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] J. F. Scinocca,[14] T. G. Shepherd,[15] K. Shibata,[16] D. Smale,[12] H. Teyssèdre,[11] W. Tian,[7] and Y. Yamashita[3]

Received 19 November 2010; revised 13 May 2011; accepted 13 June 2011; published 9 September 2011.

[1] We use observations of N_2O and mean age to identify realistic transport in models in order to explain their ozone predictions. The results are applied to 15 chemistry climate models (CCMs) participating in the 2010 World Meteorological Organization ozone assessment. Comparison of the observed and simulated N_2O, mean age and their compact correlation identifies models with fast or slow circulations and reveals details of model ascent and tropical isolation. This process-oriented diagnostic is more useful than mean age alone because it identifies models with compensating transport deficiencies that produce fortuitous agreement with mean age. The diagnosed model transport behavior is related to a model's ability to produce realistic lower stratosphere (LS) O_3 profiles. Models with the greatest tropical transport problems compare poorly with O_3 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 Stratospheric Processes and their Role in Climate (SPARC) CCMVal Report to explain the range of CCM predictions for the return-to-1980 dates for global (60°S–60°N) and Antarctic column ozone. Antarctic O_3 return dates are generally correlated with vortex Cl_y levels, and vortex Cl_y is generally correlated with the model's circulation, although model Cl chemistry and conservation problems also have a significant effect on return date. In both regions, models with good LS transport and chemistry 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 broad due to identifiable model deficiencies.

Citation: Strahan, S. E., et al. (2011), Using transport diagnostics to understand chemistry climate model ozone simulations, *J. Geophys. Res.*, *116*, D17302, doi:10.1029/2010JD015360.

1. Introduction

[2] 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* (*WMO*), 2011, 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,

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

[1]Universities Space Research Association, Columbia, 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]Met Office Hadley Centre, 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]Department of Physics, University of Toronto, Toronto, Ontario, Canada.
[16]Meteorological Research Institute, Japan Meteorological Agency, Tsukuba, Japan.

Copyright 2011 by the American Geophysical Union.
0148-0227/11/2010JD015360

Table 1. Chemistry Climate Models Participating in CCMVal and *WMO* [2011]

CCM Name	Atmospheric GCM	Reference	Horizontal Resolution of Advection, Levels, and Top
AMTRAC3	AM3	*Austin and Wilson* [2006]	~200 km, L48, 0.017 hPa
CAM3.5	CAM	*Lamarque et al.* [2008]	1.9° × 2.5°, L26, 3.5 hPa
CCSRNIES[a]	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[a,b]	ECHAM4	*Stenke et al.* [2008, 2009]	T30L39, 10 hPa
EMAC v1.6[a,b]	ECHAM5	*Jöckel et al.* [2006]	T42L90, 0.01 hPa
GEOSCCM v2	GEOS5	*Pawson et al.* [2008]	2° × 2.5°, L72, 0.015 hPa
LMDZrepro	LMDz	*Jourdain et al.* [2008]	2.5° × 3.75°, L50, 0.07 hPa
MRI	MJ98	*Shibata and Deushi* [2008a, 2008b]	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[b]	HadAM3 L64	*Austin and Butchart* [2003]	2.5° × 3.75°, L64, 0.01 hPa
UMSLIMCAT	HadAM3 L64	*Tian and Chipperfield* [2005]	2.5° × 3.75°, L64, 0.01 hPa
UMUKCA-METO	HadGEM-A	*Morgenstern et al.* [2009]	2.5° × 3.75°, L60, 84 km
UMUKCA-UCAM	HadGEM-A	*Morgenstern et al.* [2009]	2.5° × 3.75°, L60, 84 km
WACCM	CAM	*Garcia et al.* [2007]	1.9° × 2.5°, L66, 6 × 10^{-6} hPa

[a]No mean age output submitted.
[b]No REF-B2 (future scenario) submitted.

where P is O_3 photochemical production and the L terms are chemical loss processes due to different radical families [*Stolarski and Douglass*, 1985, and references therein]. In the lower stratosphere (LS), 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 prediction of O_3.

[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 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]. This presents a problem with regard to credibility of prediction. If the relative fractional losses are misrepresented in the present, the response of each term to the changing halocarbon and greenhouse gas (GHG) emissions and Brewer-Dobson circulation in the future will further change the relative size of each loss term, making it improbable that they will correctly simulate future O_3.

[4] *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 1990s with the availability of multiyear, near-global stratospheric trace gas data sets such as those from the Upper Atmosphere Research Satellite (UARS) instruments. This approach is physically based and 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.

[5] *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 that participated in the WMO 2006 assessment [*Eyring et al.*, 2007] to interpret model predictions of future O_3 levels. They found that the four CCMs showing the best agreement with a wide range of observations had a smaller variation in O_3 projections than did the full suite of 11 CCMs. The Stratospheric Processes and their Role in Climate (SPARC) CCMVal Report [*SPARC CCMVal*, 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*, 2011] were evaluated in this report and are listed in Table 1. Details of the components of these models and the reference simulations performed is provided by *Morgenstern et al.* [2010].

[6] The transport evaluation in the SCR concluded that tropical ascent (i.e., circulation) and quasi-horizontal irreversible mixing of midlatitude air into the tropics (i.e., recirculation) were two fundamental processes that strongly affect the distributions of ozone and ozone-depleting substances. The SCR concluded that at least half of the par-

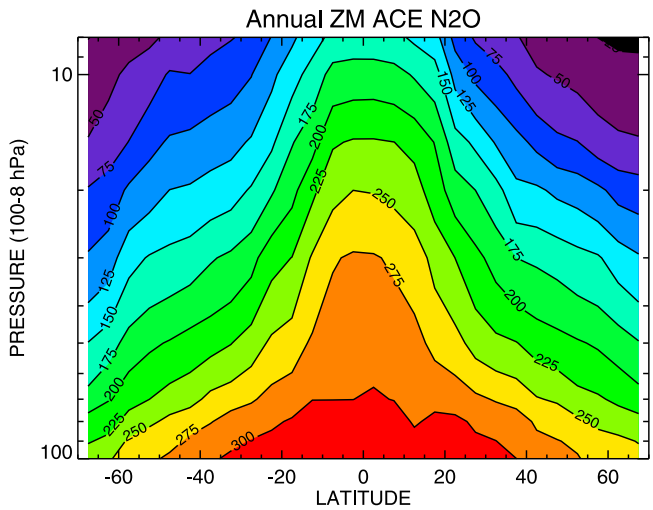

Figure 1. Annual zonal mean N_2O (in ppb) calculated from 6 years of an ACE climatology, 2004–2009.

ticipating CCMs had significant issues with lower stratosphere transport (i.e., circulation, mixing, or both). A useful transport metric developed in the SCR was the average mean age (AMA), which measured how well a model agreed with the available mean age observations covering 7 latitude and altitude ranges. Mean age depends on both circulation (ascent) and mixing (recirculation) and on the balance between them as a function of altitude. Unfortunately, mean age has only been measured at a few latitudes at pressures other than 50 hPa (~20 km), so the mean age data set necessary to fully constrain model LS transport does not exist.

[7] In the lower stratosphere, mean age is generally less than 4 years and is strongly correlated with N_2O. The correlation between simulated N_2O and mean age was described by *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 longer residence time in the stratosphere is correlated with increased maximum altitude for a parcel [*Hall*, 2000]. The longer a parcel is in the stratosphere, the higher the maximum altitude it can attain and the greater the N_2O loss. *Hall et al.* [1999] concluded that long-lived trace gas distributions depend strongly on the quality of a model's mean age, which is controlled by transport. They also concluded that inaccuracies in model transport have a significant impact on the simulated LS chemical composition.

[8] In this paper we present a simple transport diagnostic based on the N_2O/mean age compact correlation and apply this diagnostic, along with chemical diagnostics from the SCR, to explain some of the variation in model projections of future O_3. We first show how the observed compact correlation between N_2O and mean age allows us to evaluate model transport in the extrapolar lower stratosphere (i.e., 60°S–60°N, 150–30 hPa). This diagnostic is based on a well-measured species (N_2O) and is simple to use and interpret. The N_2O/mean age analysis is applied to 15 CCMs and is interpreted in terms of the models' representation of tropical ascent and recirculation. Comparison of these results with O_3 profile data investigates the physical link between this transport diagnosis and model O_3 simulations in the tropical and midlatitude lower stratosphere. The N_2O/mean age diagnostic complements the suite of transport diagnostics discussed in the SCR.

[9] The second part of the paper examines predictions of the return-to-1980 ozone columns for the 15 CCMs simulating the 21st century. Using the transport diagnostics presented here along with some of the SCR chemistry evaluations, we show that several transport and chemistry problems, by themselves, explain at least half of the range of model-predicted return dates for global (60°S–60°N) and October Antarctic ozone columns. We conclude that the current range of predicted return dates is unnecessarily large due to the modeling deficiencies identified here. While there are significant uncertainties in the return dates due to the unknown levels of future halocarbon and GHG emissions [*Charlton-Perez et al.*, 2010], the use of transport and chemistry diagnostics to identify models with credible LS transport and photochemistry can reduce the uncertainty in return dates for a given scenario caused by unrealistic representation of some of the important processes controlling stratospheric ozone.

2. N_2O and Mean Age Observations

[10] In this section we review the observations of N_2O and mean age that are used to develop empirical constraints on lower stratospheric transport. A monthly N_2O climatology derived from observations of the Atmospheric Chemistry Experiment (ACE) satellite instrument onboard SCISAT-1 (2004–2009) (A. Jones, personal communication, 2010) is used to calculate annual mean stratospheric N_2O. The use of the 6 year ACE data mean reduces the effect of the quasi-biennial oscillation on LS N_2O distributions. The latitudinal coverage of ACE measurements varies seasonally, and latitudes poleward of 70° are sampled infrequently. We therefore thus use an annual mean climatology restricted to 68°S–68°N, 150 hPa and above. Figure 1 shows zonal annual mean ACE N_2O determined from the Jones climatology. The Aura Microwave Limb Sounder (MLS) also measured N_2O during this time period [*Livesey et al.*, 2007]. Where these data sets have spatial overlap in the LS, 68°S–68°N and 100–20 hPa, the annual mean 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 to 100 hPa, with MLS showing low bias at pressures greater than 32 hPa.

[11] 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, personal communication, 2010). The average of the OMS and Engel profiles (red) and its 1σ uncertainty (yellow shading) are used in the analysis. The mean midlatitude mean age and uncertainty profiles were derived by Engel (pers. comm.) from CO_2 and SF_6 data from 27 balloon flights occurring mostly between May and October, 32°N–51°N, spanning

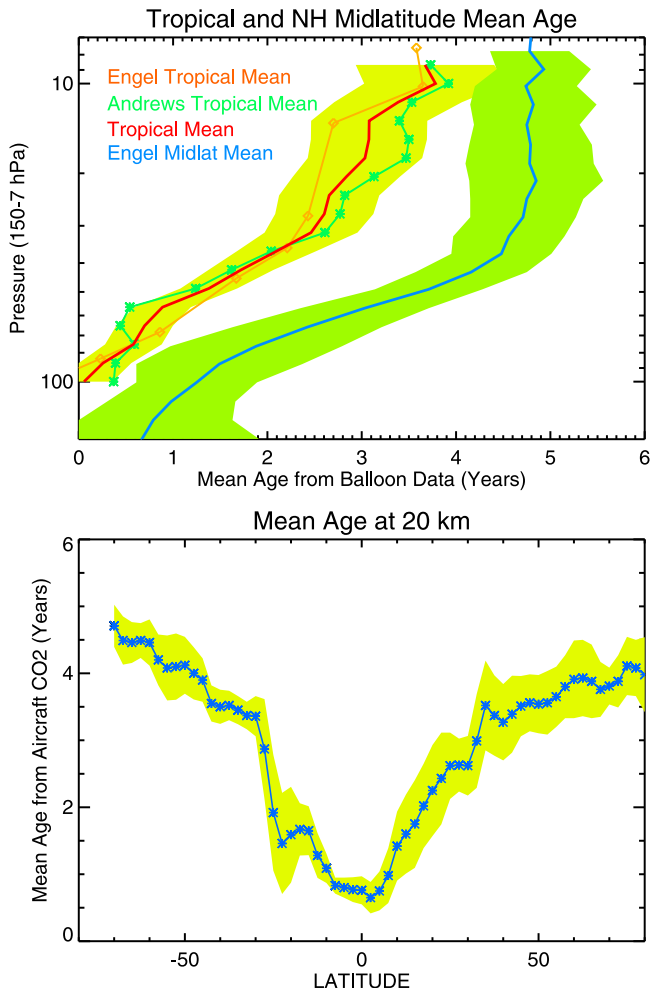

Figure 2. CO_2 and SF_6-derived mean age observations used in this study. (top) Tropical and midlatitude (32°N–51°N) mean profiles and their 1σ uncertainties [*Andrews et al.*, 2001, *Engel et al.* 2009]. The midlatitude mean was provided by A. Engels (personal communication, 2010) and is an average of 27 profiles. (bottom) The *Andrews et al.* [2001] aircraft CO_2 mean age and 1σ uncertainty.

30 years [*Engel et al.*, 2009]. Figure 2 (bottom) shows the 50 hPa (~20 km) mean age as a function of latitude [*Andrews et al.*, 2001]. It was derived from dozens of aircraft flights that sampled during most months, although there were no southern hemisphere flights in summer. These data sets represent our best estimate of the annual averaged mean age.

3. Relationship Between N_2O and Mean Age

[12] Although mean age observations have less spatial coverage than 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 shown in Figure 2. Mean age of less than 4.5 years and N_2O greater than 150 ppb form a compact and nearly linear relationship. Where mean age is >4.5 years, age and N_2O lose correlation. *Hall et al.* [1999] and *Schoeberl et al.* [2000] have explained this as a result of increasing N_2O loss (photochemical exposure) as air ages. The N_2O/mean age behavior is analogous to the compact correlation observed in the LS for N_2O and Cl_y [*Plumb and Ko*, 1992; *Schauffler et al.*, 2003] that is a consequence of slope equilibrium. Slope equilibrium, as originally explained by *Mahlman et al.* [1986], creates a compact relationship between two long-lived constituents due to the mixing on quasi-horizontal surfaces being much faster than time scales for chemical loss or transport through the surfaces. For mean ages ≤4.5 years, tropical (red) and extratropical (blue and green) points show the same compact relationship, suggesting that for the extrapolar lower stratosphere where N_2O > 150 ppb (i.e., where N_2O and mean age are in slope equilibrium), mean age can be estimated from N_2O.

[13] All collocated mean age and N_2O > 150 ppb data were fit to a line. The fitted line is shown in Figure 3 in blue, where the yellow shading represents the approximate 1σ uncertainty range for the observations, ~5% uncertainty for N_2O and ~25% for mean age. Figure 4 was created using the fit to the observed compact N_2O/mean age correlation to estimate mean age for values of N_2O over the range 160–315 ppb. The heavy yellow line marks the 160 ppb N_2O contour, indicating the upper limit of the LS domain in which the compact correlation is expected to occur. Note that this relationship cannot be determined in polar regions due to the latitude range of the N_2O climatology (68°S–68°N).

[14] Nearly 20 years ago, *Plumb and Ko* [1992] suggested that the compact correlation observed between long-lived tracers could be applied to models to evaluate their balance between the large-scale circulation and horizontal mixing.

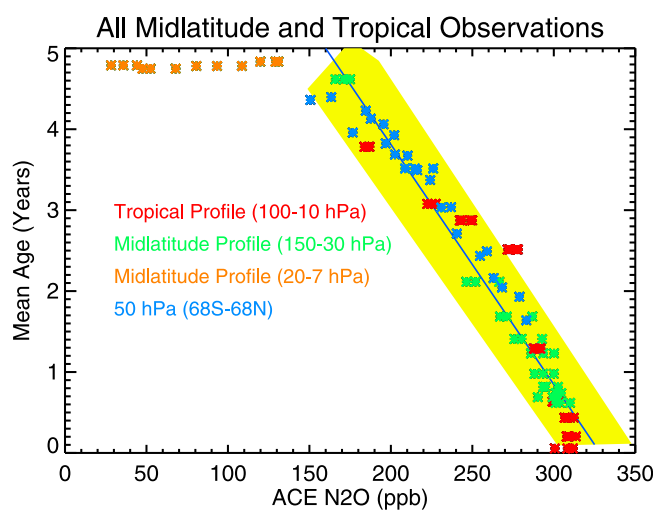

Figure 3. Scatterplot of all mean age observations and their collocated climatological ACE N_2O mixing ratios. The points with N_2O > 150 ppb have been fit to a line. The yellow shading shows the range of uncertainty for the mean age and N_2O observations with respect to the fitted line. Blue points are the 20 km (~50 hPa) aircraft mean age measurements, green points are from the midlatitude mean age profile below 20 hPa, and red points are from the average tropical mean age 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).

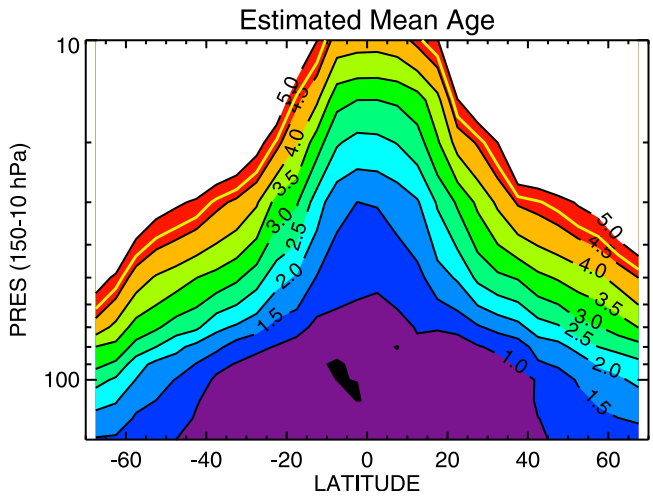

Figure 4. Estimated mean age calculated from a linear fit to the mean age and N_2O points (where N_2O > 150 ppb) shown in Figure 3. The heavy yellow line indicates the upper limit of the domain of the mean age/N_2O compact relationship (~160 ppb).

In section 4 we apply the observationally derived compact correlation between N_2O and mean age to evaluate CCM transport.

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

[15] We now examine the correlation between N_2O and mean age in the CCMs that participated in the 2010 WMO ozone assessment 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 is provided by *Morgenstern et al.* [2010]. This analysis 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) to compute annual means. This period was chosen because it reflects the present-day circulation, as do the observations; we average over 10 years to reduce any bias caused by the presence of a QBO in a model. Three of the participating CCMs did not provide mean age output and thus could not be evaluated here (CCSRNIES, E39CA, and EMAC).

[16] Figure 5 compares the observed global N_2O/mean age relationship with 15 CCMs for the LS (150–50 hPa, 65°S–65°N, and up to 30 hPa from 40°S–40°N); the observations where the compact correlation is observed and the fitted line and uncertainties from Figure 3 are shown in black with yellow shading. The interpretation may be 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 first and second rows of Figure 5 (AMTRAC3, CNRM-ACM, Niwa-SOCOL, SOCOL, and UMETRAC) show maximum mean age less than observed while showing an N_2O range about the same as observed. This relationship suggests a fast circulation that may have good horizontal mixing. 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 fifth row of Figure 5 (LMDZ and the UMUKCA models) shows model curves that are steeper than observed and a maximum mean age that is 1–3 years older than observed. For LMDZ, where the range of N_2O values is about the same as observed, the high mean age suggests a slow circulation. The UMUKCA models have both high mean age and low N_2O. This behavior may be the result of a slow circulation, but may also be caused by too much recirculation (mixing) between the tropics and midlatitudes. UMSLIMCAT and the models in the third and fourth rows of Figure 5 (CAM3.5, CMAM, MRI, WACCM, ULAQ, and GEOS) have curves that agree well with the observed slope and the range of N_2O and mean ages.

[17] While all models show a compact and nearly linear relationship throughout this LS domain, CNRM-ACM has a much flatter slope than the observations. Figure 3 showed a large change in slope for points above 30 hPa in the midlatitudes. The nearly flat slope found in CNRM-ACM may indicate that parcels found near the top of this domain (30 hPa) have already spent significant time at high altitudes where N_2O is destroyed (10 hPa or higher). N_2O loss rates below 30 hPa are very small, resulting in an N_2O lifetime $\gg$1 year in the domain considered here.

[18] The processes controlling the slope and range of the compact correlation, tropical ascent and mixing, can be better assessed by examining N_2O and mean age in the tropics. The fourth column of Figures 6a–6c shows the tropical subset of N_2O/mean age points from Figure 5 (10°S–10°N, 100–30 hPa). The first, second, and third columns of Figures 6a–6c 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). The quantity in each panel is the difference between the observed and simulated value; thus a profile centered on zero indicates exact agreement with the observations. Tropical mean age (first column of Figures 6a–6c) 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. The second and third columns of Figures 6a–6c identify how ascent and horizontal mixing individually contribute to the overall tropical transport seen in the first column. These comparisons may reveal whether a problem lies with circulation or mixing, or both. The mixing evaluated here is cumulative (i.e., total mixing that has occurred during ascent) and relative to the model's ascent rate; this differs from the analysis in the SCR that diagnosed mixing as a function of pressure. The horizontal mean age gradient is calculated as the difference between the mean ages averaged over 35°N–50°N and 10°S–10°N. For most models, their tropical residual mean vertical velocity, w^*, agrees well with the age gradient derived quantity (SCR, chapter 5) except where noted.

[19] In the limit of low vertical diffusion, the horizontal age gradient profile (second column in Figures 6a–6c) is an empirical measure of ascent rate because it reflects the transit time through Brewer-Dobson cell [*Neu and Plumb*, 1999; SCR]. It is independent of horizontal mixing across

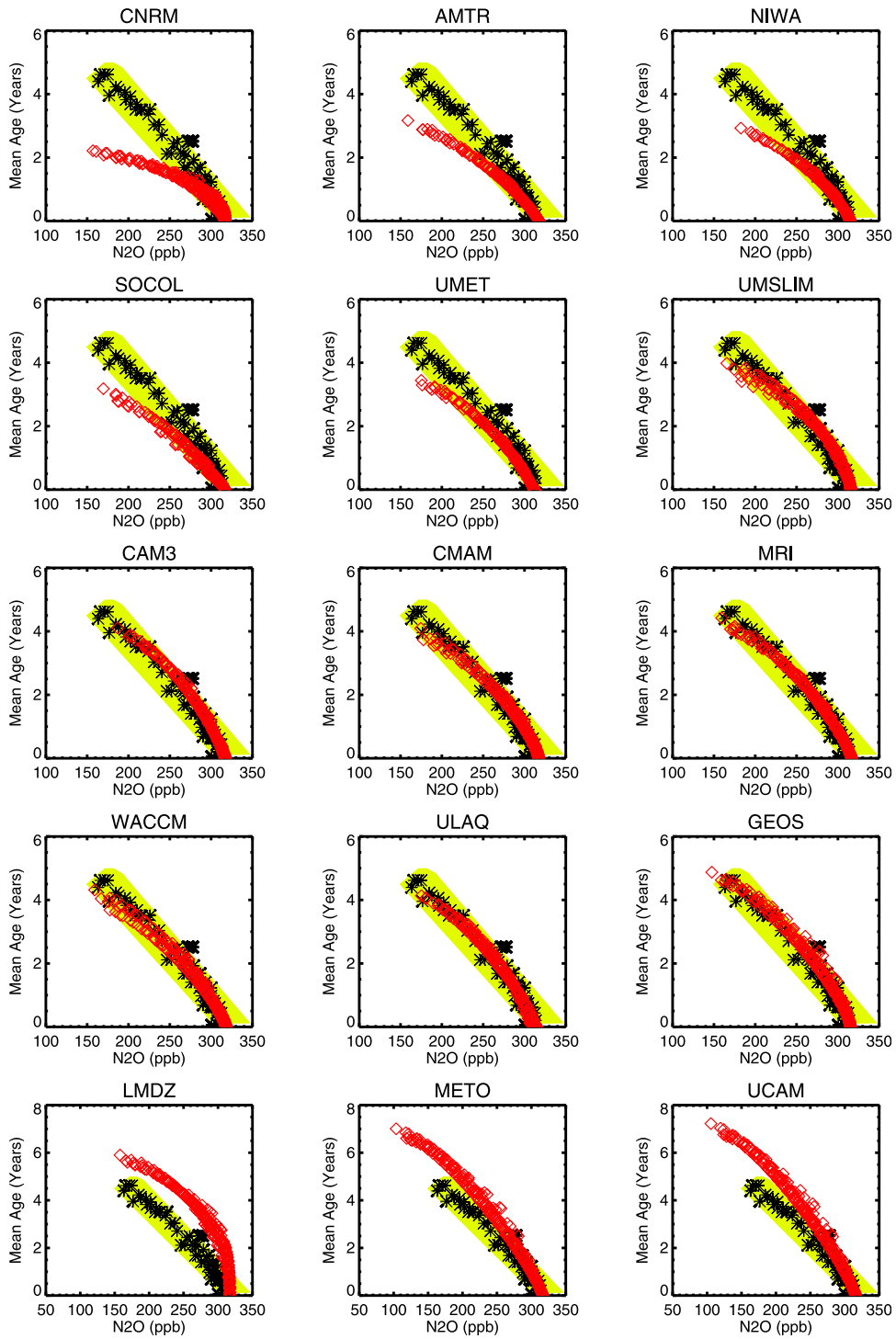

Figure 5. N_2O and mean age relationship in 15 CCMs (red) from the same domain as the compact correlation. Observations are shown in black. The yellow shaded area represents the uncertainty of the fitted line (from Figure 3). Model curves that fall off faster (slower) than observed indicate a fast (slow) circulation. Models results are sorted by fastest circulations at the top, slowest circulations at the bottom. Results are from the REF-B1 (1960–2006).

the subtropics because tropical and midlatitude mean ages will be affected equally by the mixing. (Large vertical diffusion acts to decrease mean age everywhere, which biases the ascent rate high.) 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 (third column of Figures 6a–6c) is a measure of the cumulative degree of horizontal mixing into the tropics. Below 30 hPa tropical N_2O has essentially no loss, so ascent alone has no effect on its mixing ratio. The decreasing

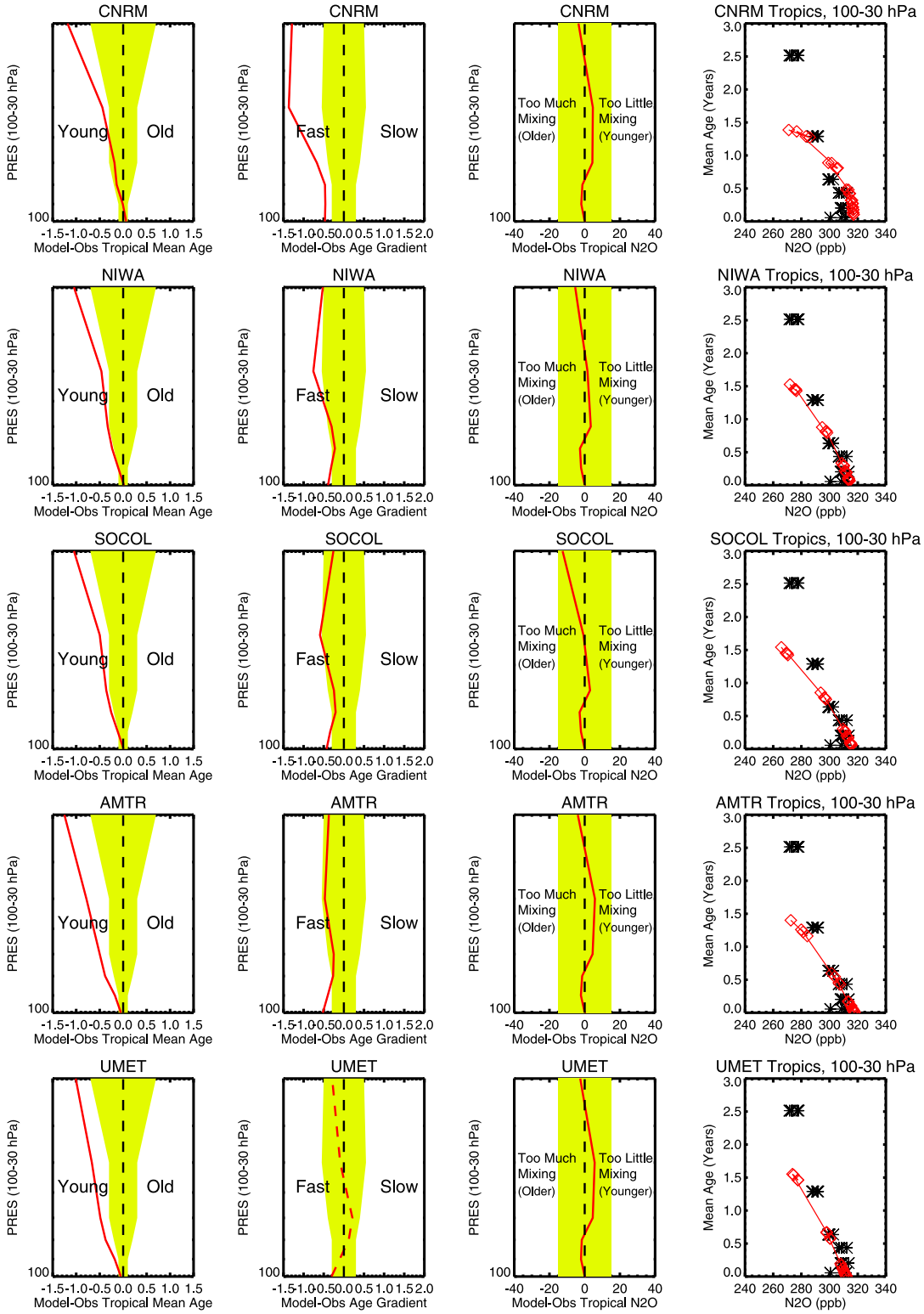

Figure 6a. Diagnostic plots evaluating tropical ascent and tropical-midlatitude (horizontal) mixing in CCMs. Results are from the REF-B1 simulation (1960–2006). The first, second, and third columns show the difference between model and observed profiles of tropical mean age (years), the horizontal mean age gradient (years), and tropical N_2O (ppb), respectively. The yellow shading indicates the $\pm 1\sigma$ uncertainty in the observations. The fourth column compares the simulated (red) and observed (black) mean age/N_2O relationship in the tropical LS, 100–30 hPa. A dashed red line in the second and third columns indicates an issue that complicates the interpretation of that diagnostic; see text. These CCMs have the youngest ages of the 15 CCMs.

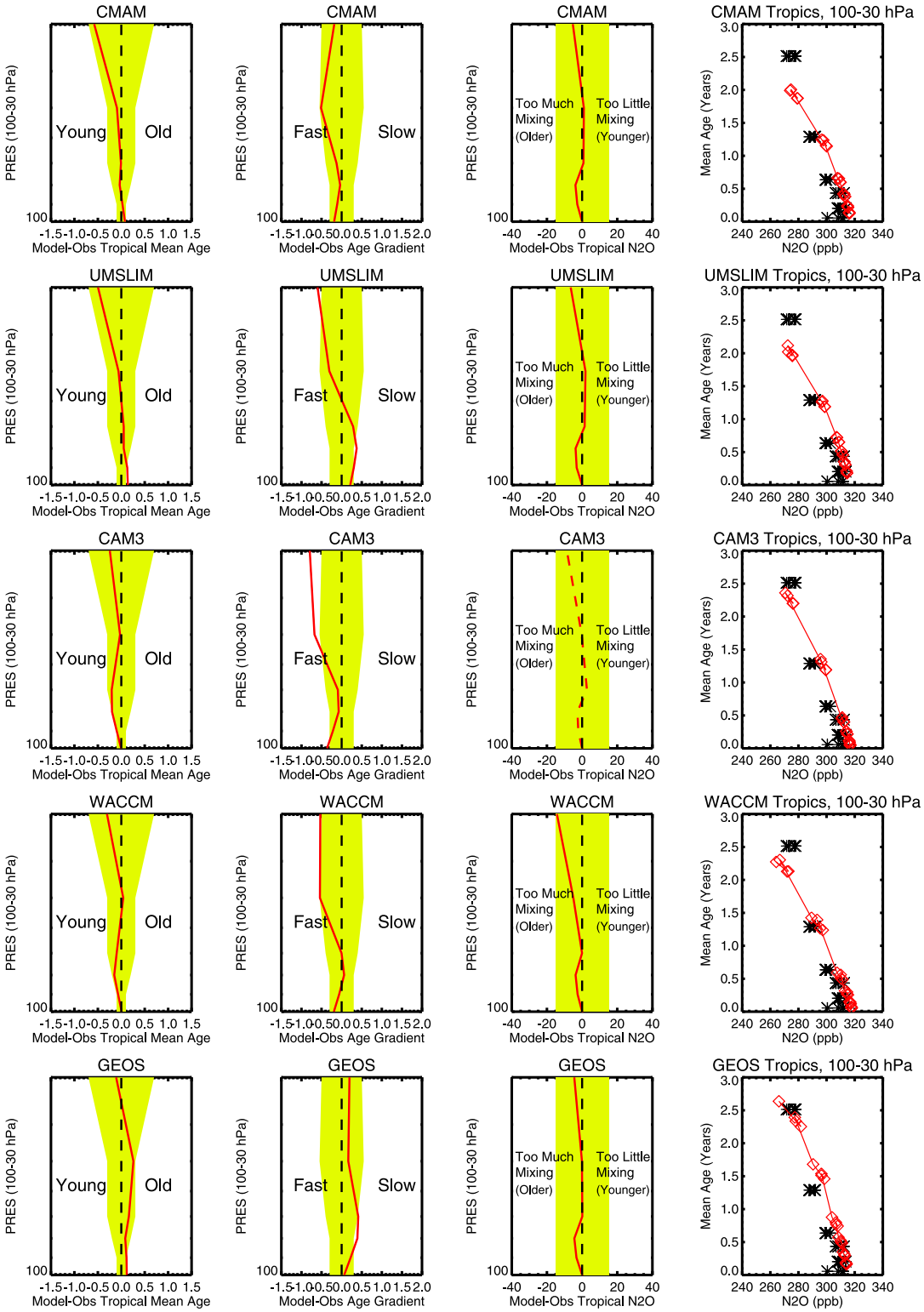

Figure 6b. Same as Figure 6a but for the 5 CCMs with the most realistic tropical circulation and mixing.

mixing ratio with height is due only to the cumulative effects of horizontal mixing as the air ascends. Model N_2O that is higher (lower) than observed can be interpreted as too little (much) mixing of older air into the tropics. The models' tropical N_2O profiles are normalized to the ACE data 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

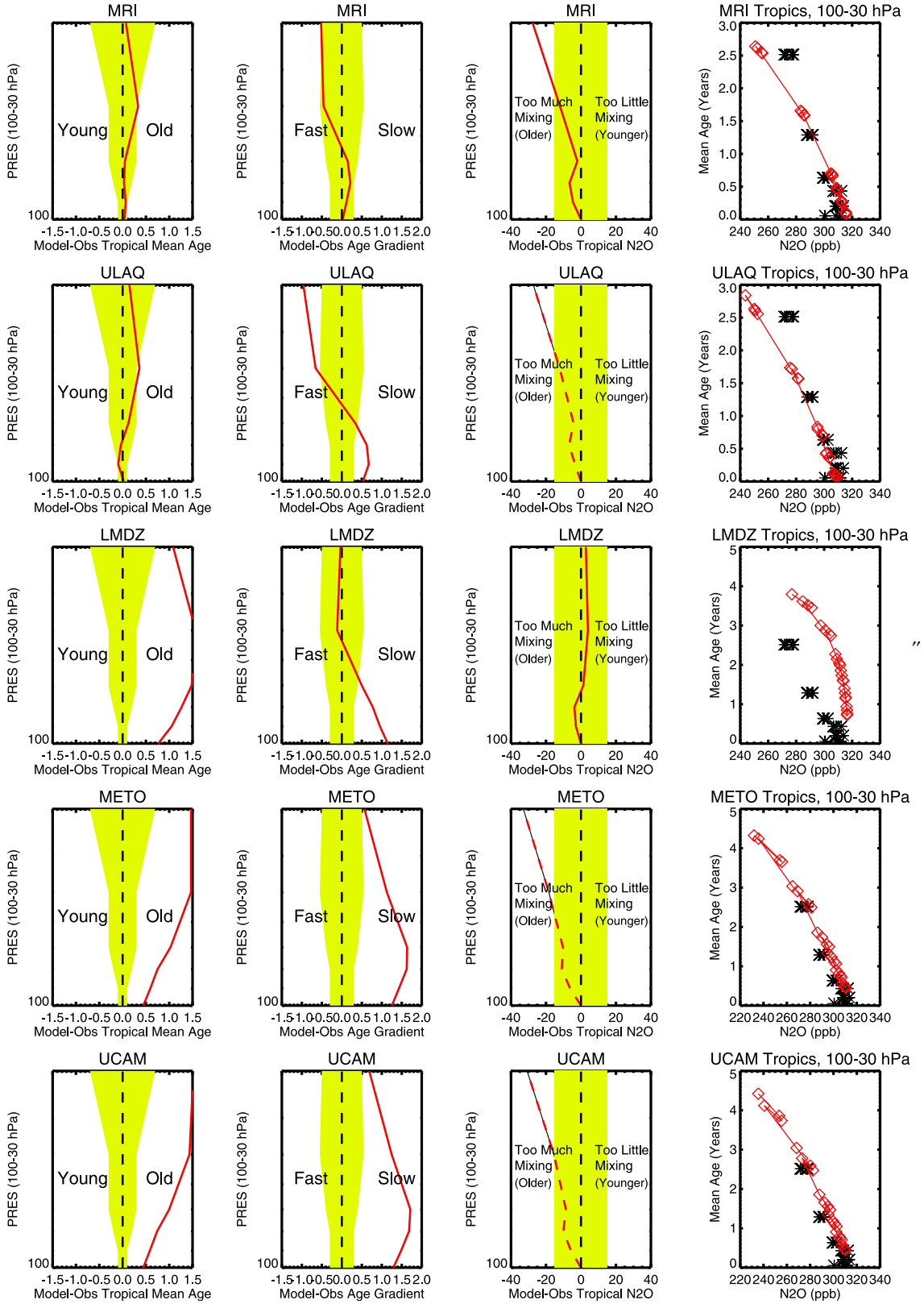

Figure 6c. Same as Figure 6a but for the 5 CCMs with slow circulations or excessive horizontal mixing. Note the dashed line in the third column for ULAQ, METO, and UCAM.

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 biased low, this would imply that mixing was actually too weak. Most models' midlatitude profiles agreed reasonably well with observations; those with bias are indicated by a dashed red line in the third column of Figures 6a–6c.

[20] All models in Figure 6a are too young in the tropics, particularly above 70 hPa (first column). The second column in Figure 6a shows that CNRM-ACM, Niwa-SOCOL, SOCOL, and AMTRAC3 have fast ascent rates with about the right amount of horizontal mixing relative to their ascent rates, although SOCOL borders on too much mixing near 30 hPa in spite of young tropical mean age there. All of these models have excellent agreement with ACE N_2O profiles at 40°N. Note that if ascent rates decreased in these models without any change in horizontal mixing, their resulting N_2O profiles would indicate too much mixing. CNRM-ACM has the youngest mean age and fastest ascent of any of the CCMs. CNRM-ACM, Niwa-SOCOL, and SOCOL have residual mean velocities that agree with the observationally derived w^* of *Schoeberl et al.* [2008] but not with the age gradient-derived ascent rate. The disagreement could be caused by excess vertical diffusion in these models that leads to a high bias for the age-derived ascent rate. UMETRAC has young tropical mean age yet appears to have a good ascent rate and good horizontal mixing. Its midlatitude N_2O is not biased. The w^* was not available for UMETRAC or AMTRAC3, so the cause for UMETRAC's young mean age is unknown. The SCR was also unable to reconcile young mean age in UMETRAC with its tropical transport diagnostics.

[21] The models in Figure 6b show the best agreement overall with all tropical 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. WACCM shows borderline fast ascent and strong mixing at 50 hPa and above, with no bias in its midlatitude N_2O. These effects combine to give WACCM a good tropical mean age profile. GEOSCCM ascent is a little slow below 50 hPa but otherwise agrees very closely with the diagnostics. All models in Figure 6b have very good agreement with ACE N_2O profiles at 40°N except for CAM3.5.

[22] The CAM3.5 model in the third row of Figure 6b has an age gradient that indicates fast ascent while tropical mean age and mixing appear correct. This is explained in part by the high bias in CAM3.5 midlatitude N_2O. While the mixing diagnostic (third column of Figure 6b) suggests mixing is about right (dashed line), it must actually be too strong because its midlatitude N_2O is too high: more mixing must occur in order to sufficiently lower the tropical mixing ratio. Thus, tropical mean age appears correct because the fast ascent, which lowers mean age, is compensated by excessive mixing, which increases age. However, this model has w^* that agrees closely with observations. The inconsistency between the age gradient diagnostic and w^* may be a result of the CAM3.5 low lid, giving it a shorter Brewer-Dobson circulation path and hence a faster age-derived transit time. The CAM3.5 w^* is the same as WACCM and it uses the same atmospheric GCM, but WACCM has a much higher lid than CAM3.5 (5×10^{-6} hPa versus 2.2 hPa).

[23] The models in Figure 6c show a mixture of tropical transport problems. Of these models, only LMDZ and MRI have good agreement with ACE N_2O profiles at 40°N. MRI produces a good tropical mean age profile at all levels but the ascent and mixing diagnostics indicate this is fortuitous. MRI ascent and mixing look good from 100 to 70 hPa, but at 50 hPa and above mixing is too strong while ascent is too fast. The results for ULAQ are similar above 50 hPa, where fast ascent combines with too much mixing to produce good tropical mean age. Below 50 hPa, ULAQ has slow ascent, yet apparently good mean age and horizontal mixing. ULAQ has low midlatitude N_2O from 70 to 100 hPa, so the good performance on the mixing diagnostic actually means that there is too little mixing. At 30 hPa the ULAQ midlatitude N_2O is biased high, thus mixing is even stronger than indicated by this diagnostic. The ULAQ w^* is considerably slower than the age gradient ascent rate from 100 to 30 hPa, suggesting that its age derived ascent rate may be affected by vertical diffusion. LMDZ has very slow ascent from 100 to 70 hPa. By 50 hPa, the ascent rate appears correct and mixing is good at all levels. The UMUKCA models show very slow ascent, especially below 50 hPa. However, both models have very low midlatitude N_2O. 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 contributes to much older than observed age; the mixing diagnostic is not useful for these models. The results presented here are generally consistent with the transport conclusions in the SCR.

5. Effects of Transport and Chemistry on CCM Ozone Simulations

5.1. Transport Effects on O_3 in the Lower Stratosphere

[24] In this section we examine models' LS O_3 profiles to see if their agreement with observations can be related to transport characteristics. We include all models whose REF-B1 (1960–2006) transport could be evaluated or that ran the REF-B2 (1960–2100) scenario. *Oman et al.* [2010] report that among 14 CCMs, differences in the lower stratospheric O_3 are responsible for most of the column differences in the evolution of O_3 in the 21st century. Because O_3 is strongly influenced by gas phase chemistry and temperature above 30 hPa and by heterogeneous chemistry in the polar LS, the best place to look for a link between ozone and transport is the extrapolar LS.

[25] Figure 7 shows the difference between annual mean O_3 profiles from Aura MLS and 16 CCMs for the SH and NH midlatitudes and the tropics. Although the CCSRNIES model did not submit the necessary mean age output for the transport diagnostic, its O_3 profiles are included here. The model annual means are a 10 year average calculated from the last 10 years of the REF-B1 run, usually 1997–2006. An annual zonal mean was calculated from 4 years (2005–2008) 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. The shaded area about the zero

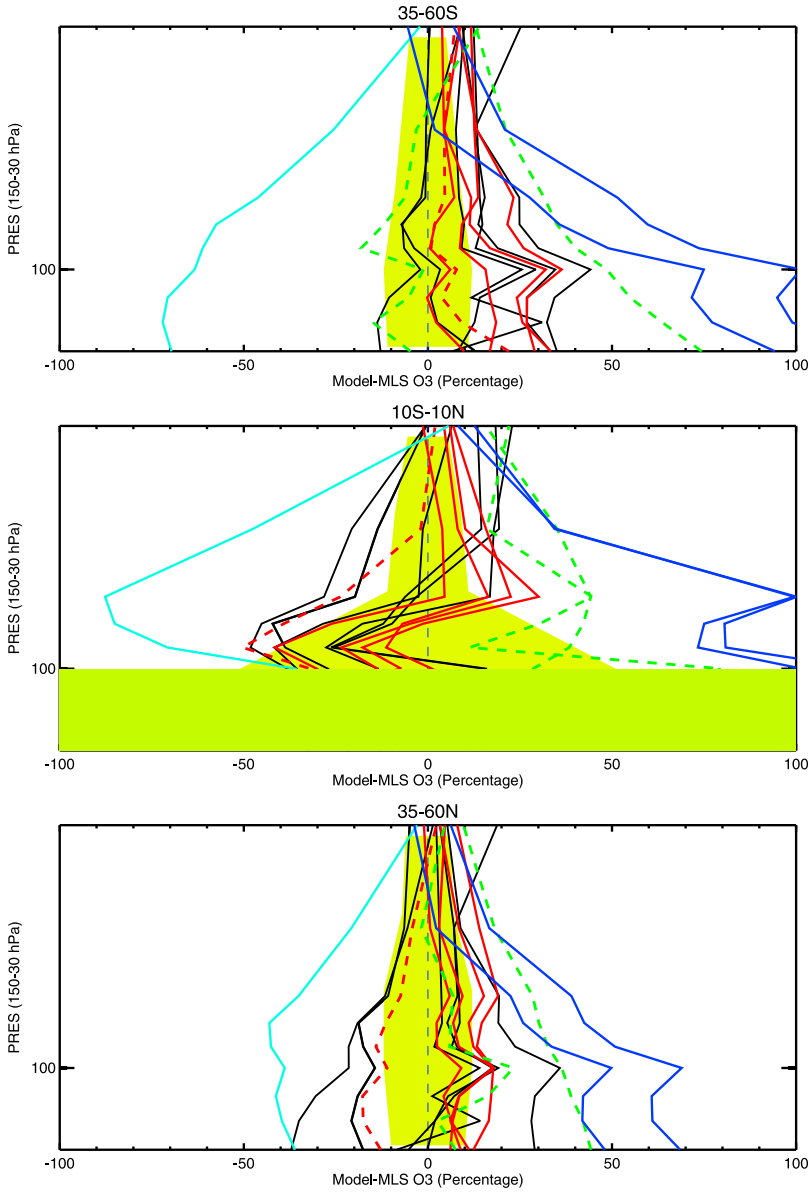

Figure 7. Differences between models and observed O_3 profiles for the SH and NH midlatitudes and the tropics as a percentage of the MLS annual mean values. Model results are from the REF-B1 simulation (1960–2006). The models with the most realistic tropical transport processes (Figure 6b) are shown in red. 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 too much mixing. Yellow shading indicates MLS O_3 2σ systematic uncertainty [*Livesey et al.*, 2007]. Most of the CCMs do not have a tropospheric chemistry scheme so tropical O_3 below 100 hPa should not be compared.

line indicates MLS systematic uncertainties. 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 Figure 6b, are plotted in red. (CAM3.5 is shown as dashed red, indicating lower confidence in its transport.) Overall, these models show closer agreement with MLS O_3 than most models, and the spread among them, except for CAM3.5, is small. The five models shown in blue or green have the greatest LS transport problems, either very fast ascent (CNRM-ACM, light blue), very slow ascent (UMUKCA-METO and UMUKCA-UCAM, dark blue), or combined ascent and mixing problems (MRI and ULAQ, green dashed). 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

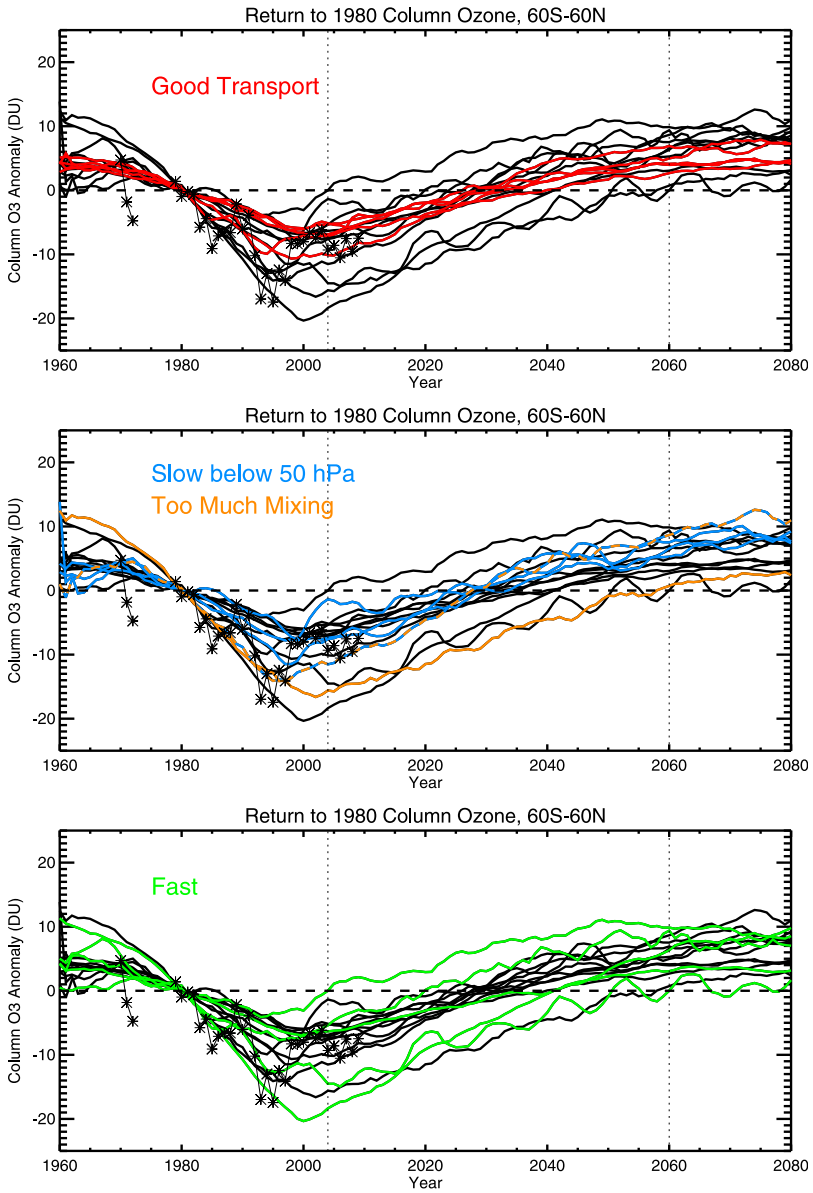

Figure 8. Annual mean 60°S–60°N column O_3 anomalies with respect to 1980 for 15 CCMs. Results from the 15 CCMs that integrated the REF-B2 simulation (1960–2100) are shown in each plot. Observations from the Merged Ozone Data Set [*Stolarski and Frith*, 2006] are shown with black asterisks. (top) The 5 CCMs (red) with the most realistic transport (CAM3.5, CMAM, GEOSCCM, UMSLIMCAT, and WACCM). (middle) The 3 CCMs in blue have been diagnosed with slow ascent (LMDZ, UMUKCA-METO, and UMIUKA-UCAM). The CCM in orange is diagnosed with fast ascent and too much mixing (MRI). The CCM with a dashed blue/orange line has slow LS ascent and too much mixing (ULAQ). (bottom) The 5 CCMs (green) with a fast circulation (AMTRAC3, CCSRNIES, CNRM-ACM, Niwa-SOCOL, and SOCOL).

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.

[26] The six models shown in black (AMTRAC3, CCSRNIES, LMDZ, Niwa-SOCOL, SOCOL, and UMETRAC) have identifiable transport deficiencies yet they often compare as well with observations from 100 to 50 hPa as do the models with realistic transport. (The SCR transport evaluation characterized CCRSNIES as having a slightly fast ascent rate.) These models have less serious transport deficiencies and smaller disagreements with the observed O_3 than the blue and green line models discussed above. The 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 or temperature problem. Figure 7 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. The ability of models with identifiable problems to reasonably simulate O_3 profiles serves as a reminder that O_3 itself is not a good diagnostic quantity.

5.2. Return-to-1980 Column Ozone 60°S–60°N

[27] Roughly half of the ozone column resides in the LS, but to what degree can transport diagnostics explain the entire model column or the return-to-1980 date? Figure 8 shows the annual global mean (60°S–60°N) of the column O_3 anomaly (with respect to 1980) from the REF-B2 (1960–2100) simulation for 15 CCMs and observed annual mean column O_3 anomaly from the Merged Ozone Data Set [*Stolarski and Frith*, 2006]. Most CCMs do not explicitly calculate tropospheric ozone, but because comparisons are made with column anomalies this should not be a source of bias. Differences in approaches to tropospheric O_3 treatment, as well as differences between in the implementation of the REF-B2 forcings, such as SST projections, will have some impact on the models' O_3 predictions; details of these differences are provided by *Morgenstern et al.* [2010]. E39CA, EMAC, and UMETRAC did not a submit REF-B2 simulation and are not included in the remainder of the discussion.

[28] Figure 8 groups the return-to-1980 column O_3 predictions for the 15 CCMs by the results of the transport diagnosis in section 4. In Figure 8 (top) the CCMs with the best transport characteristics (in red: CMAM, GEOSCCM, UMSLIMCAT, WACCM, and CAM3.5) produce a narrow range of predicted return dates (2026–2040) compared to the full range of predicted dates (2004–2060). Figure 8 (middle) shows the four models that demonstrated slow ascent below 50 hPa (in blue: LMDZ, ULAQ, UMUKCA-METO, and UMUKCA-UCAM). They too show a narrow range of return dates, 2023–2035, about 4 years earlier than the models with the best LS transport. The difference in the mean return dates for the two sets of models results is not statistically significant. Two of the slow models (UMUKCA) showed considerable high bias with respect to MLS O_3 in the LS (Figure 7), yet the return dates are not significantly different from the models with good transport and little or no bias. While transport affects LS O_3 distributions, the return date for ozone columns depends on model O_3 sensitivities, especially to temperature and Cl [*Stolarski and Douglass*, 1985; *Oman et al.*, 2010]. The models with strong tropical recirculation (MRI and ULAQ) are shown in orange in Figure 8 (middle).

[29] Figure 8 (bottom) shows that the five CCMs diagnosed with fast circulations (in green: AMTRAC3, CCSRNIES, CNRM-ACM, Niwa-SOCOL, and SOCOL) span the full range of predicted return dates (2004–2060). It is interesting that two models with very similar transport diagnostics, AMTRAC3 and Niwa-SOCOL (e.g., Figure 6a), that also showed reasonable agreement with present-day O_3 profiles produce widely different return dates, 2060 and 2004, respectively. While these CCMs have similar transport diagnostics, clearly they must have a different balance of processes controlling O_3.

5.3. Role of Chemistry

[30] The role of chemistry must be considered in order to understand differences in CCM O_3 predictions. *Oman et al.* [2010] have evaluated the contributions of halogens (Cl_y+5Br_y), NO_y, and temperature to tropical (25°S–25°N) O_3 profiles changes over the 21st century for most of the CCMs evaluated in this study. AMTRAC3 stands out from the other models by having a much greater NO_y contribution to O_3 loss and a much smaller halogen contribution to O_3 increase in the tropical upper stratosphere during the 21st century. AMTRAC3 has just over half the abundance of halogens compared to most other models. While most models show a LS NO_y decrease and upper stratospheric increase over the 21st century, AMTRAC3 shows NO_y increases in both the lower and upper stratosphere, leading to more O_3 loss (relative to other models) and a slower recovery. Low halogens and high NO_y levels in the AMTRAC3 tropics, compared to the other CCMs, explain why AMTRAC3′s predicted return date is much later than other CCMs.

[31] Niwa-SOCOL and SOCOL have the earliest 60°S–60°N return-to-1980 dates of the 15 CCMs, 2004 and 2020. The Niwa-SOCOL REF-B2 simulation was not evaluated by *Oman et al.* [2010], however they did evaluate SOCOL, which has the same chemical mechanism and transport scheme as Niwa-SOCOL [*Morgenstern et al.*, 2010]. *Oman et al.* [2010] show that the halogen contribution to SOCOL's tropical O_3 increase during the 21st century is much larger than AMTRAC3′s and most of the CCMs. Of all the CCMs, SOCOL shows nearly the highest level of halogens in the year 2000 in the tropical stratosphere between 20 and 1 hPa. Levels of ~4 ppb are simulated, which is in excess of the Cl mixing ratio boundary conditions. While *Oman et al.* [2010] show that SOCOL has a tropical O_3 sensitivity to changing halogens that is very similar to other models, SOCOL has a much larger abundance of Cl_y and photochemistry that produces too much ClO for a given Cl_y (chapter, 6, SCR). At the same time, SOCOL has a low NO_x/NO_y ratio and thus a smaller sensitivity to NO_y (chapter 6, SCR). In general, increasing NO_y during the 21st century causes additional O_3 loss and hence slows down O_3 recovery. SOCOL's and implicitly Niwa-SOCOL's early return dates may be a consequence of a larger fraction of their middle and upper stratospheric O_3 losses being due to halogens while the relatively small fraction of NO_y loss does not retard the recovery.

[32] Chapter 6 of the SCR evaluated chemistry in CCMs by looking at radicals and radical precursors, and photolysis rates. The remainder of this section summarizes some of the chemistry issues identified in the SCR that affect O_3 simulations. In the SCR, a photochemical steady state (PSS) model was used to calculate radicals using each model's radical precursor and temperature profiles as input. If the PSS model results match the CCM's radicals, then the CCM is deemed to have the same (i.e., correct) chemical mechanism as the PSS. Niwa-SOCOL did not participate in the PSS comparisons, but it uses the same photolysis and chemistry schemes as SOCOL. Model precursors (e.g., Cl_y, NO_y, O_3) were evaluated based on their correlation with N_2O.

[33] CCSRNIES, CNRM-ACM, MRI, and SOCOL had the most disagreements with observations for the precursors, particularly Cl_y, and MRI and SOCOL had poor agreement with several radicals (e.g., O^1D, HO_x, NO_x, and ClO). MRI greatly overestimates ClO/Cl_y due to a missing loss reaction

for ClO. AMTRAC3 has minor disagreements with the ClO/Cl_y ratio in the middle and upper stratosphere. CCSRNIES, Niwa-SOCOL, SOCOL, and UMUKCA-METO have total chlorine in the upper stratosphere that exceeds 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.) Simulated tropospheric Cl_y was much higher than expected for CCSRNIES, CNRM-ACM, MRI, SOCOL, ULAQ, and UMUKCA-METO. In these models, the excess Cl_y was found to extend into the lowermost stratosphere. This presents a larger problem for CCSRNIES, CNRM-ACM, SOCOL, and ULAQ because they also have LS ClO/Cl_y that is too high, increasing these models' potential for ozone loss for a given level of Cl_y. HO_x is quite low and NO_x/NO_y is quite high in the LS in UMUKCA-METO; UMUKCA-UCAM did not participate in the PSS comparison but uses the same chemistry scheme as UMUKCA-METO and its results are expected to be similar. The PSS comparison identified no major issues with radicals or precursors for CMAM, EMAC, GEOSCCM, LMDZ, UMSLIMCAT, or WACCM.

[34] The SCR photochemical intercomparison ('Photo-Comp') evaluated the accuracy of J values; nine CCMs participated. AMTRAC3, CCSRNIES, and Niwa-SOCOL showed the greatest inaccuracies, while UMSLIMCAT, WACCM, GEOSCCM, and LMDZ had highly accurate J values.

[35] Models with both transport and chemistry problems will have different responses to the changing forcings of the 21st century, i.e., increasing SSTs and the concomitant circulations changes, and changing GHG and halocarbon emissions. All the fast circulation models in Figure 8 (bottom), as well as most of the models in Figure 8 (middle), have transport and Cl chemistry problems. Together they span a 56 year range of return dates. MRI has too much tropical recirculation and significant ClO/Cl_y problems including a missing loss process for ClO. Niwa-SOCOL has a fast circulation, lack of conservation of Cl_y, and inaccuracies in its J values. CCSRNIES, CNRM-ACM and SOCOL have fast circulations, excessive ClO/Cl_y in the LS, and disagreements with radical precursors; CCSRNIES, Niwa-SOCOL, and SOCOL also lack Cl_y conservation. 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. ULAQ and UMUKCA-METO have circulation and/or mixing problems and a problem with high Cl_y in the lowermost stratosphere.

[36] Of the four models with the best LS transport (CMAM, GEOSCCM, UMSLIMCAT, and WACCM), none had any major chemistry problems. CAM3.5 has fairly good transport and minor problems with ClO/Cl_y and NO_x/NO_y. All five of these CCMs, except for UMSLIMCAT, agree closely with column O_3 observations (discussed in section 6). It is worth noting that the LMDZ model performed very well in the photolysis, radical, and precursor evaluations and its only identified transport deficiency is slow ascent in the tropics below 50 hPa. *Oman et al.* [2010] examined partial column O_3 recovery in the tropics in these CCMs and found very small changes to the 500–20 hPa columns over the 21st century; the greatest O_3 recovery in all CCMs occurs outside the tropics. Because much of midlatitude O_3 is transported from the production region in the tropical middle stratosphere, tropical circulation problems below 50 hPa may have only a minor effect on the 60°S–60°N recovery. The LMDZ return date is within the range predicted by the five models with the best LS transport.

5.4. Return to 1980 Column O_3 in the Antarctic

[37] The N_2O/mean age transport diagnostic cannot be directly applied to polar regions because most high latitude N_2O and mean age values there are outside the range of the compact correlation. However, the results of the transport analysis are indirectly relevant because polar descent is linked to tropical ascent through the circulation. In fact, CCMs that performed best on tropical ascent evaluations also performed best on polar descent (SCR, chapter 5). In this section we look at the relationship between vortex Cl_y, Cl chemistry, and column O_3 return dates. Vortex Cl_y is a more sensitive diagnostic of transport than mean age because Cl_y is both time scale and pathway dependent [*Waugh et al.*, 2007]. For example, two models may have the same mean age in the Antarctic vortex, but if one model's circulation takes parcels to higher altitudes where more CFCs are photolyzed, that model will have higher vortex Cl_y. Differences in modeled Antarctic Cl_y have been shown to be related to differences in the return-to-1980 O_3 values [*Eyring et al.*, 2006, 2007]. Activated Cl, i.e., ClO, is required for O_3 loss, so a realistic response to Cl requires both a good simulation of Cl_y (transport) and good photochemistry (ClO/Cl_y). Other factors are important too, such as polar temperatures, PSC distributions, and the presence of a mixing barrier at the edge of the Antarctic vortex.

[38] Figure 9 compares each model's Antarctic return date with its simulated October 2005 mean Cl_y at 80°S. The models shown in red have no significant problem with Cl chemistry or conservation. For this set of 8 models there is a clear, positive correlation between Cl_y in the LS vortex and the return-to-1980 date. This is relationship is generally true for the full set of CCMs not just the ones having correct Cl chemistry, although there is more scatter.

[39] Each of the models in black was identified in the SCR as having high Cl_y and/or high ClO/Cl_y in the upper troposphere/lower stratosphere. A high ClO/Cl_y allows a model to have a later return date in spite of low Cl_y because more ClO is produced for a given Cl_y level. In the case of MRI, the cause of the high ratio is likely the missing ClO loss reaction. UMUKCA-METO has a known problem in the removal of tropospheric HCl that results in excessive stratospheric Cl_y (by about 0.5 ppb) but its ClO/Cl_y in the LS is not high. Understanding this model's return date is more complex because it also has high NO_x/NO_y, low HO_x, and high vortex temperatures. The other five models in black have high ClO/Cl_y in the UT/LS but the cause was not determined in the SCR (chapter 6). This high ratio alone may explain why Niwa-SOCOL, SOCOL, and ULAQ have a later return date than would be expected for their low Cl_y levels. However, CCSRNIES and CNRM-ACM also have a high ratio but do fall on the fitted line, so the ClO/Cl_y problem cannot be the only factor. CNRM-ACM has very young mean age in the Antarctic and incorrect fractional release for CFC-11 and CFC-12 (SCR, chapter 5); this explains why high Cl_y is found in the CNRM-ACM LS

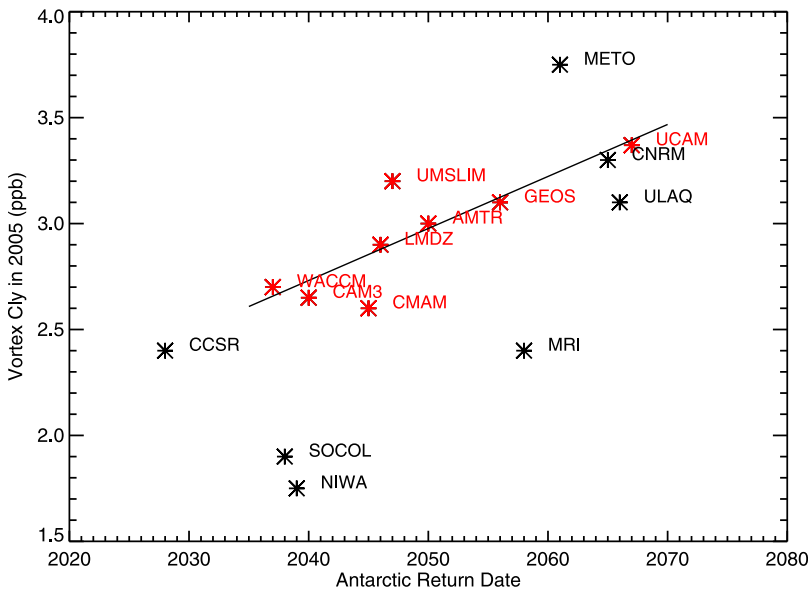

Figure 9. The relationship between model return-to-1980 date for the October Antarctic (65°S–90°S) column O_3 and model Cl_y at 80°S 50 hPa in 2005. Results are from the REF-B2 simulation (1960–2100). Models in black have a problem with Cl conservation (CCSRNIES, Niwa-SOCOL, and SOCOL), tropospheric HCl removal (UMUKCA-METO), or ClO/Cl_y (CNRM-ACM, MRI, and ULAQ). Models in red have no Cl chemistry problems. The linear relationship between vortex Cl_y and Antarctic return date is shown for the eight CCMs that have no Cl problems.

vortex in spite of the transport problems, but the cause for the excessive fractional release is unknown. Fractional release could not be assessed for CCSRNIES, but overall CCSRNIES performed below the multimodel average for precursors and radicals.

[40] Antarctic O_3 recovery is not strictly correlated with vortex Cl_y (i.e., transport) because the amount of Cl_y that a model converts into active Cl is dependent on so many processes. The amount of vortex Cl_y in these CCMs is roughly correlated with their transport circulation: many fast models have low Cl_y (e.g., CCSRNIES, Niwa-SOCOL, SOCOL) while those with slower circulations have high Cl_y (e.g., UMUKCA-METO, UMUKCA-UCAM, ULAQ). However, model problems with Cl photochemistry and conservation prohibit a strong correlation with circulation strength from standing out in Figure 9.

[41] For models that do not have significant chemistry problems, we find a smaller range of predicted Antarctic return dates. Figure 9 shows a range of 39 years (2028–2067) for Antarctic column O_3 return-to-1980 dates for the 15 CCMs. When we consider only the models that do not have Cl chemistry problems (8 models), we find a range of 30 years (2037–2067). Of these 8 models, 3 have been diagnosed with transport problems (UMUKCA-UCAM, LMDZ, and AMTRAC3). The five models that demonstrated the most realistic LS transport and chemistry (CAM3.5, CMAM, GEOSCCM, UMSLIMCAT, and WACCM) produce a 19 year range of return dates, 2037–2056.

6. Discussion and Conclusions: Understanding Model Predictions

[42] Figure 10 shows 60°S–60°N column O_3 and column anomalies for 15 CCMs. In Figure 10 (top) the models whose column O_3 agrees to within 5% of observations (~15 DU) are plotted in blue. Figure 10 (middle) shows column anomalies, indicating when each model returns to its 1980 60°S–60°N mean value. The 10 models that agree best with global mean observations (blue) span the full 54 year 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, although a reduction in range for a randomly selected group of 10 results would be expected 57% of the time. Figure 10 (bottom) shows the same column anomalies where the models with the best LS transport, as determined by this analysis, are plotted in red; the range of predicted recovery dates is 14 years (2026–2040). This reduction is unlikely to result from the reduced sample size (5 CCMs instead of 15). The random selection of 5 results gives a range of 14 years or less only 7% of the time. The 10 models shown in black have identifiable problems with transport, such as fast or slow ascent, or inappropriate tropical isolation. Seven of those 10 models also have a problem with their ClO/Cl_y, and three of those also lack Cl_y conservation.

[43] 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 to gauge model credibility. Similarly, CCMs with good LS transport and chemistry will not necessarily produce good agreement with observed columns for several reasons, including lack of tropospheric chemistry, problems with upper stratospheric temperatures, and representation of heterogeneous chemical processes.

[44] The overall goal of this paper is to understand ozone predictions in CCMs. The initial approach was to evaluate model transport, then determine if transport problems had a

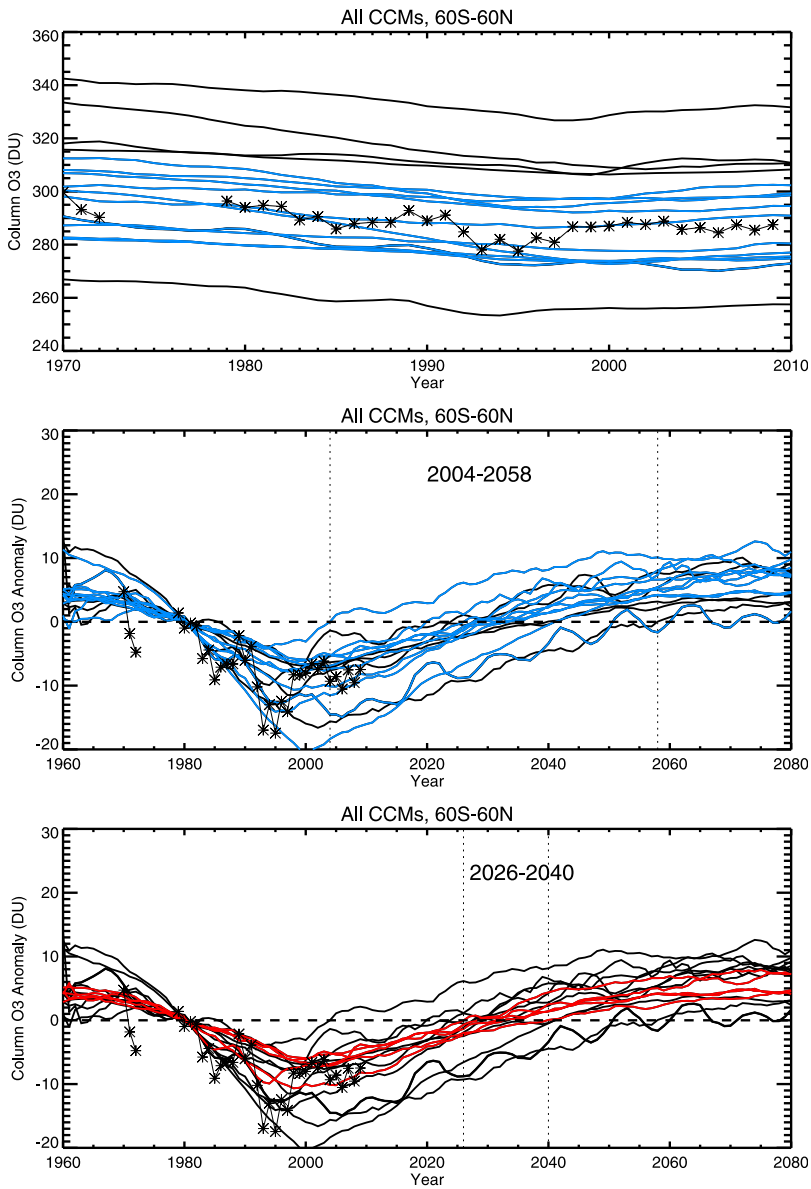

Figure 10. (top) Annual mean 60°S–60°N column O_3 in 15 CCMs from 1970 to the present. Results are from the REF-B2 simulation (1960–2100). Observations are shown with black asterisks in each plot. The models in blue show agreement within 5% of the observations; models in black do not agree within 5% of the observations. (middle) The same CCM output and observations, but plotted as anomalies with respect to 1980, from 1960 to 2080. The dashed vertical lines show the earliest and latest predicted return dates for the models with the best agreement with observations (blue), which is the same 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 models in red, 2026–2040.

noticeable impact on simulated LS O_3. We found that large transport discrepancies (e.g., very fast or slow circulations, or strong mixing) did affect O_3 profiles in the lower stratosphere, while smaller transport problems did not. Later, using some of the SCR chemistry evaluations and results from *Oman et al.* [2010], we showed how O_3 simulations are strongly affected by chemistry problems, particularly those involving the Cl_y family.

[45] Ozone recovery in CCMs is a result of decreasing CFCs that are forced by the mixing ratio boundary conditions used in the scenario. Chemistry and transport evaluations of CCMs provide the means to explain much of the variation in return-to-1980 dates. Antarctic recovery is a function of reactive Cl. If a model's photochemistry is correct, then realistic Cl_x can be produced if LS Cl_y is also realistic, and the Cl_y levels are to a large extent controlled by transport. If there are problems with ClO/Cl_y, then the relationship between transport and Cl_x, and hence return date, will be less clear. For the Antarctic and the 60°S–60°N average, the models with good chemistry and good LS

transport, as identified by close agreement with observations of N_2O and mean age, show a range of return dates that is much less than the range of the full set of 15 CCMs. In the Antarctic the range is reduced by half, from 39 to 19 years, and for the 60°S–60°N mean, the range is reduced by 75%, from 56 to 14 years. Realistic chemistry may have a greater impact than realistic transport on O_3 simulations, but this can't be assessed for this set of CCMs because none was found to have excellent chemistry but poor transport. However, given the importance of Cl_y to Antarctic recovery and the sensitivity of Cl_y to transport, we conclude that both good chemistry and good transport are required for credible predictions.

[46] While there is much uncertainty in future ozone levels due to the uncertainties in future emissions of ozone-depleting substances and greenhouse gases [e.g., *Charlton-Perez et al.*, 2010], the current range of predicted return dates is unnecessarily large due to identifiable model transport and chemistry deficiencies. It is remarkable, and perhaps an encouraging sign of progress in chemistry climate modeling, that a group of models with different dynamical cores, transport schemes, chemical solvers, and spatial resolutions can 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 diagnostics presented here: these models have credible representation of important physical and chemical processes that affect the distribution of ozone and other trace constituents involved in O_3 chemistry. Having models with credible physical processes that produce similar O_3 predictions increases our confidence in our current understanding of the essential chemical and dynamical processes controlling ozone.

[47] **Acknowledgments.** We thank Ashley Jones for use of the ACE N_2O climatology, Andreas Engel for use of the mean age profiles, and the Aura MLS team for daily O_3 data. We thank Michael Prather, Huisheng Bian, Ross Salawitch, and Tim Canty for their contributions to the photochemical evaluations in the SCR. We thank Luke Oman for helpful conversations. CCSRNIES research was supported by the Global Environmental Research Foundation of the Ministry of the Environment of Japan (A-071). CCSRNIES and MRI simulations were completed with the supercomputer at NIES, Japan. This work was supported by the NASA 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.

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, doi:10.1029/2001JD000465.

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, doi:10.1256/qj.02.203.

Austin, J., and R. J. Wilson (2006), Ensemble simulations of the decline and recovery of stratospheric ozone, *J. Geophys. Res.*, *111*, D16314, doi:10.1029/2005JD006907.

Charlton-Perez, A. J., et al. (2010), The potential to narrow uncertainty in projections of stratospheric ozone over the 21st century, *Atmos. Chem. Phys.*, *10*, 9473–9486, doi:10.5194/acp-10-9473-2010.

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, doi:10.1016/j.gloplacha.2006.11.030.

Douglass, A. R., R. B. Rood, S. R. Kawa, and D. J. Allen (1997), A three-dimensional simulation of the evolution of middle latitude winter ozone in the middle stratosphere, *J. Geophys. Res.*, *102*(D15), 19,217–19,232, doi:10.1029/97JD01043.

Douglass, A. R., M. J. Prather, T. M. Hall, S. E. Strahan, P. J. Rasch, L. C. Sparling, L. Coy, and J. M. Rodriguez (1999), Choosing meteorological input for the global modeling initiative assessment of high-speed aircraft, *J. Geophys. Res.*, *104*(D22), 27,545–27,564, doi:10.1029/1999JD900827.

Douglass, A. R., R. S. Stolarski, S. E. Strahan, and P. S. Connell (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. Geosci.*, *2*, 28–31, doi:10.1038/ngeo388.

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

Eyring, V., et al. (2007), Multi-model projections of stratospheric ozone in the 21st century, *J. Geophys. Res.*, *112*, D16303, doi:10.1029/2006JD008332.

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*, D09301, doi:10.1029/2006JD007485.

Hall, T. M. (2000), Path histories and timescales in stratospheric transport: Analysis of an idealized model, *J. Geophys. Res.*, *105*(D18), 22,811–22,823, doi:10.1029/2000JD900329.

Hall, T. M., D. W. Waugh, K. A. Boering, and R. A. Plumb (1999), Evaluation of transport in stratospheric models, *J. Geophys. Res.*, *104*(D15), 18,815–18,839, doi:10.1029/1999JD900226.

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, doi:10.5194/acp-6-5067-2006.

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, doi:10.5194/angeo-26-1391-2008.

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.

Livesey, N., et al. (2007), Earth Observing System (EOS) Microwave Limb Sounder (MLS) Version 2.2 Level 2 data quality and description document, *Tech. Rep. JPL D-32381*, Jet Propul. Lab., Pasadena, Calif.

Mahlman, J. D., H. Levy, and W. J. Moxim (1986), Three-dimensional simulations of stratospheric N_2O—Predictions for other trace constituents, *J. Geophys. Res.*, *91*(D2), 2687–2707, doi:10.1029/JD091iD02p02687.

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, doi:10.5194/gmd-2-43-2009.

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 R. A. Plumb (1999), Age of air in a "leaky pipe" model of stratospheric transport, *J. Geophys. Res.*, *104*(D16), 19,243–19,255, doi:10.1029/1999JD900251.

Oman, L. D., et al. (2010), Multimodel assessment of the factors driving stratospheric ozone evolution over the 21st century, *J. Geophys. Res.*, *115*, D24306, doi:10.1029/2010JD014362.

Pawson, S., R. S. Stolarski, A. R. Douglass, P. A. Newman, J. E. Nielsen, S. M. Frith, and M. L. Gupta (2008), Goddard Earth Observing System chemistry-climate model simulations of stratospheric ozone-temperature coupling, *J. Geophys. Res.*, *113*, D12103, doi:10.1029/2007JD009511.

Perliski, L. M., S. Solomon, and J. London (1989), On the interpretation of seasonal variations of stratospheric ozone, *Planet. Space Sci.*, *37*, 1527–1538, doi:10.1016/0032-0633(89)90143-8.

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, doi:10.1175/1520-0469(2002)059<0414:IOFCAE>2.0.CO;2.

Plumb, R. A., and M. K. W. Ko (1992), Interrelationship between mixing ratios of long-lived stratospheric constituents, *J. Geophys. Res.*, *97*(D9), 10,145–10,156, doi:10.1029/92JD00450.

Schauffler, S. M., E. L. Atlas, S. G. Donnelly, A. Andrews, S. A. Montzka, J. W. Elkins, D. F. Hurst, P. A. Romashkin, G. S. Dutton, and V. Stroud

(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., L. C. Sparling, C. H. Jackman, and E. L. Fleming (2000), A Lagrangian view of stratospheric trace gas distributions, *J. Geophys. Res.*, *105*(D1), 1537–1552, doi:10.1029/1999JD900787.

Schoeberl, M. R., A. R. Douglass, R. S. Stolarski, S. Pawson, S. E. Strahan, and W. Read (2008), Comparison of lower stratospheric tropical mean vertical velocities, *J. Geophys. Res.*, *113*, D24109, doi:10.1029/2008JD010221.

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, doi:10.5194/acp-8-5957-2008.

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, doi:10.5194/acp-8-7055-2008.

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, doi:10.5194/angeo-26-1299-2008.

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.

SPARC CCMVal (2010), SPARC report on the evaluation of chemistry-climate models, edited by V. Eyring, T. G. Shepherd, and D. W. Waugh, *SPARC Rep. 5*, Toronto, Ont., Canada. (Available at http://www.atmosp.physics.utoronto.ca/SPARC.)

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, doi:10.1007/s00382-007-0347-5.

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, doi:10.5194/acp-9-5489-2009.

Stolarski, R. S., and A. R. Douglass (1985), Parameterization of the photochemistry of stratospheric ozone including catalytic loss processes, *J. Geophys. Res.*, *90*(D6), 10,709–10,718, doi:10.1029/JD090iD06p10709.

Stolarski, R. S., and S. M. Frith (2006), Search for evidence of trend slowdown in the long-term TOMS/SBUV total ozone data record: The importance of instrument drift uncertainty, *Atmos. Chem. Phys.*, *6*, 4057–4065, doi:10.5194/acp-6-4057-2006.

Strong, K., et al. (2008), Validation of ACE-FTS N_2O measurements, *Atmos. Chem. Phys.*, *8*, 4759–4786, doi:10.5194/acp-8-4759-2008.

Teyssèdre, H., et al. (2007), A new tropospheric and stratospheric chemistry and transport model MOCA—Climate for multi-year studies: Evaluation of the present-day climatology and sensitivity to surface processes, *Atmos. Chem. Phys.*, *7*, 5815–5860, doi:10.5194/acp-7-5815-2007.

Tian, W., and M. P. Chipperfield (2005), A new coupled chemistry-climate model for the stratosphere: The importance of coupling for future O_3-climate predictions, *Q. J. R. Meteorol. Soc.*, *131*, 281–303, doi:10.1256/qj.04.05.

Waugh, D. W., and V. Eyring (2008), Quantitative performance metrics for stratospheric-resolving chemistry-climate models, *Atmos. Chem. Phys.*, *8*, 5699–5713, doi:10.5194/acp-8-5699-2008.

Waugh, D. W., S. E. Strahan, and P. A. Newman (2007), Sensitivity of stratospheric inorganic chlorine to differences in transport, *Atmos. Chem. Phys.*, *7*, 4935–4941, doi:10.5194/acp-7-4935-2007.

World Meteorological Organization (WMO) (2007), Scientific assessment of ozone depletion: 2006, *WMO Global Ozone Res. and Monit. Proj. Rep. 50*, Geneva Switzerland.

World Meteorological Organization (WMO) (2011), Scientific assessment of ozone depletion: 2010, *WMO Global Ozone Res. and Monit. Proj. Rep. 52*, Geneva Switzerland.

H. Akiyoshi, T. Nakamura, and Y. Yamashita, National Institute of Environmental Studies, 16-2 Onogawa, Tsukuba, Ibaraki 305-8506, Japan.

S. Bekki, D. Cugnet, and M. Marchand, LATMOS, IPSL, UVSQ, UPMC, CNRS, INSU, 4, Place Jussieu, Boite 102-T45-E3 Tour 45-46, E3 F-75252 Paris CEDEX 05, France.

P. Braesicke and J. A. Pyle, NCAS Climate-Chemistry, Centre for Atmospheric Science, Department of Chemistry, Cambridge University, Lensfield Road, Cambridge CB2 1EW, UK.

N. Butchart and S. C. Hardiman, Met Office Hadley Centre, FitzRoy Road, Exeter EX1 3PB, UK.

M. P. Chipperfield, S. Dhomse, and W. Tian, School of Earth and Environment, University of Leeds, Leeds LS2 9JT, UK.

A. R. Douglass, S. M. Frith, S. Pawson, R. S. Stolarski, and S. E. Strahan, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA. (susan.e.strahan@nasa.gov)

S. M. Frith, Science Systems and Applications Inc., Lanham, Maryland, USA.

A. Gettelman, D. E. Kinnison, and J.-F. Lamarque, National Center for Atmospheric Research, 1850 Table Mesa Dr., Boulder, CO 80305, USA.

E. Mancini and G. Pitari, Dipartimento di Fisica, Università degli Studi dell' Aquila, Via Vetoio Coppito, I-67010 L'Aquila, Italy.

M. Michou and H. Teyssèdre, GAME/CNRM, Météo-France, CNRS, 42. Av. G. Coriolis F-31057 Toulouse, France.

O. Morgenstern and D. Smale, National Institute of Water and Atmospheric Research, Private Bag 50061, Omakau, Lauder, Central Otago 9352, New Zealand.

D. Olivié, Department of Geosciences, University of Oslo, N-0315 Oslo, Norway.

D. A. Plummer and J. F. Scinocca, Canadian Centre for Climate Modeling and Analysis, Environment Canada, Victoria, BC V8W 3V6, Canada.

T. G. Shepherd, Department of Physics, University of Toronto, 60 St. George St., Toronto ON M5S 1A7, Canada.

K. Shibata, Meteorological Research Institute, Japan Meteorological Agency, 1-1 Nagamine, Tsukuba, Ibaraki 305-0052, Japan.