transport, black lines are models with good agreement with O_3 profiles but known transport problems, and the green lines are models with known transport problems that have poor agreement with O_3 profiles. The top panel shows that the models that most closely agree with observations (to within ±10%) include models with good transport and models with poor transport. The same is true for the models that are biased high. Good agreement with column O_3 observations is not, by itself, an indication of realistic transport.

The bottom panel shows column O_3 relative to each model's own 1980 value from 1960-2080. The 5 models with transport problems and poor agreement with O_3 profiles (green) have the latest return

dates of the 15 CCMs. The proximate cause for four of them is vortex Cl_y that is as high or higher than most models (only MRI has <3.0 ppb Cl_y). In addition to their previously identified transport problems, these five models also have additional problems affecting polar ozone loss processes. CNRM-ACM, MRI, and ULAQ lack the physical containment necessary to sequester high Cl_y/low NO_y air in the vortex, yet CNRM-ACM and ULAQ produce more than 3 ppb at 80°S (SCR, Ch. 5). CNRM-ACM has a Cl_y non-conservation problem in the LS (Ch. 6, SCR). All five models except CNRM-ACM have a problem with the ClO/Cl_y ratio. The UMUKCA models have good vortex isolation but have Antarctic spring temperatures that are 4-7K higher than the climatological mean over the 200-50 hPa region (SCR, Ch. 4). The later-than-average return dates for the UMUKCA models are consistent with their slow circulations. In spite of having only 2.4 ppb Cl_y in the present day vortex, MRI has the latest Antarctic return date of any CCM (2067). The likely cause for this is a missing ClO loss reaction (SCR, Ch. 6). This leads to higher levels of ClO for a given level of Cl_y, allowing greater polar ozone loss to occur in spite of low Cl_y.

The red and black lines in the bottom panel of Fig. 11 form two groups with a similar spread of return dates, about 20 years each. The black group, representing models with transport problems but reasonable O_3 profiles, is shifted ~6 years earlier than the red group (the models with good transport and good profile agreement). The difference between the two groups may reflect the faster circulation of most of the black group models and generally lower LS Cl_y levels found in them (SCR, Ch. 5). AMTRAC3 is an exception to this, having high vortex Cl_y (3 ppb) and a return date of 2050, comparable to the return dates of the green group models, most of which also have 3 ppb or higher Cl_y. These results point to a relationship between vortex Cl_y and return date.

Figure 12 compares the Antarctic return date with each model's simulated October 2005 mean 80°S Cl_y. There are two important results. The first is that the Antarctic return date is related to vortex Cl_y, and the higher the present day Cl_y simulated, the later the return date. This is true for 12 of the CCMs examined. The exceptions are Niwa-SOCOL, SOCOL, and MRI. Each has a problem with ClO/ Cl_y and a much later return date than expected based on their 80°S Cl_y. While UMUKCA-METO is not an exception, it is worth noting a known problem in this model's removal of tropospheric HCl results in excessive stratospheric Cl_y, affecting where vortex Cl_y falls on this line. The second result is that the faster circulation models (green) tend to have low Cl_y (~1.8-2.4 ppb) while the slower circulation models (blue) have higher Cl_y (2.9-3.8 ppb), so return date is generally related to the model's circulation strength. There are exceptions. As previously discussed, AMTRAC3 and CNRM-ACM both have high Cl_y at 80°S inconsistent with their fast circulation, giving them return dates similar to other high Cl_y simulations. For

AMTRAC3 this may be related to the parameterization used for production of Cl_y from CFCs. For CNRM-ACM, high Cl_y due to lack of conservation was found in the lowermost stratosphere, but stratospheric temperatures and jet locations and strengths are also reported to be very biased (Chapters 4 and 6, SCR). In summary, the higher the Cl_y, the later the Antarctic return date, but Cl_y in the vortex is not solely influenced by transport and ClO/Cl_y matters as well.

6. Conclusions: Understanding Model Predictions

Figure 13 shows 60°S-60°N column O_3 and column anomalies for 15 CCMs. The top panel shows annual zonal mean column O_3 observations along with 15 CCMs; the models that agree to within 5% of observations (~15 DU) are plotted in green. The middle panel shows column anomalies, indicating when each model returns to its 1980 60°S-60°N mean column value. The 10 models that agree best with global mean observations (green) span the full 54-yr range of return dates (2004-2058). No reduction in the range of predictions is found by selecting models based on their ability to reproduce ozone observations, nor is there any reduction caused by comparing a smaller set of results (10 rather than 15). A random selection of 10 return dates would result in a reduction of the range 57% of the time. The bottom panel of Figure 13 shows the same column anomalies, where the models having the best LS transport are plotted in red. The best LS transport was determined by N_2O and mean age agreement (Figures 7 and 8). The spread of predicted recovery dates is reduced to 14 years (2026-2040) when credible LS transport is considered. This reduction is unlikely to occur solely due to the reduction of the sample size (5 CCMs instead of 15), which would produce a range of more than 14 years 93% of the time when selected randomly. The models shown in black have identifiable problems with transport, such as fast or slow ascent, inappropriate tropical or vortex isolation, problems with radical/reservoir ratios (e.g., ClO/Cl_y), or lack of Cl_y conservation.

Column ozone is influenced by so many processes that a simulated column may be close to the observed value due to compensating effects from multiple problems. For this reason, column O_3 should not be used as a surrogate for model performance. Similarly, CCMs with good LS transport and good stratospheric chemistry will not necessarily produce good agreement with observed columns for several reasons, including lack of tropospheric chemistry, problems with upper stratospheric temperatures, and heterogeneous chemical processes. The primary focus of this paper is the effect of transport on O_3 in

CCM simulations. Other issues affecting O_3 were evaluated in the SCR; Morgenstern et al. (2010) discusses the approximations used for tropospheric chemistry in these CCMs.

Ozone recovery in CCMs is a result of decreasing CFCs that are forced by the mixing ratio boundary conditions used in the future scenario. The chemistry and transport evaluations that are applied to CCMs provide the means to explain much of the variation in return-to-1980 dates. Overall, Antarctic recovery is a function of LS Cl_y, and Cl_y levels are to a large extent controlled by transport. The models with LS transport in agreement with observations (i.e., N_2O and mean age) show a reduced range of return dates. Models with slow circulation have high Cl_y and a later recovery, but some models have high Cl_y due to a parameterization or a conservation issue. One model has low Cl_y but has a missing ClO loss reaction, so its recovery is similar to models with high Cl_y. Except for AMTRAC3, all the models with a fast circulation (shown in green in Fig. 12), also have a problem with excess Cl_y in the UT/LS (Ch. 6, SCR). The full range of recovery dates in the Antarctic for the 15 CCMs is 2028-2067 (39 years), but the range for models that do not have a Cl_y conservation problem is 2037-2067 (30 years). Transport problems also contribute to this range: slow circulation in the UMUKCA-UCAM model, which results in high vortex Cl_y, is responsible for this model's 2067 predicted return. The range of predicted return dates is 2037-2056 (19 years) for models whose chemistry and transport have been shown to be free of major disagreement with observations.

While much uncertainty in prediction exists because we can only guess what future ODS and GHG emissions will be, the current range of predicted return dates is unnecessarily large due to identifiable model transport and chemistry deficiencies. This study shows that much of the range of future ozone predictions can be directly attributed to identifiable modeling problems in chemistry and transport. It is remarkable, and perhaps an encouraging sign of progress in chemistry climate modeling, that a handful of models having different dynamical cores, transport schemes, chemical solvers, spatial resolutions, etc., all produce very similar lower stratospheric O_3 and column O_3 return dates. The explanation for the similar behavior in these models is suggested by the transport diagnostic presented in this study: these models have credible representation of important physical processes that affect the distribution of ozone and other trace constituents involved in O_3 chemistry. Having models with credible physical processes in the lower stratosphere that simulate similar O_3 behavior increases our confidence in the current understanding of the essential chemical and dynamical processes controlling ozone.

Acknowledgments. We thank Ashley Jones for use of the ACE N_2O climatology and Andreas Engel for use of the tropical mean age profiles. We thank Michael Prather, Huisheng Bian, Ross Salawitch, and Tim Canty for their contributions to the photochemical evaluations in the SCR. CCSRNIES research was supported by the Global Environmental Research Foundation of the Ministry of the Environment of Japan. CCSRNIES and MRI simulations were completed with the supercomputer at NIES, Japan. This work was supported by the NASA Atmospheric Chemistry, Modeling, Analysis, and Prediction Program.

References

Akiyoshi, H., L. B. Zhou, Y. Yamashita, K. Sakamoto, M. Yoshiki, T. Nagashima, M. Takahashi, J. Kurokawa, M. Takigawa, and T. Imamura (2009), A CCM simulation of the breakup of the Antarctic polar vortex in the years 1980–2004 under the CCMVal scenarios, J. Geophys. Res., 114, D03103, doi:10.1029/2007JD009261.

Andersen, S.B., et al. (2006), Comparison of recent modeled and observed trends in total column ozone, J. Geophys. Res., 111, D02303, doi:10.1029/2005JD006091.

Andrews, A. E., et al. (2001), Mean age of stratospheric air derived from in situ observation of CO_2, CH_4, and N_2O, J. Geophys. Res., 106(D23), 32,295-32,314.

Austin, J., and N. Butchart (2003), Coupled chemistry-climate model simulations of the period 1980–2020: Ozone depletion and the start of ozone recovery, Q. J. R. Meteorol. Soc., 129, 3225–3249.

Austin, J., and Wilson, R.J. (2006), Ensemble simulations of the decline and recovery of stratospheric ozone, *J. Geophys. Res.*, 111, doi: 10.1029/2005JD006907.

Austin, J., et al. (2010), Chemistry-climate model simulations of spring Antarctic ozone, *J. Geophys. Res.*, 115, D00M11, doi:10.1029/2009JD013577.

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

Déqué, M. (2007), Frequency of precipitation and temperature extremes over France in an anthropogenic scenario: Model results and statistical correction according to observed values, Global Planet. Change, 57, 16–26.

Douglass, A.R., Rood, R.B., Kawa, S.R., and Allen, D.J. (1997), A three-dimensional simulation of the evolution of middle latitude winter ozone in the middle stratosphere, J. Geophys. Res., 102, 19,217-19,232.

Douglass, A.R., Prather, M.J., Hall, T.M., Strahan, S.E., Rasch, P.J., Sparling, L.C., Coy, L., and Rodriguez, J.M. (1999), Choosing meteorological input for the global modeling initiative assessment of high-speed aircraft, J. Geophys. Res., 104, 27,545-27,564.

Douglass, A.R., Stolarski, R.S., Strahan, S.E., and Connell, P.S. (2004), Radicals and reservoirs in the GMI chemistry and transport model: Comparison to measurements, J. Geophys. Res., 109, D16302, doi: 10.1029/2004JD004632.

Engel, A., et al. (2009), Age of stratospheric air unchanged within uncertainties over the past 30 years, Nat. Geo., 2, 28-31.

Eyring et al. (2006), Assessment of temperature, trace species, and ozone in chemistry climate model simulations of the recent past2006, J. Geophys. Res, 111, D22308, doi:10.1029/2006JD007327.

Garcia, R. R., D. Marsh, D. E. Kinnison, B. Boville, and F. Sassi (2007), Simulations of secular trends in the middle atmosphere, 1950-2003, *J. Geophys. Res.*, 112, doi:10.1029/2006JD007485.

Hall, T.M., Waugh. D.W., Boering, K.A., and Plumb, R.A. (1999), Evaluation of transport in stratospheric models, J. Geophys. Res., 104, 18,815-18,839.

Jackman, C.H., Douglass, A.R., Brueske, K.F., and Stein, S.A.(1991), The Influence of Dynamics on Two-Dimensional Model Results: Simulations of ^{14}C and Stratospheric Aircraft NO_x Injections, J. Geosphys. Res., 96, 22,259-22,572.

Jöckel, P., et al. (2006), The atmospheric chemistry general circulation model ECHAM5/MESSy1: Consistent simulation of ozone from the surface to the mesosphere, Atmos. Chem. Phys., 6, 5067–5104.

Jones, A, et al., ACE climatologies (submitted?)

Jourdain, L., S. Bekki, F. Lott, and F. Lefèvre (2008), The coupled chemistry-climate model LMDz-REPROBUS: Description and evaluation of a transient simulation of the period 1980–1999, Ann. Geophys., 26, 1391–1413.

Lamarque, J.-F., D. E. Kinnison, P. G. Hess, and F. M. Vitt (2008), Simulated lower stratospheric trends between 1970 and 2005: Identifying the role of climate and composition changes, J. Geophys. Res., 113, D12301, doi:10.1029/2007JD009277.

Lambert et al. (2007), Validation of the Aura Microwave Limb Sounds middle atmosphere water vapor and nitrous oxide measurements, J. Geophys. Res., 112, doi:10.1029/2007JD8724.

Livesey, N., Read, W.G., Filipiak, M.J, et al.: Earth Observing System (EOS) Microwave Limb Sounder (MLS) Version 2.2 Level 2 data quality and description document, JPL D-32381, 2007.

Morgenstern, O., P. Braesicke, F. M. O'Connor, A. C. Bushell, C. E. Johnson, S. M. Osprey, and J. A. Pyle (2009), Evaluation of the new UKCA climate-composition model–Part 1: The stratosphere, Geosci. Model Dev., 2, 43–57.

Morgenstern, O., et al. (2010), Review of the formulation of present-generation stratospheric chemistry-climate models and associated external forcings, J. Geophys. Res., 115, D00M02, doi:10.1029/2009JD013728.

Neu, J.L. and Plumb, R.A. (1999), Age of air in a "leaky pipe" model of stratospheric transport, J. Geophys. Res., 104, 19,243-19,255.

Pawson, S., Stolarski, R.S., Douglass, A.R., Newman, P.A., Nielsen, J.E., Frith, S.M., and Gupta, M.L. (2008), Goddard Earth Observing System chemistry-climate model simulations of stratospheric

ozone-temperature coupling between 1950 and 2005, J. Geophys. Res., 113, D12103, doi:10.1029/2007JD009511.

Perliski, L.M., Solomon, S., and London, J. (1989), On the interpretation of seasonal variations of stratospheric ozone, Planet. Space Sci, 37, 1527-1538.

Pitari, G., E. Mancini, V. Rizi, and D. T. Shindell (2002), Impact of future climate and emission changes on stratospheric aerosols and ozone, J. Atmos. Sci., 59, 414–440.

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. (2003), Chlorine budget and partitioning during the Stratospheric Aerosol and Gas Experiment (SAGE) III Ozone Loss and Validation Experiment (SOLVE), J. Geophys. Res., 108(D5), 4173, doi:10.1029/2001JD002040.

Schoeberl, M.R., Sparling, L.C., Jackman, C.H., and Fleming, E.L. (2000), A lagrangian view of stratospheric trace gas distributions, J. Geophys. Res., 105, 1537-1552.

Schraner, M., et al. (2008), Technical note: Chemistry-climate model SOCOL–Version 2.0 with improved transport and chemistry/microphysics schemes, Atmos. Chem. Phys., 8, 5957–5974.

Scinocca, J. F., N. A. McFarlane, M. Lazare, J. Li, and D. Plummer (2008), Technical note: The CCCma third generation AGCM and its extension into the middle atmosphere, Atmos. Chem. Phys., 8, 7055–7074.

Shibata, K., and M. Deushi (2008a), Long-term variations and trends in the simulation of the middle atmosphere 1980–2004 by the chemistry-climate model of the Meteorological Research Institute, Ann. Geophys., 26, 1299–1326.

Shibata, K., and M. Deushi (2008b), Simulation of the Stratospheric Circulation and Ozone During the Recent Past (1980–2004) With the MRI Chemistry-Climate Model, CGER' s Supercomput. Monogr. Rep., vol. 13, 154 pp., Cent. for Global Environ. Res., Natl. Inst. for Environ. Studies, Tsukuba, Japan.

Stenke, A., V. Grewe, and M. Ponater (2008), Lagrangian transport of water vapor and cloud water in the ECHAM4 GCM and its impact on the cold bias, Clim. Dyn., 31, 491–506.

Stenke, A., M. Dameris, V. Grewe, and H. Garny (2009), Implications of Lagrangian transport for coupled chemistry - climate simulations, Atmos. Chem. Phys., 9, 5489–5504.

SPARC CCMVal (2010), SPARC Report on the Evaluation of Chemistry-Climate Models, V. Eyring, T.G. Shepherd, D.W. Waugh (Eds.), SPARC Report No. 5, WCRP-132, WMO/TD-No. 1526, http://www.atmosp.physics.utoronto.ca/SPARC

Stolarksi, R.S. and Douglass, A.R. (1985), Parameterization of the Photochemistry of Stratospheric Ozone Including Catalytic Loss Processes, J. Geophys. Res., 90, 10,709-10,718.

Stolarski, R.S. and Frith, S.M. (2006), Search for evidence of trend slow-down in the long-term TOMS/SBUV total ozone data record: the importance of instrument drift uncertainty, Atmos. Chem. Phys., 6, 4057-4065.

Strahan, S.E., Duncan, B.N., and Hoor, P. (2007), Observationally derived transport diagnostics for the lowermost stratosphere and their application to the GMI chemistry and transport model, Atmos. Chem. Phys., 7, 2435-2445.

Strong, K. et al. (2008), Validation of ACE-FTS N_2O measurements, Atmos. Chem. Phys., 8, 4759-4786.

Teyssèdre, H., et al. (2007), A new tropospheric and stratospheric chemistry and transport model MOCAGE - Climate for multi-year studies: Evaluation of the present - day climatology and sensitivity to surface processes, Atmos. Chem. Phys., 7, 5815–5860.

Tian, W., and M. P. Chipperfield (2005), A new coupled chemistry-climate model for the stratosphere: The importance of coupling for future O3 -climate predictions, Q. J. R. Meteorol. Soc., 131, 281–303.

Waugh, D.W. and Eyring, V. (2008), Quantitative performance metrics for stratospheric-resolving chemistry-climate models, Atmos. Chem. Phys., 8, 5699-5713.

Waters, J., Froidevaux, L., Harwood, R.S., et al., The Earth Observing System Microwave Limb Sounder (EOS MLS) on the Aura satellite, IEEE Trans. Geosci. Remote Sensing, 44, 1075-1092, 2006.

World Meteorological Organization (WMO) (2003), Scientific assessment of ozone depletion: 2002, WMO Report No. 47, Geneva, Switzerland.

World Meteorological Organization (WMO) (2007), Scientific assessment of ozone depletion: 2006, WMO Report No. 50, Geneva Switzerland.

World Meteorological Organization (WMO) (in press), Scientific assessment of ozone depletion: 2010, WMO Report No. X, Geneva Switzerland.

Table 1. Chemistry Climate Models participating in CCMVal and WMO (2011)

CCM Name	**Atmospheric GCM**	**Reference**	**Horiz. Res. of Advection, Levels, Top**
AMTRAC3	AM3	Austin and Wilson (2008)	~200 km, L48, 0.017 hPa
CAM3.5	CAM	Lamarque et al. (2008)	1.9° x 2.5°, L26, 3.5 hPa
CCSRNIES*	CCSR/NIES AGCM 5.4g	Akiyoshi et al. (2009)	T42L34, 0.012 hPa
CMAM	AGCM3	Scinocca et al. (2008)	T31L71, 0.00081 hPa
CNRM-ACM	ARPEGE-Climate v4.6	Déqué (2007); Teyssèdre et al. (2007)	T42L60, 0.07 hPa
E39CA*†	ECHAM4	Stenke et al. (2008, 2009)	T30L39, 10 hPa
EMAC v1.6*†	ECHAM5	Joeckel et al. (2006)	T42L90, 0.01 hPa
GEOSCCM v2	GEOS5	Pawson et al. (2008)	2° x2.5°, L72, 0.015 hPa
LMDZrepro	LMDz	Jourdain et al. (2008)	2.5° x3.75°, L50, 0.07 hPa
MRI	MJ98	Shibata and Deushi (2008a,b)	T42L68, 0.01 hPa
Niwa-SOCOL	MAECHAM4	Schraner et al. (2008)	T30L39, 0.01 hPa
SOCOL	MAECHAM4	Schraner et al. (2008)	T30L39, 0.01 hPa
ULAQ	ULAQ-GCM	Pitari et al. (2002)	R6L26, 0.01 hPa
UMETRAC†	HadAM3 L64	Austin and Butchart (2003)	2.5°x3.75°, L64, 0.01 hPa
UMSLIMCAT	HadAM3 L64	Tian and Chipperfield (2005)	2.5°x3.75°, L64, 0.01 hPa
UMUKCA-METO	HadGEM-A	Morgenstern et al. (2008,2009)	2.5°x3.75°, L60, 84 km
UMUKCA-UCAM	HadGEM-A	Morgenstern et al. (2009)	2.5°x3.75°, L60, 84 km
WACCM	CAM	Garcia et al (2007)	1.9°x2.5°, L66, $6x10^{-6}$ hPa

*No mean age output submitted.

†No REF-B2 (future scenario) submitted.

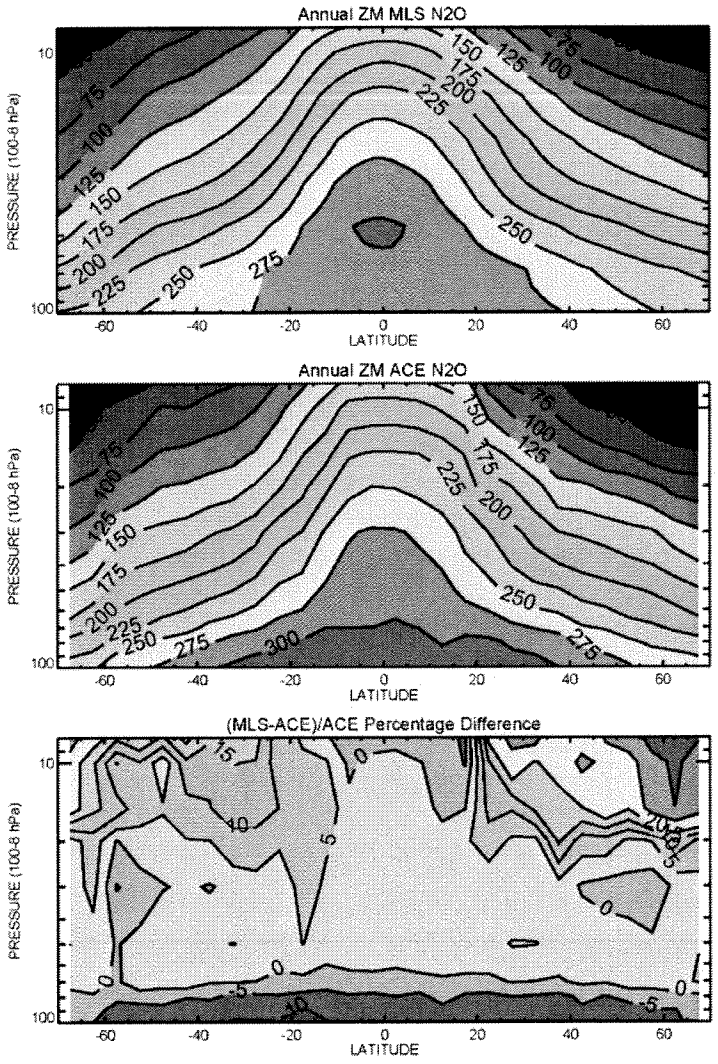

Figure 1. Top: Annual zonal mean N_2O calculated from 5 years of ACE observations, 2004-2008. Middle: Annual zonal mean N_2O calculated from 5 years of Aura MLS N_2O observations, 2005-2009. Bottom: (MLS-ACE)/ACE percentage difference. Lowest MLS N_2O level retrieved is 100 hPa.

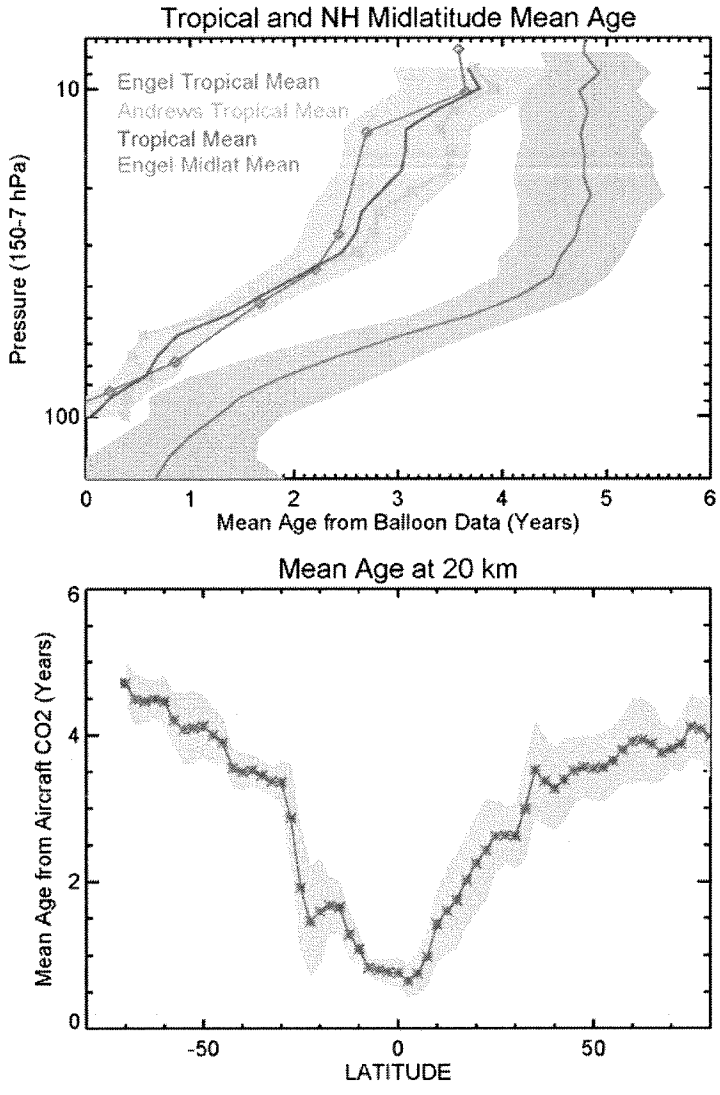

Figure 2. $CO_2$and SF_6-derived mean age observations used in this study. Top: Tropical and midlatitude (35-50 °N) mean profiles and their 1σ uncertainties (Andrews et al., 2001, and Engel, pers. comm.). The Engel et al. (2009) midlatitude mean is an average of 27 profiles. Bottom: The Andrews et al. (2001) aircraft CO_2 mean age and 1σ uncertainty.

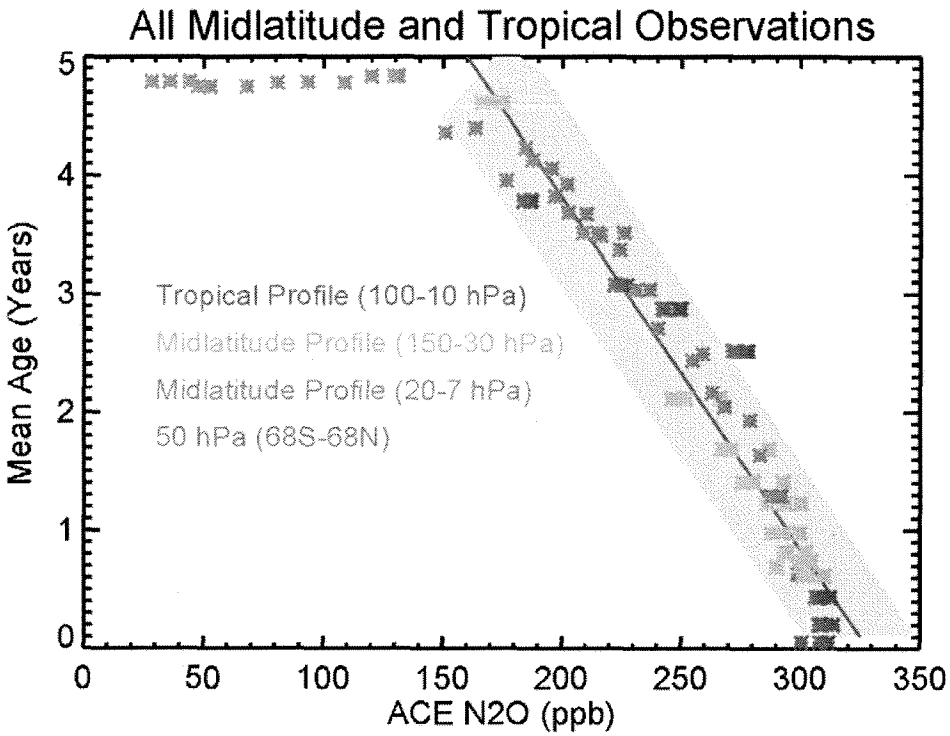

Figure 3. Scatterplot of all mean age observations and their co-located climatological ACE N_2O mixing ratios. The points between 0-4.5 yrs mean age and 150-315 ppb N_2O have been fit to a line. The yellow shading shows the range of uncertainty for the mean age and N_2O observations with respect to the fitted line. Blue points are the 20 km (50 hPa) measurements, green points are from the midlatitude profile below 20 hPa, and red points are from the tropical profile at 10 hPa and below. A breakdown in the correlation between mean age and N_2O is found at 20 hPa and above in the midlatitudes where mean age is >4.5 years (orange points).

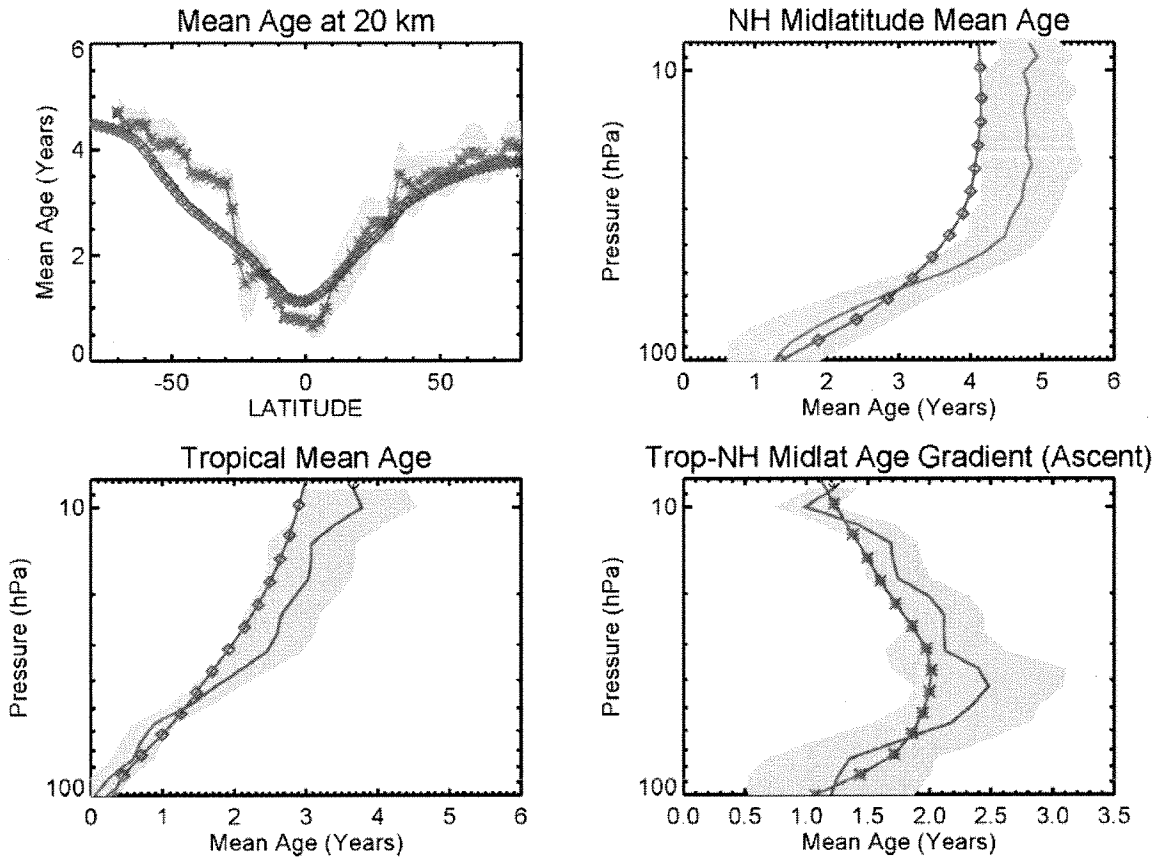

Figure 4. Mean age from a 20-yr simulation of the GMI-G4GCM (red) and mean age observations (blue points or lines with yellow shading for the 1σ uncertainty. The close agreement with the tropical-midlatitude mean age gradient indicates a realistic tropical ascent rate. The good agreement with the tropical mean age profile indicates a good balance between ascent and subtropical mixing in the simulation. The simulated mean age agrees within the observed uncertainties everywhere except for the SH midlatitude LS and the NH midlatitudes 40-10 hPa.

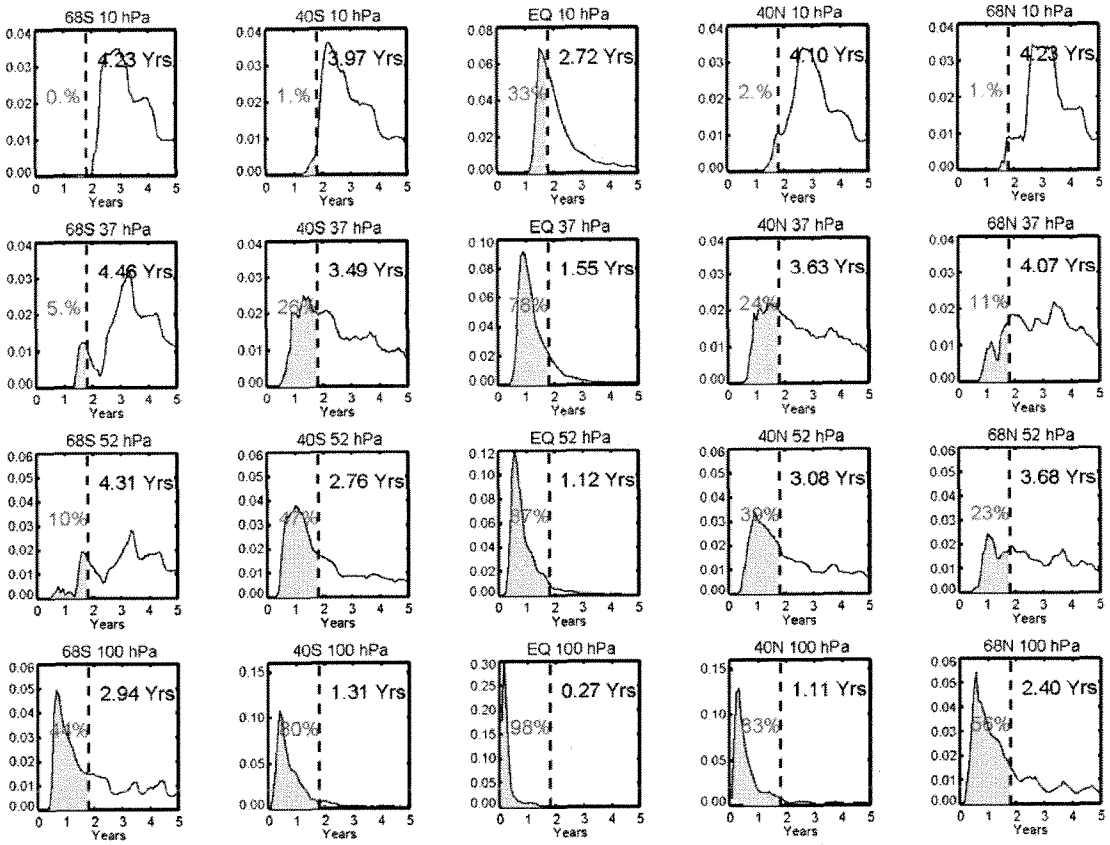

Figure 5. Age spectra from the GMI-G4GCM, from 100-10 hPa and 68°S-68°N. The y-axis is spectral density/year.The mean age at each location is shown in the upper right corner of each panel. The yellow shaded area under the curve shows the part of the spectrum containing the youngest 1.8 years. The area under that curve as percentage of all the age elements is shown in red.

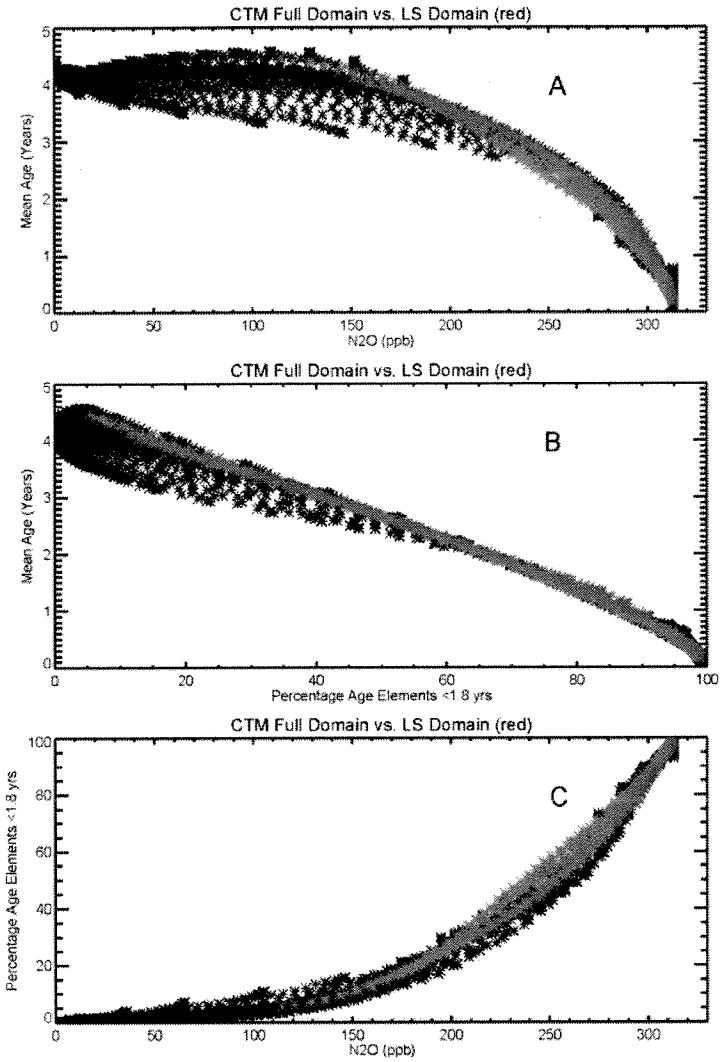

Figure 6. Relationship in the GMI-G4GCM between a) mean age and N_2O, b) mean age and the percentage of young age elements, c) the percentage of young age elements and N_2O over the entire model domain (black). Points in red show the relationship over the domain of the observations used in this study (68°S-68°N, 30-150 hPa). In each panel, the relationship in the LS domain is nearly linear and mostly monotonic. There are slight hemispheric differences, mostly noticeable in the form of two slightly offset red curves in panel c). Panel b) shows a linear relationship between mean age and the percentage of the mean coming from young elements in the LS domain. There is a unique relationship between N_2O and mean age in the lower stratosphere that is determined by transport.

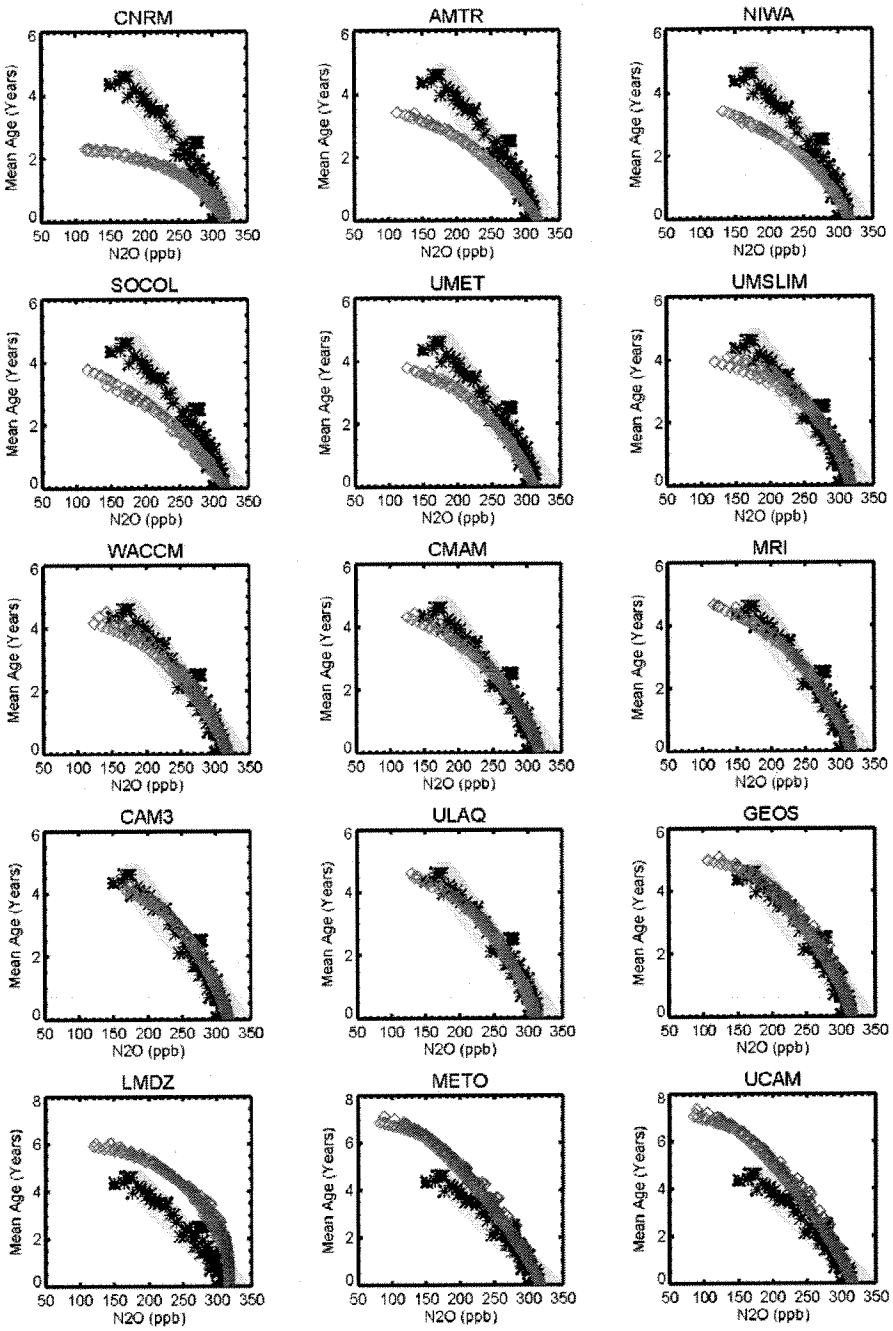

Figure 7. N_2O and mean age relationship in 15 CCMs (red) in the same domain as the observations (black). The yellow shaded area represents the uncertainty of the fitted line (from Fig. 3). Model curves that fall off faster (slower) than observed indicate a fast (slow) circulation. Models results are sorted by fastest circulations at the top, slowest circulations at the bottom. Results are from the REF-B1 simulation of the recent past.

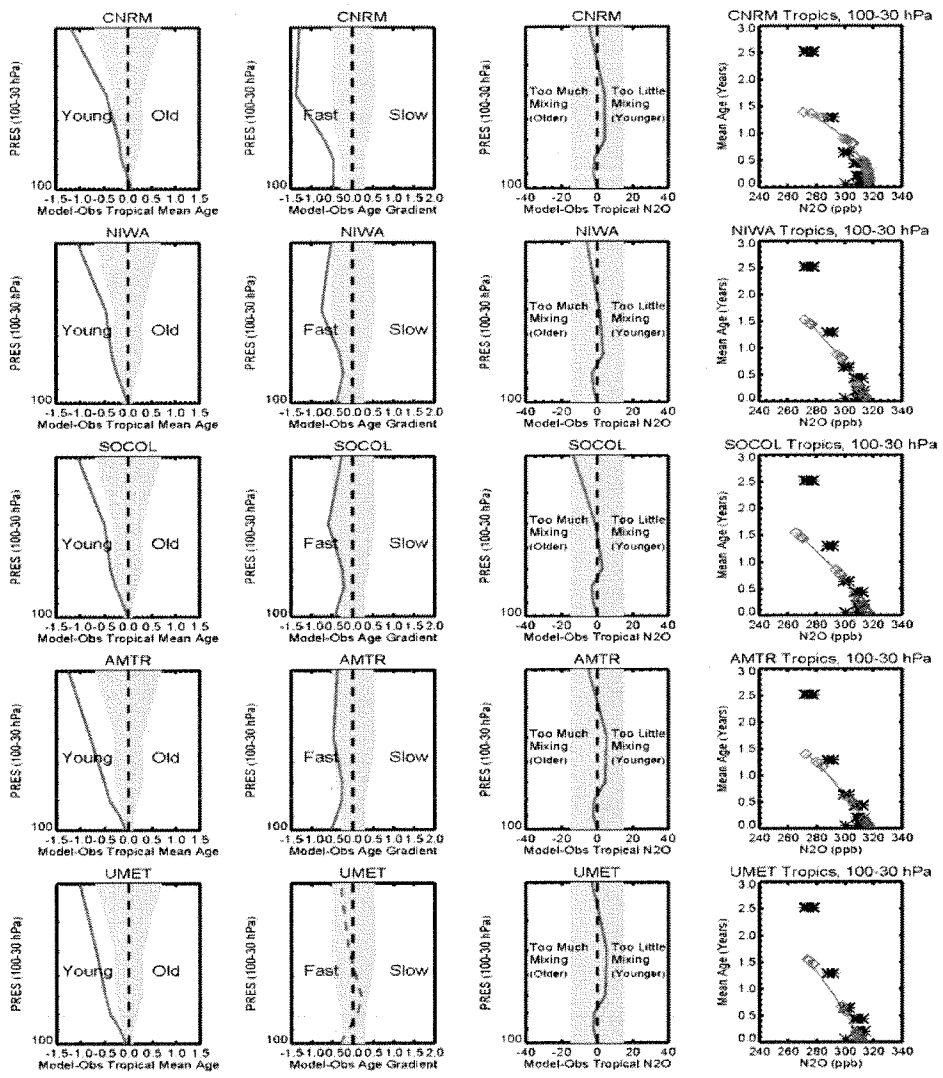

Figure 8a. Diagnostic plots evaluating tropical ascent and tropical-midlatitude (horizontal) mixing in CCMs. Results are from the REF-B1 simulation of the recent past. The first 3 columns show the difference between model and observed profiles of 1) tropical mean age (years), 2) the horizontal mean age gradient (years), and 3) tropical N_2O (ppb). The yellow shading indicates the ±1σ uncertainty in the observations. The rightmost column compares the simulated (red) and observed (black) mean age/N_2O relationship in the tropical LS. The middle columns diagnose the reason for the (dis)agreement with the tropical mean age profile (far left) by identifying ascent and or mixing problems. Fast ascent decreases mean age everywhere while strong mixing across the subtropics increases mean age everywhere. A dashed red line indicates there is an issue that complicates the interpretation of that diagnostic – see text. The 5 CCMs shown have the fastest circulations.

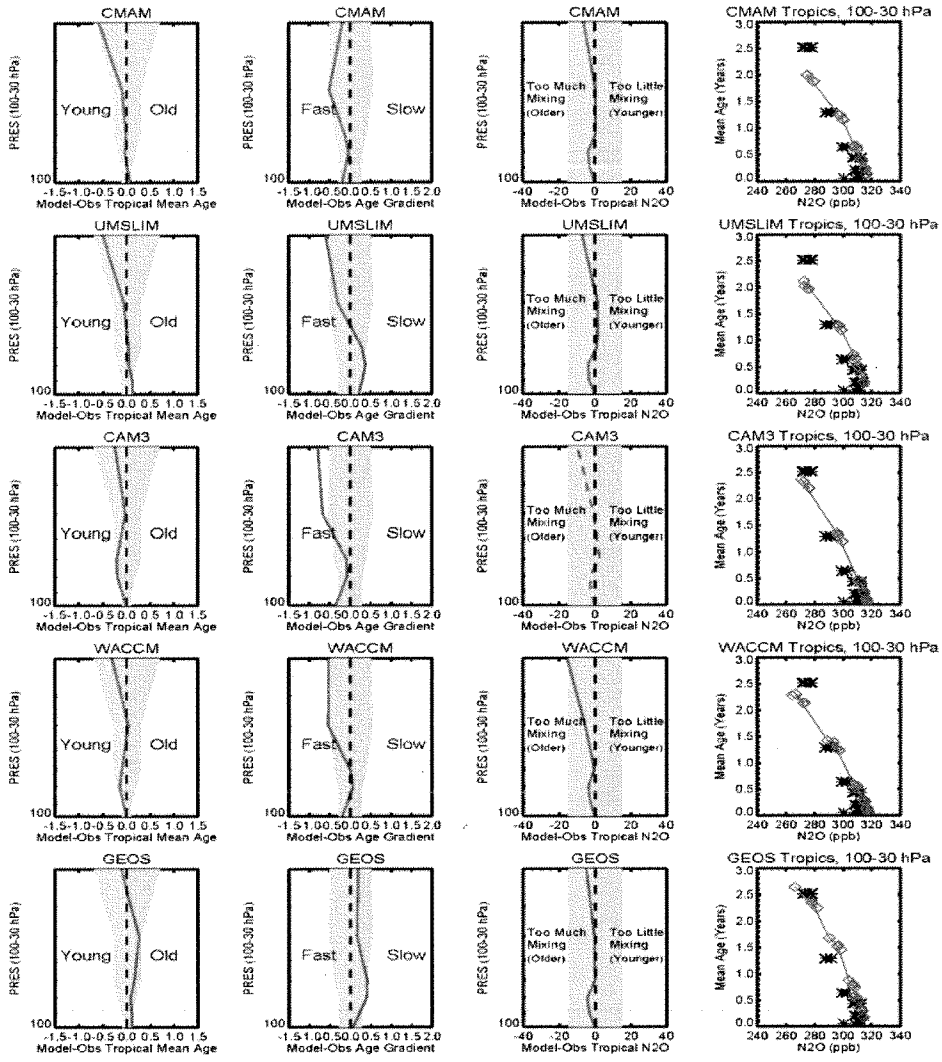

Figure 8b. Same as Figure 8a except for the the 5 CCMs with the most realistic tropical circulation and mixing.

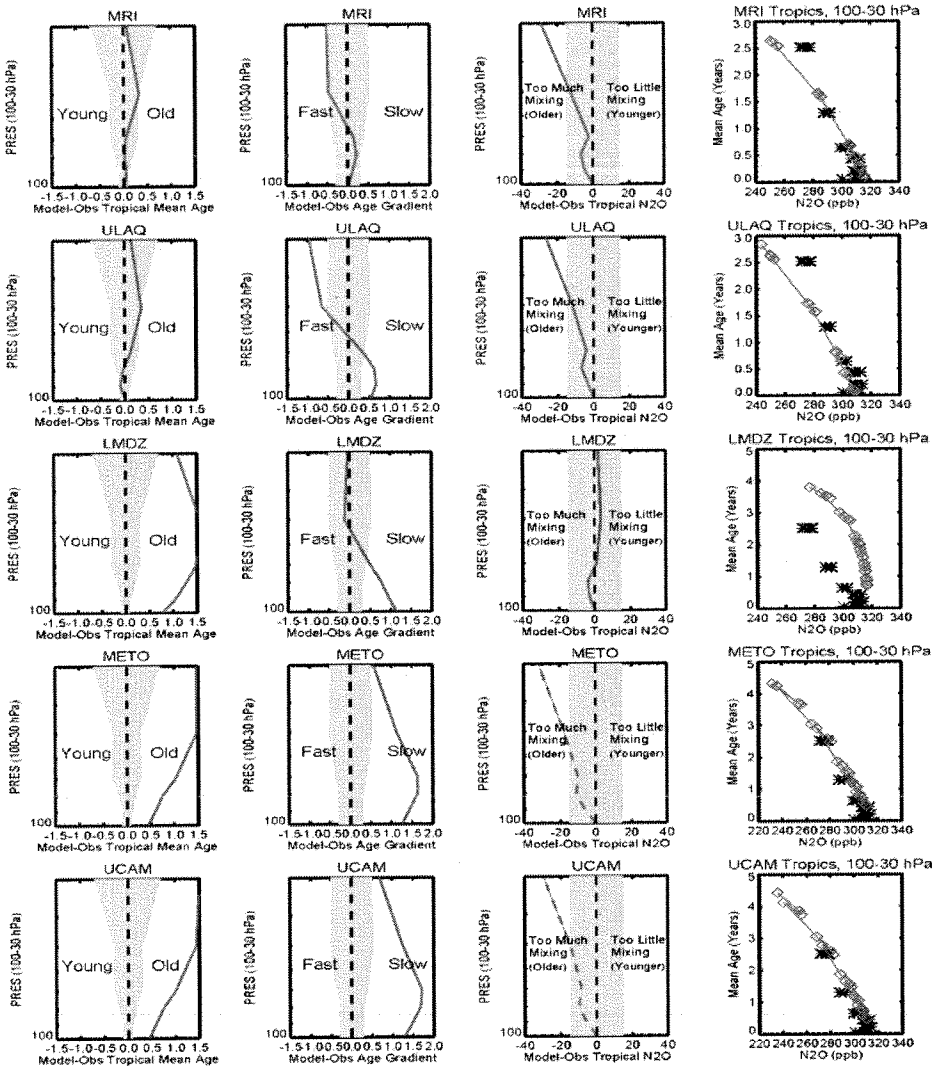

Figure 8c. Same as Figure 8a except for the 5 CCMs with slow circulations or excessive horizontal mixing.

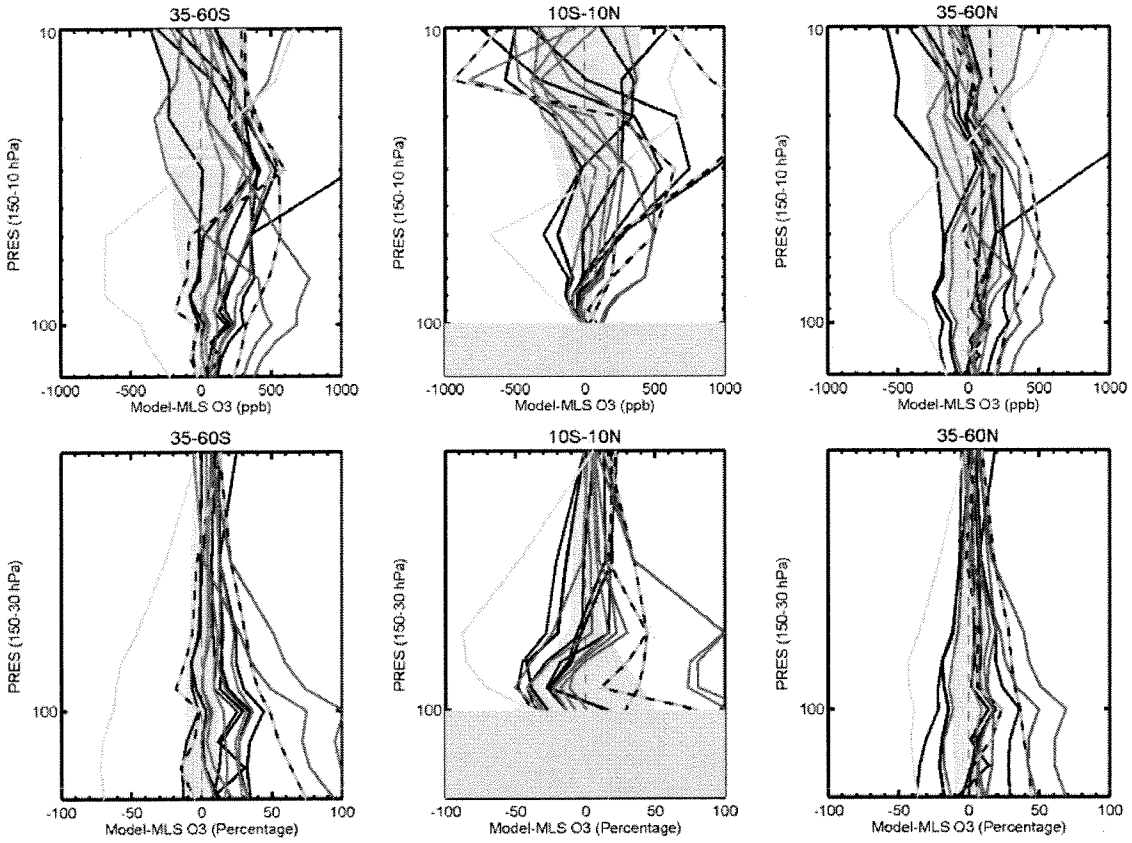

Figure 9. Differences between models and observed O_3 profiles for the SH and NH midlatitudes and the tropics. Model results are from the REF-B1 simulation of the recent past. The models with the most realistic tropical transport processes (Fig. 8b) are shown in red. The top panels show O_3 differences in ppb while the bottom panels express the difference as a percentage. The models in dark blue have the highest O_3 (slowest circulation and oldest LS mean age), while the model shown in light blue has the lowest O_3 (fastest circulation and youngest mean age). The models shown in dashed green have competing transport errors (fast or variable ascent rates and strong mixing). Yellow shading indicates MLS O_3 2σ systematic uncertainty (Livesey, 2007). Most of the CCMs do not have a tropospheric chemistry scheme so tropical O_3 below 100 hPa should not be compared.

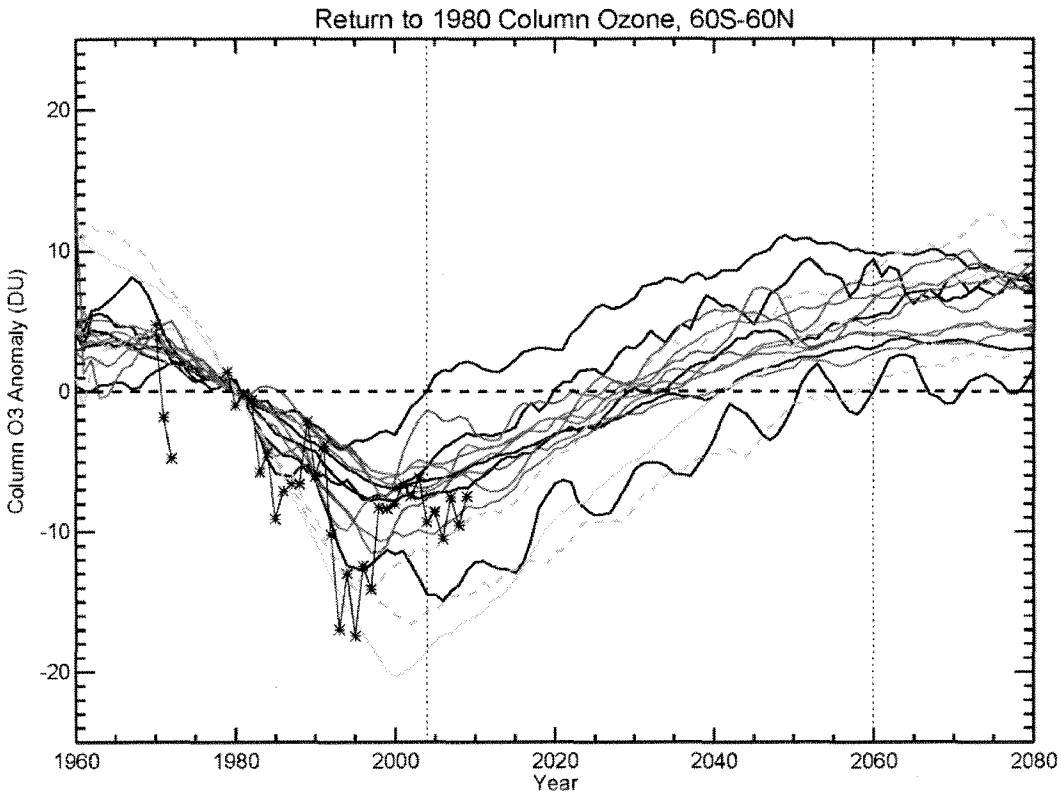

Figure 10. Annual mean 60°S-60°N column O_3 anomalies with respect to 1980 for 15 CCMs. Results are from the REF-B2 future simulation. Observations are shown with black asterisks. Column anomalies for the 5 CCMs with the best representation of LS transport processes, which also have good agreement with LS O_3 profiles, are shown in red. Models with significant transport issues that do not agree with O_3profiles are shown in green and blue (same colors as in Figure 9). The models shown in black have transport issues but have O_3profiles in reasonable agreement with observations (Fig. 9)

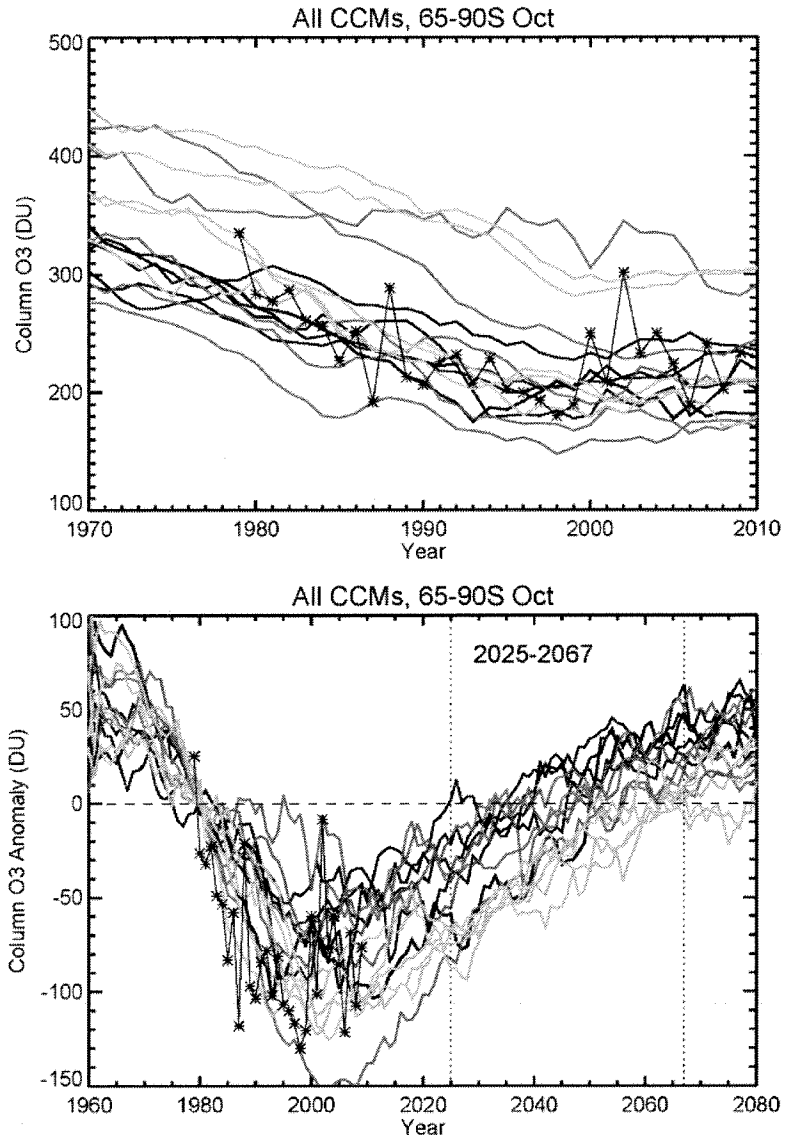

Figure 11. October zonal mean column O_3 (top) and column anomalies with respect to 1980 (bottom) for 15 CCMs and observations. Results are from the REF-B2 future simulation. Observations are shown by black asterisks. As in the previous figures, CCMs in red have the best LS transport as well as good agreement with LS O_3 profiles. CCMs shown in black have problems with LS transport but good agreement with LS O_3 profiles. The 5 CCMs shown in green have the greatest LS transport problems and do not agree well with LS O_3 profiles.

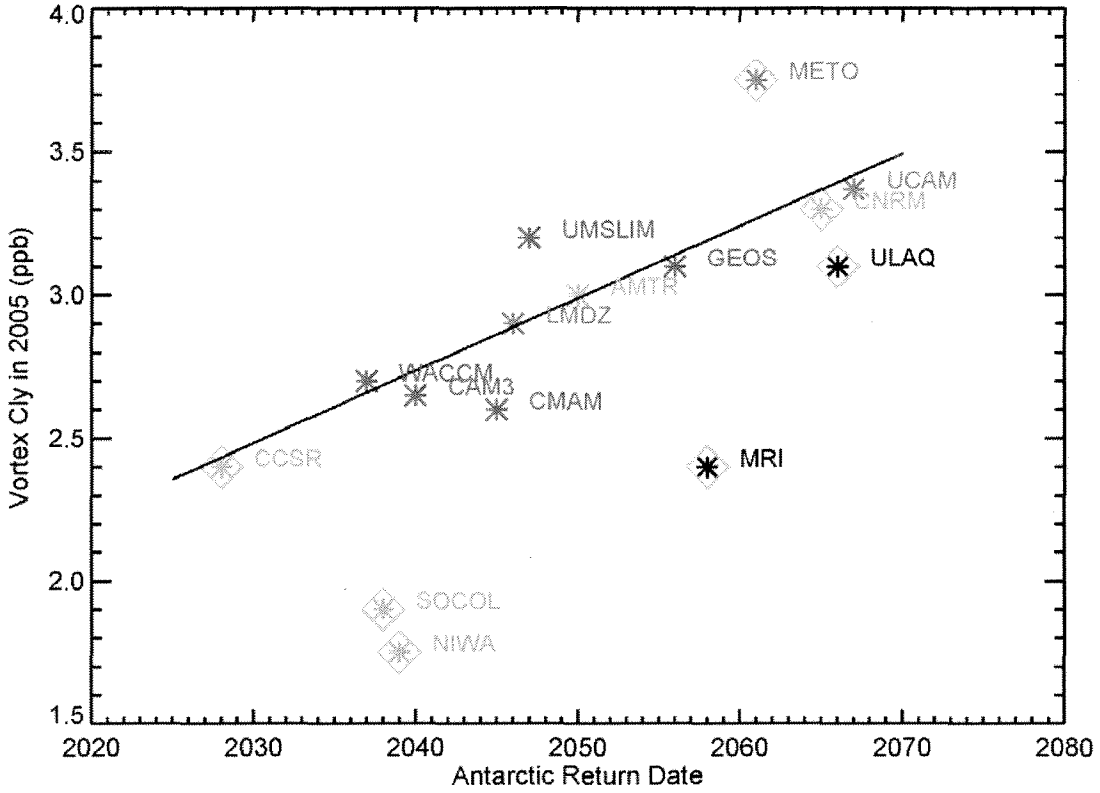

Figure 12. The relationship between model return-to-1980 date for the October Antarctic column O_3 and model Cl_y at 80°S 50 hPa in 2005. Results are from the REF-B1 simulation of the recent past. Models in red have good LS transport. Models in green (blue) have fast (slow) LS circulation. Models in black have excessive subtropical mixing and some problems with the tropical ascent rate. Models with an orange diamond either have a problem with Cl_y conservation (MRI, Niwa-SOCOL, and SOCOL), or excessive Cl_y in the UT/LS (as diagnosed by Chapter 6 of the SCR). MRI, Niwa-SOCOL, and SOCOL also all have diagnosed problems with ClO/Cl_y; they are not included in the fit shown.

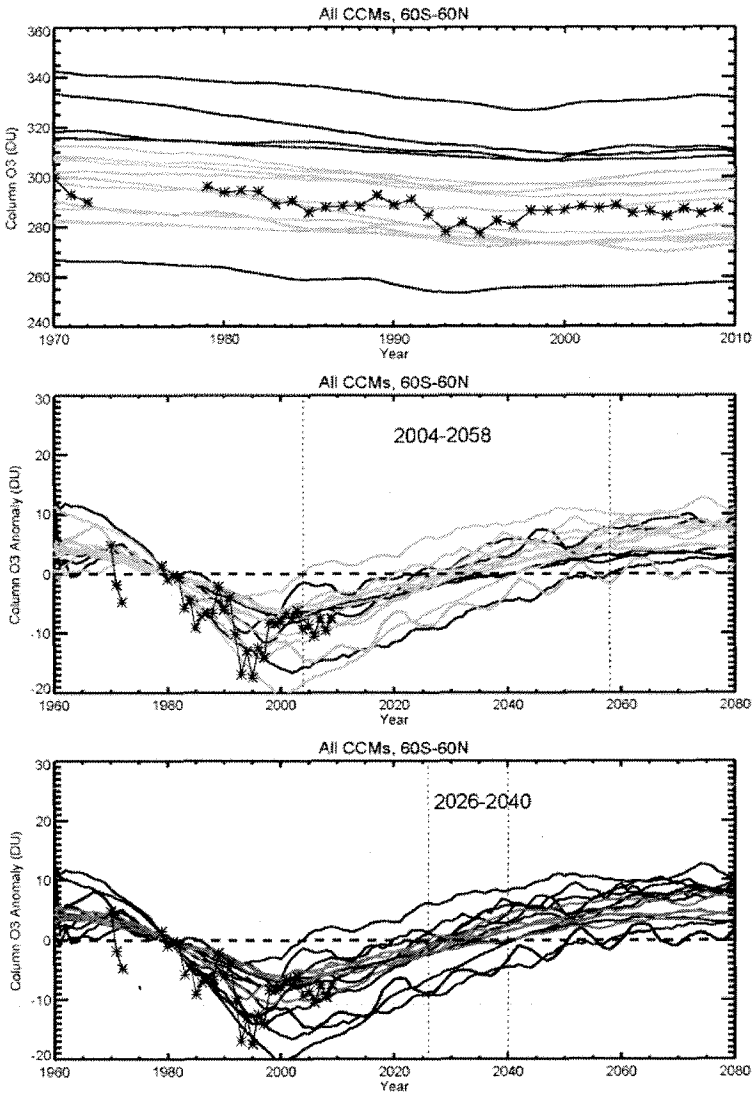

Figure 13. Top: Annual mean 60°S-60°N column O_3 in 15 CCMs from 1970 to the present. Results are from the REF-B2 future simulation. Observations are shown with black asterisks in each panel. The models in green show agreement with the observations to within 5%; models in black do not agree with observations to within 5%. Middle: the same CCM output and observations, but plotted as anomalies with respect to 1980. The time period is 1960-2080. The dashed vertical lines show the earliest and latest predicted return dates for the models with the best agreement with observations, which is the same as the range of all 15 CCMS (2004-2058). Bottom: the same CCM output but models with the best LS transport are shown in red. The dashed vertical lines show the earliest and latest predicted return dates for the red models, 2026-2040.