Chemistry-climate model simulations of spring Antarctic ozone

John Austin,[1,2] H. Struthers,[3] J. Scinocca,[4] D. A. Plummer,[5] H. Akiyoshi,[6] A. J. G. Baumgaertner,[7] S. Bekki,[8] G. E. Bodeker,[9] P. Braesicke,[10] C. Brühl,[7] N. Butchart,[11] M. P. Chipperfield,[12] D. Cugnet,[8] M. Dameris,[13] S. Dhomse,[12] S. Frith,[14,15] H. Garny,[13] A. Gettelman,[16] S. C. Hardiman,[11] P. Jöckel,[7] D. Kinnison,[16] A. Kubin,[17] J. F. Lamarque,[16] U. Langematz,[17] E. Mancini,[18] M. Marchand,[8] M. Michou,[19] O. Morgenstern,[20] T. Nakamura,[6] J. E. Nielsen,[14,15] G. Pitari,[18] J. Pyle,[10] E. Rozanov,[21,22] T. G. Shepherd,[23] K. Shibata,[24] D. Smale,[20] H. Teyssèdre,[19] and Y. Yamashita[6]

Received 19 November 2009; revised 1 June 2010; accepted 11 June 2010; published 5 November 2010.

[1] Coupled chemistry-climate model simulations covering the recent past and continuing throughout the 21st century have been completed with a range of different models. Common forcings are used for the halogen amounts and greenhouse gas concentrations, as expected under the Montreal Protocol (with amendments) and Intergovernmental Panel on Climate Change A1b Scenario. The simulations of the Antarctic ozone hole are compared using commonly used diagnostics: the minimum ozone, the maximum area of ozone below 220 DU, and the ozone mass deficit below 220 DU. Despite the fact that the processes responsible for ozone depletion are reasonably well understood, a wide range of results is obtained. Comparisons with observations indicate that one of the reasons for the model underprediction in ozone hole area is the tendency for models to underpredict, by up to 35%, the area of low temperatures responsible for polar stratospheric cloud formation. Models also typically have species gradients that are too weak at the edge of the polar vortex, suggesting that there is too much mixing of air across the vortex edge. Other models show a high bias in total column ozone which restricts the size of the ozone hole (defined by a 220 DU threshold). The results of those models which agree best with observations are examined in more detail. For several models the ozone hole does not disappear this century but a small ozone hole of up to three million square kilometers continues to occur in most springs even after 2070.

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

1. Introduction

[2] Since its discovery [*Farman et al.*, 1985], the Antarctic ozone hole has been a frequent topic for research using both observations [e.g., *Bodeker et al.*, 2002, 2005] and models [e.g., *Struthers et al.*, 2009; *Austin and Wilson*, 2006; *Eyring et al.*, 2006]. The phenomenon is well understood [e.g., *Solomon*, 1999] and can now be simulated quantitatively by many models. Essentially, heterogeneous reactions take place on the

[1]Geophysical Fluid Dynamics Laboratory, NOAA, Princeton, New Jersey, USA.
[2]University Corporation for Atmospheric Research, Boulder, Colorado, USA.
[3]Department of Applied Environmental Science, University of Stockholm, Stockholm, Sweden.
[4]CCCMA, University of Victoria, Victoria, British Columbia, Canada.
[5]Science and Technology Branch, Environment Canada, Toronto, Ontario, Canada.
[6]National Institute for Environmental Studies, Tsukuba, Japan.
[7]Max-Planck-Institut für Chemie, Mainz, Germany.
[8]LATMOS, IPSL, UVSQ, UPMC, CNRS, INSU, Paris, France.
[9]Bodeker Scientific, Alexandra, New Zealand.
[10]NCAS-Climate-Chemistry, Centre for Atmospheric Science, Department of Chemistry, Cambridge University, Cambridge, UK.
[11]Hadley Centre, Met Office, Exeter, UK.
[12]School of Earth and Environment, University of Leeds, Leeds, UK.
[13]Deutsches Zentrum für Luft- und Raumfahrt, Institut für Physik der Atmosphäre, Wessling, Germany.
[14]NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.
[15]Science Systems and Applications, Inc., Beltsville, Maryland, USA.
[16]NCAR, Boulder, Colorado, USA.
[17]Institut für Meteorologie, Freie Universität, Berlin, Germany.
[18]Dipartimento di Fisica, University of L'Aquila, L'Aquila, Italy.
[19]GAME, CNRM, Meteo-France, Toulouse, France.
[20]National Institute of Water and Atmospheric Research, Lauder, New Zealand.
[21]Physical-Meteorological Observatory, World Radiation Center, Davos, Switzerland.
[22]ETH Zürich, Zurich, Switzerland.
[23]Department of Physics, University of Toronto, Toronto, Ontario, Canada.
[24]Meteorological Research Institute, Japan Meteorological Agency, Tsukuba, Japan.

Copyright 2010 by the American Geophysical Union.
0148-0227/10/2009JD013577

surfaces of polar stratospheric clouds, transforming unreactive chlorine and bromine reservoir species (HCl, $ClONO_2$, HBr, $BrONO_2$) into active forms (Cl_2, HOCl, etc.). Owing to the presence of higher amounts of chlorine species, related to anthropogenic emissions of chlorofluorocarbons and halons, photolysis and catalytic ozone destruction cycles with the ClO dimer now play a greater role than prior to the formation of the ozone hole. Recommended rates of the most relevant reactions have not changed substantially since the review by *Solomon* [1999] and although the details of the heterogeneous reactions are uncertain, the above summary remains unchallenged.

[3] In this paper, we investigate simulations of the ozone hole using many chemistry-climate models which have contributed to the Stratospheric Processes and their Role in Climate (SPARC) Chemistry-Climate Model Validation (CCMVal) project [*Eyring et al.* 2005]. A consistent chemical reaction set is taken from *Sander et al.* [2006], and all the models used here are capable of simulating an ozone hole given the right physical conditions: sufficient polar stratospheric clouds (PSCs), sufficient halogen amounts as well as sunlight. Diagnostics may be separated into 'simple' diagnostics, which can be calculated directly from the ozone field itself, and 'complex' diagnostics, which require the incorporation of additional fields such as potential vorticity or information of the polar vortex.

[4] Prior to 1980 the Antarctic ozone column was rarely observed to be less than 220 DU, which is now commonly taken as the threshold for the occurrence of the ozone hole. In this paper, we investigate commonly used, 'simple' diagnostics: the area of total ozone less than 220 DU, the minimum spring total ozone outside the tropics, and the ozone mass deficit [*Bodeker et al.*, 2005]. These diagnostics have their advantages as well as their undoubted weaknesses. The disadvantage of some simple diagnostics is that they sometimes obscure underlying model shortcomings which may be unrelated to the physics of the ozone hole itself. For example some models may have a high ozone bias, which artificially restricts the size of the ozone hole when that hole is defined in terms of a fixed column (220 DU). Other models may have a realistic ozone hole but it is displaced upward or downward due to the model thermal structure. Still other models may simulate the meridional mixing barrier at the polar vortex edge that is less sharp than observed, or displaced in latitude relative to observations [e.g., *Struthers et al.*, 2009]. Several methods have been proposed to examine the performance of the simulated polar ozone loss [e.g., *Huck et al.*, 2007; *Tilmes et al.*, 2008] and they point out some of the disadvantages of the above simple diagnostics. On the positive side, the diagnostics are easy and fast to calculate and most of the essential physics emerges. The diagnostics do not require additional fields for their computation and have stood the test of time in that they are still being used after, in some cases, almost 20 years of publication. These diagnostics address the past and future state of total column ozone, since this is directly related to changes in solar ultraviolet radiation, whether the ozone changes are due to chemistry, dynamics or radiation. It is certainly plausible to refine the definition of relevant diagnostics, simple or otherwise. For example instead of choosing the absolute minimum, the ozone diagnostic could be the average poleward of a fixed equivalent latitude value [e.g., *Müller et al.*, 2008], which would tend to reduce the chance of sampling local minima due to synoptic-scale variability in the dynamics. Although this would be useful to focus more on chemical loss for the Arctic, it would be less useful for the Antarctic for model simulations which often differ significantly from observations. For example, restricting the latitudinal extent of the ozone hole can lead to artificial conclusions in those models in which the low-ozone columns extend to low latitudes. Finally, because of the coupling of the chemistry with the temperature, problems in simulating the ozone hole will lead to impacts on the model dynamics, whether these problems relate to a fixed 220 DU column, or are revealed by more complex diagnostics.

[5] For the CCMVal model assessment [*Eyring et al.*, 2006], simple Antarctic ozone hole diagnostics were determined for the 1990s. In general, models underestimated the size of the ozone hole (using the classical 220 DU definition) as well as the ozone mass deficit, calculated as the mean loss relative to the 220 DU ozone amount averaged for September and October. There was no clear consensus on the simulated minimum Southern Hemisphere value, which observationally has remained robust at about 100 DU throughout most of the last two decades. In this paper, we explore these issues using a combination of diagnostics to identify problems in the simulation of the ozone hole and to suggest a strategy for resolving these problems where possible.

2. Model Descriptions and Simulations

[6] Results are taken from the CCMs described by *Morgenstern et al.* [2010], and the CCMs have well-resolved stratospheres. In addition results are included from the future simulation for EMAC, which is here indicated as EMAC-FUB. This model is a modified version of EMAC with improved representation of PSCs but lower vertical resolution (39 levels compared with 90 for EMAC). Of particular significance for the current study: heterogeneous reactions are taken to occur on the surfaces of supercooled ternary solution (STS) droplets as well as nitric acid trihydrate (NAT) and ice PSCs. CMAM does not include reactions on NAT, but the reaction rates on STS increase rapidly near the NAT PSC temperature threshold (195 K). Although the general characteristics are the same for all models, there are detailed differences between different schemes as described by *Morgenstern et al.* [2010]. The implications of these differences for the results of the current paper are discussed in section 7.

[7] The model simulations cover the period 1950–2099 or a subset thereof, in two experiments, REF-B1 and REF-B2 [*Eyring et al.*, 2008]. REF-B1 covered the past, from 1950 or so, to 2007, with sea surface temperatures (SSTs) and sea ice specified from observations. REF-B2 covered the period 1950–2099 (or a subset) with SSTs supplied from a coupled atmosphere-ocean experiment, depending on the model used. The main period of investigation is 1980–2008 when observations exist for comparisons, and the ozone hole was at least partially present in the observations. We also use the REF-B2 simulations to investigate the short- and long-term behavior of the ozone hole in the model simulations. Table 1 (to be described later) shows the models contributing to the analyses. The main difference between experiments REF-B1 and REF-B2 for the overlapping period is the SSTs, but with the low-frequency external forcings (Solar, quasi-biennial oscillation (QBO)) absent in REF-B2, except for those models

Table 1. Mean Low-Temperature Areas ($T < 195$ K, in Units of 10^6 km^2) for the Period July to September for the Years 1980–2007 in Comparison With Observations for the Models Used in Each Group of Experiments[a]

Model	REF-B1	REF-B2	Reference
NCEP data	21.4 ± 0.8		
AMTRAC3	19.8 ± 0.8	20.6 ± 0.6	*Austin and Wilson* [2010]
CAM3.5	17.5 ± 1.5	14.5 ± 1.7	*Lamarque et al.* [2008]
CCSRNIES	25.8 ± 0.5	26.5 ± 0.6	*Akiyoshi et al.* [2009]
CMAM	19.8 ± 0.6	20.3 ± 0.3	*Scinocca et al.* [2008]; *de Grandpré et al.* [2000]
CNRM-ACM	19.1 ± 1.4	19.2 ± 1.2	*Déqué* [2007]; *Teyssèdre et al.* [2007]
EMAC	19.0 ± 1.4	22.5 ± 0.7	*Jöckel et al.* [2006]
E39CA	24.1 ± 0.7	24.0 ± 0.7	*Stenke et al.* [2009]; *Garny et al.* [2009]
GEOSCCM	17.5 ± 0.4		*Pawson et al.* [2008]
LMDZrepro	21.5 ± 0.8		*Jourdain et al.* [2008]
MRI	22.6 ± 0.4	22.7 ± 0.4	*Shibata and Deushi* [2008]
Niwa-SOCOL	23.1 ± 0.6	22.0 ± 0.5	*Schraner et al.* [2008]
SOCOL	21.6 ± 0.4	20.5 ± 0.4	*Schraner et al.* [2008]
ULAQ	21.7 ± 1.6	21.7 ± 1.4	*Pitari et al.* [2002]
UMETRAC	18.6 ± 1.2		*Austin and Butchart* [2003]; *Struthers et al.* [2004]
UMSLIMCAT	18.9 ± 0.7	19.7 ± 0.9	*Tian and Chipperfield* [2005]
UMUKCA-METO	14.0 ± 0.8	16.3 ± 0.5	*Morgenstern et al.* [2009]
UMUKCA-UCAM	14.7 ± 0.7	15.8 ± 0.6	*Morgenstern et al.* [2009]
WACCM	23.9 ± 1.1	22.0 ± 1.5	*Garcia et al.* [2007]
Multimodel mean	20.2 ± 1.5	20.6 ± 1.6	

[a]The uncertainties indicated are approximate 95% confidence intervals for the random error, given by $2s/\sqrt{(n-1)}$,where s is the standard deviation of the annual values and n is the number of years included. The WACCM values are for August and September only. The EMAC REF-B2 results are from the EMAC-FUB model.

which had a naturally occurring QBO. Most models completed one simulation of each experiment, but several models completed ensembles of simulations. In this work we consider each of these ensemble members, but we find that intermodel differences are much larger than the differences between ensemble members. The greenhouse gas concentrations for all the simulations were specified from observations for the past and SRES scenario A1b for the future [*Intergovernmental Panel on Climate Change* (*IPCC*), 2001, Appendix II]. According to this scenario, CO_2 increases by 94% from 2000–2100, N_2O increases by 18% from 2000–2100, and CH_4 increases by 36% (reached in 2050) before declining. Tropospheric CFC and Halon concentrations are specified from the A1 profile of *World Meteorological Organization* (*WMO*) [2007, Table 8.5]. However, many models also specified additional bromine of about 6 pptv to allow for that contained in the very short lived species [*WMO*, 2007, chapter 2].

[8] In the case of GEOSCCM, the REF-B2 simulation was for 2000–2099, and for this paper, the results of REF-B1 and REF-B2 are spliced together on 1 January 2000. In the Niwa-SOCOL REF-B2 simulation there is a change in the observed SST data set between 2003 and 2004, although this does not alter the major features of the results in high latitudes. For E39CA and EMAC-FUB the REF-B2 simulation was not completed, and instead results are taken from the sensitivity experiment SCN-B2d, which is similar to REF-B2, but includes a solar cycle and the quasi-biennial oscillation. For convenience the SCN-B2d runs used and the GEOSCCM spliced run are referred to as REF-B2 runs. Similarly, the AMTRAC3 REF-B2 simulation contains a solar cycle for the whole integration.

3. Evolution of the Ozone Hole

3.1. Sensitivity of Results to the Edge of the Ozone Hole

[9] Figure 1a shows the zonal mean column ozone in the model REF-B2 simulations, averaged for 10 days on either side of the ozone minimum for 1996–2005. Several models agree quite well with the National Institute of Water and Atmosphere (NIWA) combined ozone database [*Bodeker et al.*, 2005], updated (http://www.bodekerscientific.com/data/ozone) version 2.7 'LongPatched' daily data). However, several models are biased high over a wide latitude range, and in particular place the 220 DU ozone column too far poleward compared with the observations.

[10] One method of trying to correct for the apparent bias in the models is to adjust the model results relative to the preozone hole minimum. As noted above, 220 DU was a rarely observed column ozone prior to the ozone hole, and hence in Figure 1b the model results are adjusted relative to the minimum attained in the southern extratropics throughout the period 1960–1965, using daily data. For example, the 1960–1965 minimum for AMTRAC3 was 199 DU implying that the model is biased low by about 21 DU. Figure 1b therefore shows the AMTRAC3 results increased by 21 DU to compensate. All the other model results were also adjusted by an amount appropriate to each model. The effect of these corrections has been to improve some model results relative to observations but others are made worse. The implication is that the discrepancies from observations shown in Figure 1a are typically not a simple column bias.

[11] An alternative correction is illustrated in Figure 1c, described below, based on the position of the maximum meridional gradient in each model. The edge of the ozone hole is typically within the polar vortex, the edge of which is denoted by the steepest ozone gradients [*Bodeker et al.*, 2002; *Newman et al.*, 2007; *Struthers et al.*, 2009]. The magnitudes of the gradients as a function of latitude are shown in Figure 2. Again several models agree reasonably well with observations, but several models are systematically in error, placing the steepest gradients too close to the pole. The range of model results for the maximum in the ozone gradient is shown in Figure 3, as a function of the ozone at that position. As seen in Figure 3, most models

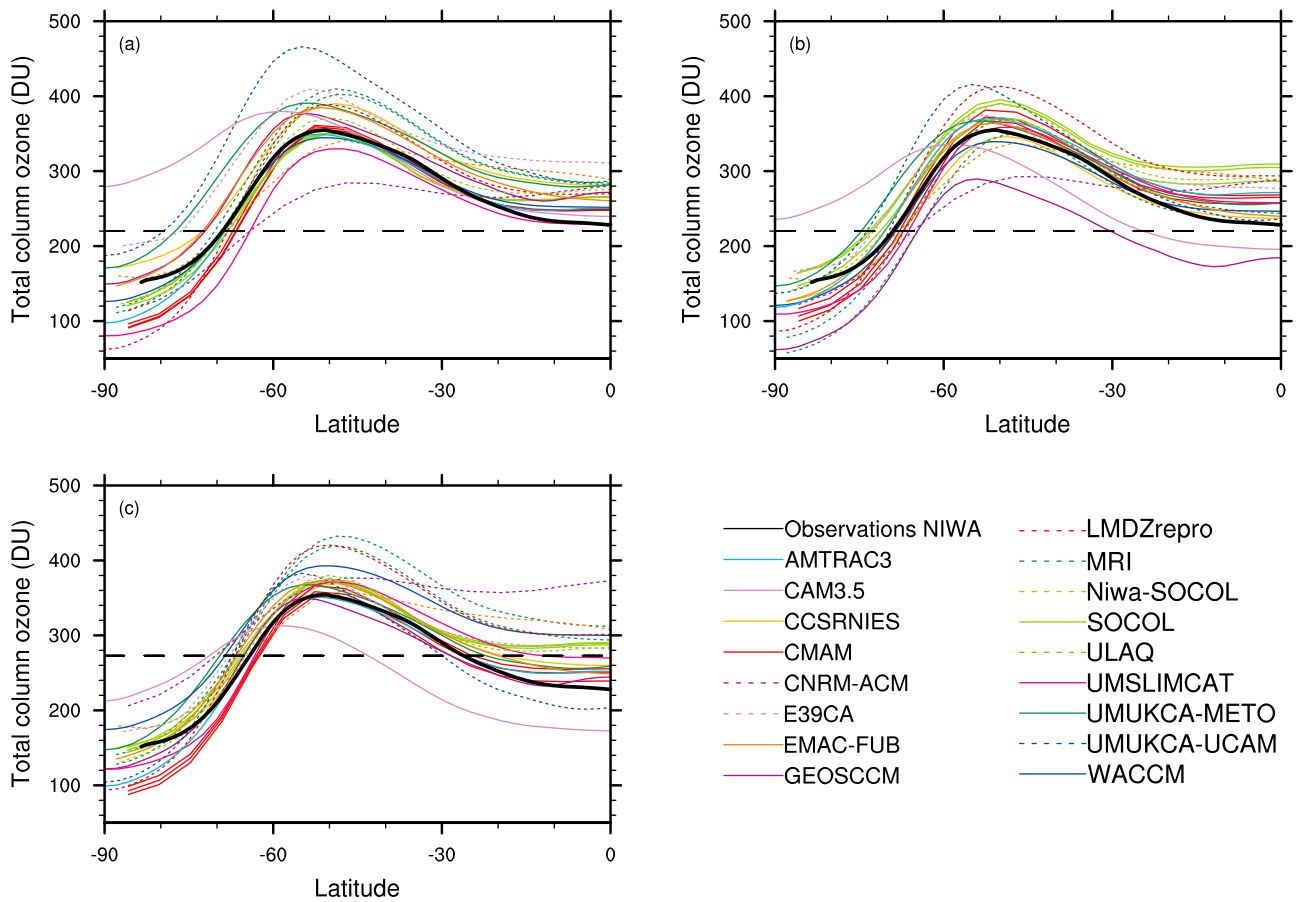

Figure 1. Total column ozone as a function of latitude, averaged for the period 1996–2005 for 10 days before and after the minimum column ozone. (a) No adjustments to the model results. (b) Model results have been adjusted relative to the 1960–1965 minimum (see text). (c) Model results have been adjusted relative to the ozone maximum meridional gradient (see text). The results have been obtained from the REF-B2 simulations.

place the steepest gradient poleward of the observations (64°S). Hence, most models will have a restricted ozone hole. In Figure 1c, the ozone latitudinal variation is adjusted to try to correct for this deficiency. For each model, the adjustment is given by the displacement of the model results from observations on the ordinate in Figure 3. For example, for UMUKCA-METO, the peak gradient occurs at 69.5°S (296 DU) whereas in the observations the peak gradient occurs at 64.2°S with a corresponding ozone value of 273 DU. This may suggest that UMUKCA-METO is biased high by 23 DU and in Figure 1c the UMUKCA-METO results have been reduced by 23 DU to compensate. After the adjustments, applied individually to each model, the results obtained are shown in Figure 1c and are generally closer to observations than in Figures 1a and 1b.

3.2. Ozone Hole Area

[12] The maximum ozone hole areas in the model simulations, computed using the above criteria are illustrated in Figure 4 from the REF-B2 simulations. For the ozone hole area defined as the area with ozone column less than 220 DU (Figure 4a), many of the models agree with observations to a reasonable approximation, but quantitative differences remain. For CCMVal, the simulated ozone hole area was typically smaller than observed by about 20% [*Eyring et al.*, 2006]. For CCMVal-2, many models have improved (AMTRAC3, CMAM, MRI, SOCOL, UMSLIMCAT, and WACCM) but several have become worse (CCSRNIES, E39CA, GEOSCCM, and LMDZrepro) while the ULAQ ozone hole area is about the same. Thus the mean model ozone hole remains about 20% smaller than observed.

[13] For the ozone hole area based on the 1960–1965 minimum (Figure 4b), CCSRNIES, E39CA, GEOSCCM, and MRI results are much improved, suggesting that their problem in simulating the ozone hole area is mainly due to an overall ozone high bias. UMUKCA-METO and CAM3.5 are improved by a smaller margin, while SOCOL (and Niwa-SOCOL) results are worse in this framework. This is because SOCOL (and Niwa-SOCOL) simulates low-ozone columns, due to the dynamical characteristics of the vortex, even when there is little chemical destruction. Measured relative to the steepest ozone gradients (Figure 4c) the models are generally more consistent with observations. In particular CAM3.5 and UMUKCA-METO results are considerably improved, suggesting that a large part of the problem in these models is dynamical in origin.

[14] Overall these results suggest that some models do not simulate well the vortex structure, including for example a

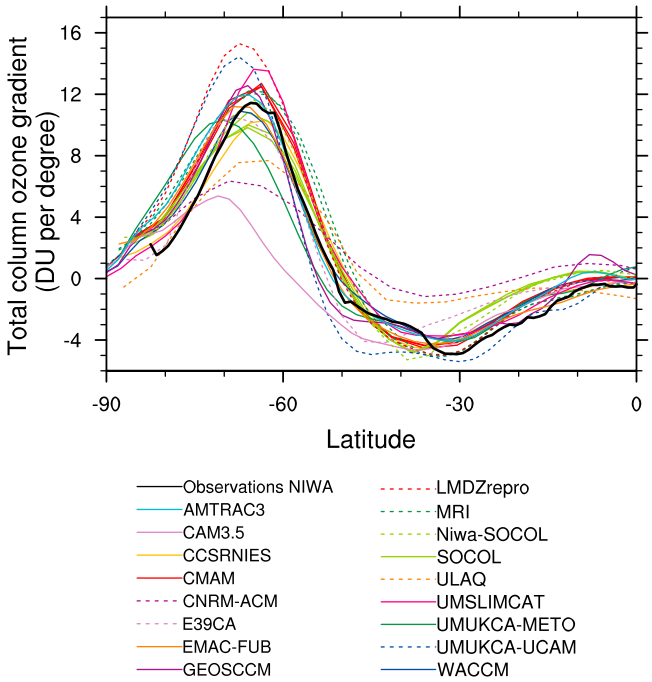

Figure 2. Meridional gradient in total column ozone averaged for the period 1996–2005 for the 10 days on either side of the ozone minimum. The results have been obtained from the REF-B2 simulations.

delay in the final warming [*Eyring et al.*, 2006; *Hurwitz et al.*, 2010]. With such a large spread in model results for both 1980 and 2060, predictions of the disappearance of the ozone hole remain unreliable, and in any case, Figure 1 indicates that these predictions are likely to be definition dependent. Overall, there has not been any clear overall improvement in the simulation of the ozone hole since CCMVal [*Eyring et al.*, 2006].

[15] While it is useful to compare diagnostics of the ozone hole in an adjusted framework, such as relative to simulated steep gradients as in Figure 4, correcting for model chemistry or dynamics weaknesses is problematic. This is illustrated in Figure 4c, which shows for some models a 'large ozone hole' at the end of the simulation in the adjusted framework. The problem arises in part from the use of coupled simulations. Once the model dynamics does not agree with observations, then its temperature or transport behavior can degrade the ozone simulation. Likewise, if there is some homogeneous gas phase chemistry problem preventing good agreement with ozone measurements, the dynamics (hence the low-temperature areas driving the PSCs) can be affected. For the remainder of the manuscript, therefore, the unadjusted model results are used, referring as needed to the conclusions inferred from the adjusted results.

3.3. Antarctic Ozone Minima

[16] Figure 5 and Table 2 show the model results for the minimum Antarctic ozone in each spring season (September to November). Results have been taken from experiment REF-B1 to which more models contributed. There is a wide spread in model results, although the mean of all the models (Table 2) is close to that observed. Several models (MRI, ULAQ, and WACCM) agree with observations throughout the period, while other models (AMTRAC3, UMETRAC, and CMAM) agree with observations prior to about 1990 but then drift lower. However, ULAQ has occasional very low values in the fall (not shown) when the ozone hole is not present in the observational record. Niwa-SOCOL and SOCOL agree best with observations during the later part of the record, and are systematically lower in the early period. The other models tend to be systematically low throughout the period (CNRM-ACM, LMDZrepro, UMSLIMCAT) or high throughout the period (CAM3.5, CCSRNIES, EMAC, E39CA, GEOSCCM and the UMUKCA pair of models). Although CCSRNIES and EMAC tend to be high throughout the period of observation, the discrepancy is more marked once the ozone hole reaches its maturity. This is discussed later in the context of the simulated PSCs (section 4).

[17] In many cases, the pedigree of the individual models is clear from Figure 5. Niwa-SOCOL and SOCOL are identical except for the different lower boundary conditions. UMUKCA-METO and UMUKCA-UCAM share a common core climate model and their results are very similar. CNRM-ACM and LMDZrepro have a common chemical scheme and their results are very similar. AMTRAC3 and UMETRAC share a common chemical solver, and although the halogen parameterization has been changed for AMTRAC3 [*Austin and Wilson*, 2010] the results are also very similar.

[18] Those models which contributed data for REF-B2 produced results similar to the REF-B1 results (Figure 6). Again several models (AMTRAC3, CMAM, LMDZrepro, and UMSLIMCAT) simulated a deeper ozone hole than observed. The only models which indicated ozone recovery to over 220 DU by the end of the simulation had an ozone high bias. The other models yield an ozone recovery to 1980 values by about 2070, but thereafter the ozone increase is simulated to be small.

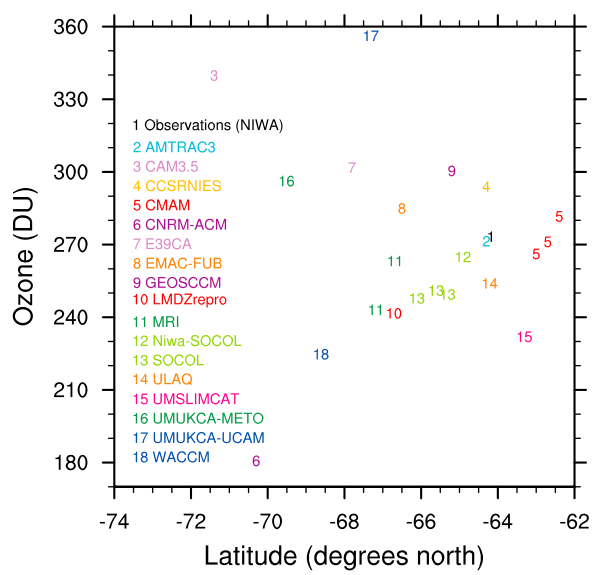

Figure 3. Latitude of maximum meridional gradient in total column ozone, as a function of the ozone value at that latitude. The results have been obtained from the REF-B2 simulations.

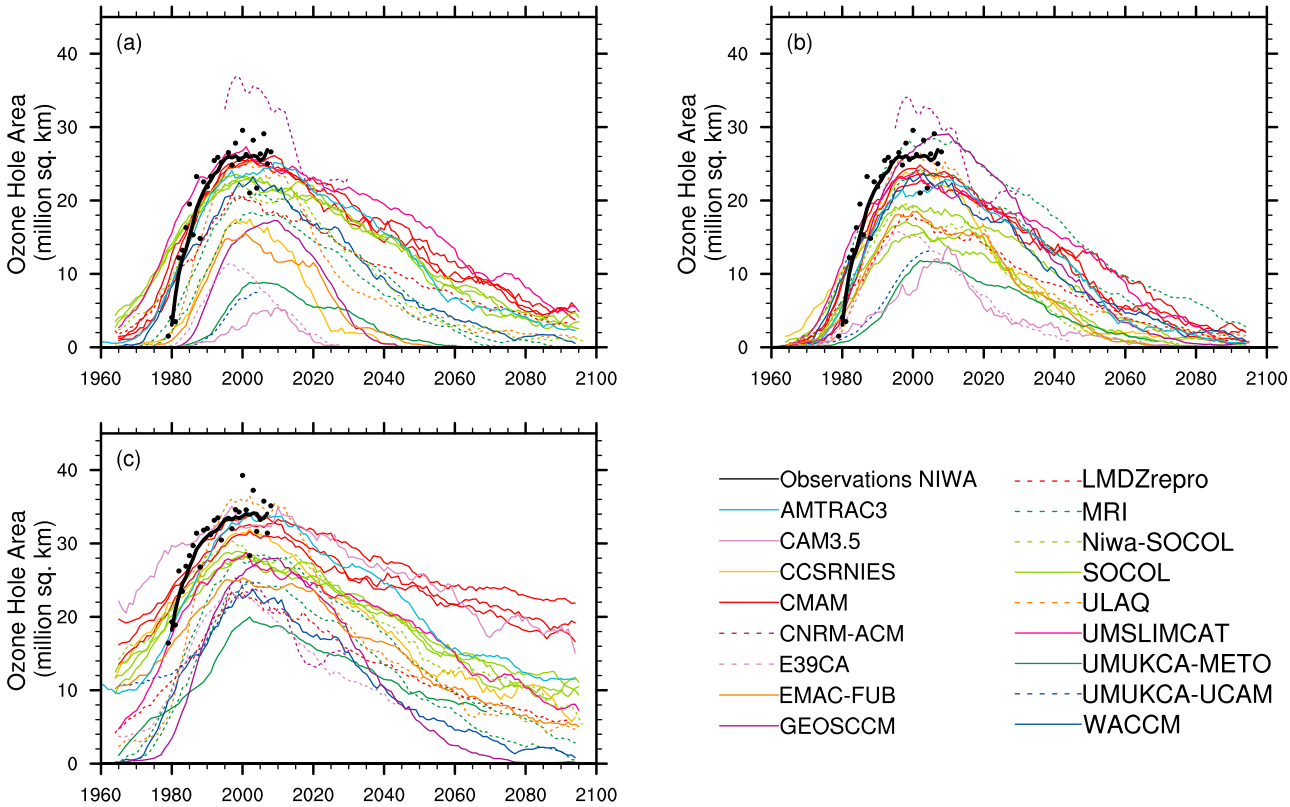

Figure 4. Simulated and observed ozone hole areas, (a) based on a fixed, 220 DU amount, (b) based on the 1960–1965 minimum, and (c) based on the value at the maximum gradient. The curves indicate 11 year running means of the observations and model results for individual years. The model results have been obtained from the REF-B2 simulations.

3.4. Date of the Ozone Minimum

[19] As the ozone hole has become deeper, the date on which the ozone minimum occurred has tended to drift earlier in the season [e.g., *Bodeker et al.*, 2005]. This is essentially due to the increase in halogen amounts which allow the ozone loss to be accelerated. The REF-B1 simulations (Figure 7) are about evenly divided, with half of the models agreeing with the observed tendency of -3 ± 2 d/decade (1σ) (AMTRAC3, CAM3.5, CMAM, CNRM-ACM, GEOSCCM, LMDZrepro, MRI, UMETRAC, UMSLIMCAT, and WACCM). SOCOL and ULAQ also have the same sign, but their trends are smaller and larger, respectively, than observed. Another group of models has the opposite tendency to that observed (CCSRNIES, EMAC, E39CA, Niwa-SOCOL, UMUKCA-METO and UMUKCA-UCAM). In the models in which the date of the minimum increases, there is a tendency for the ozone hole to be less prominent than observed (see Figures 4, 5, and 6).

[20] There is a large interannual variability in the date of the minimum, as indicated in the observations in Figure 7. Model results show comparable variability. Therefore detecting a trend of order 6 days in the timing of the minimum is challenging, particularly for those models which supplied only 10 day frequency output (UMSLIMCAT, WACCM). However, in both those cases, the uncertainty in the trend is approximately the same as for the other models, 3 d/decade. Combining all the model results, except for EMAC and UMETRAC which were stopped prematurely, the mean trend in the date of the minimum is -2.0 ± 0.8 d/decade for the period 1990 to 2004.

[21] Eleven of the 18 models reach the minimum later than observed, by typically 10 days, and in many cases the ozone holes are deeper than observed (AMTRAC3, CMAM, LMDzrepro). For those models in particular, a shorter ozone hole season would result in improved agreement with observations. In comparison, SOCOL benefits from a more punctual ozone minimum, although as noted above, that model doesn't reproduce the trend as well as, possibly for the same reason, the depth of the ozone hole early in its development stage (Figure 5).

3.5. Ozone Mass Deficit

[22] Figures 8 and 9 and Table 2 show the model results for the ozone mass deficit, defined as the total mass of ozone lost below 220 DU averaged over the months of September and October [*Bodeker et al.*, 2005]. For those models which supplied data for several simulations (CMAM, MRI, SOCOL), the individual members of each ensemble agreed, except for the MRI model for which a slight difference is present. A large depletion over a large area will have a substantially higher mass deficit than a small depletion over a small area. For this diagnostic the models have an even wider spread than for the Antarctic ozone minimum, indicating that the diagnostic is a sensitive test of model performance. This arises

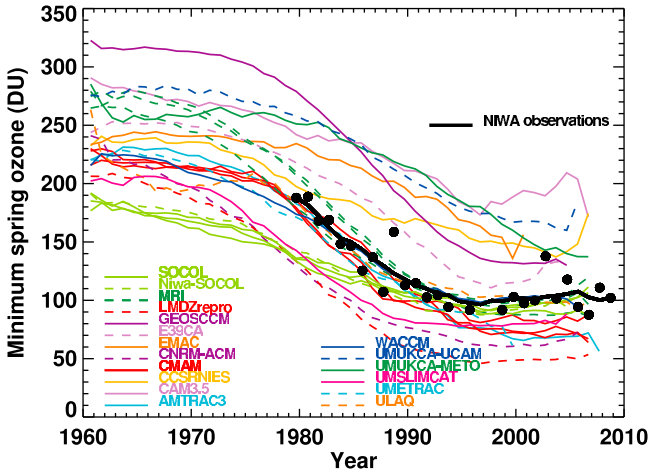

Figure 5. Minimum Antarctic spring ozone for the period 1960–2008 for each model, compared with observations (thick black line). The results were calculated from the REF-B1 simulations. A running mean filter has been applied to the results of the models and observations. The filter has a width of 11 years in the main part of the results reducing to unfiltered results at the beginning and end of the data period. For the models UMSLIMCAT and WACCM, data were available only at a 10 day frequency.

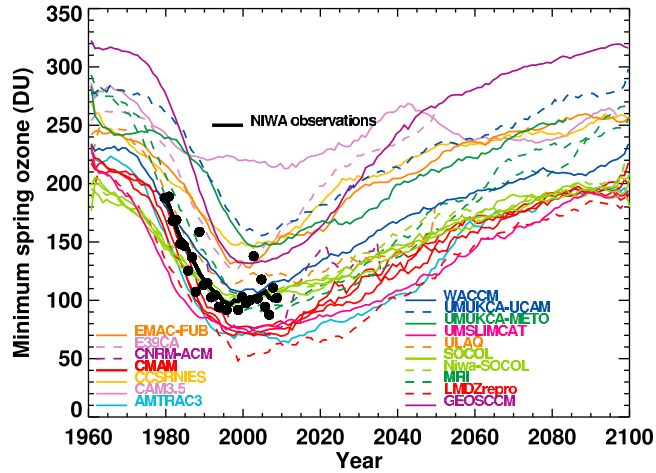

Figure 6. As in Figure 5 but for the experiment REF-B2, and covering the period to 2100. For GEOSCCM, the REF-B1 results have been included, joined at the year 2000 to provide a more complete picture.

from the compounding of errors noted in the ozone hole area and ozone minimum diagnostics. CAM3.5, UMUKCA-METO and -UCAM perform particularly poorly in this diagnostic, since their 220 DU ozone holes are too small and shallow. Most models tend to simulate too small an ozone deficit, but several models (CNRM-ACM, LMDZrepro, Niwa-SOCOL, SOCOL, UMSLIMCAT) simulate an ozone deficit that is too large. Other models (AMTRAC3 and CMAM) agree with observations up until about the middle 1990s, and then apparently exceed the observations. However, the diagnostic is strongly dependent on the dynamics of the polar vortex and on two years (2002 and 2004) the data were considerably below those observed on other years. Ignoring these points would yield agreement within about 10% between observations and the models AMTRAC3, CMAM, Niwa-SOCOL, SOCOL and WACCM. Most models produced very similar results for REF-B2 as for REF-B1. The main exceptions were Niwa-SOCOL, SOCOL, ULAQ and WACCM, which all simulated substantially less ozone mass deficit for the REF-B2 than for the REF-B1 simulations. This would suggest a sensitivity in some models to the SSTs, as well as the need for accurate prediction of the SSTs to simulate the recovery of ozone. See also *Garny et al.* [2009] and *Austin and Wilson* [2010].

4. PSCs and the Relationship With Ozone Hole Area

4.1. PSCs

[23] Observations of PSCs are not readily available to the global extent needed for a complete comparison of CCMs with measurements over the several decade time scale needed (although see section 7). All the models except CMAM included NAT and ice, with different assumptions regarding particle sizes [*Morgenstern et al.*, 2010]. Most models, including CMAM simulated STS, although STS reaction rates increase rapidly at about the thermodynamic equilibrium temperature of NAT. With typical concentrations of H_2O and HNO_3 of 4.5 ppmv and 10 ppbv, respectively, the 50 hPa ice and NAT PSC temperatures [*Hanson and Mauersberger*, 1988] are 187.9 and 195.4 respectively. In recognition of this, the CCMVal-2 project archived the areas within the 188K and 195K temperature

Table 2. Commonly Used Antarctic Ozone Hole Diagnostics, Averaged Over the Period 1990–2008, or the End of the REF-B1 Simulations, Depending on the Model[a]

Model	Minimum Antarctic Ozone	Maximum Ozone Hole Area	Ozone Mass Deficit
NIWA data	103 ± 6	26.1 ± 1.2	22.0 ± 2.7
AMTRAC3	74 ± 8	21.8 ± 1.8	24.4 ± 3.8
CAM3.5	187 ± 19	7.5 ± 2.5	1.1 ± 0.5
CCSRNIES	148 ± 10	16.9 ± 2.3	6.6 ± 2.1
CMAM	79 ± 6	23.2 ± 0.8	25.2 ± 2.2
CNRM-ACM	63 ± 4	38.2 ± 3.5	42.4 ± 4.1
EMAC	167 ± 16	10.6 ± 2.5	2.6 ± 1.6
E39CA	121 ± 12	11.7 ± 1.6	3.7 ± 1.1
GEOSCCM	139 ± 8	13.4 ± 1.3	4.6 ± 1.1
LMDZrepro	48 ± 3	22.9 ± 1.1	31.0 ± 2.5
MRI	97 ± 3	14.7 ± 0.6	14.2 ± 1.2
Niwa-SOCOL	92 ± 6	26.0 ± 1.6	28.2 ± 3.7
SOCOL	95 ± 4	26.6 ± 0.8	28.7 ± 2.4
ULAQ	102 ± 7	22.5 ± 2.2	15.2 ± 3.4
UMETRAC	91 ± 13	18.6 ± 3.1	16.3 ± 6.0
UMSLIMCAT	79 ± 4	25.0 ± 1.3	34.8 ± 3.5
UMUKCA-METO	168 ± 16	6.2 ± 2.0	2.2 ± 1.1
UMUKCA-UCAM	172 ± 8	5.0 ± 0.8	0.9 ± 0.4
WACCM	101 ± 7	26.4 ± 2.3	22.9 ± 4.3
Multimodel mean	112 ± 19	18.7 ± 4.0	16.9 ± 6.1

[a]The uncertainties indicated are approximate 95% confidence intervals for the random error, given by $2s/\sqrt{(n-1)}$, where s is the standard deviation of the annual values and n is the number of years included. For the multimodel mean, the uncertainty given is $2s/\sqrt{(n-1)}$ where s is the standard deviation of the individual model mean values and n is the number of models (18). The units are: minimum ozone, DU; maximum ozone hole area, 10^6 km^2; and ozone mass deficit, Mt.

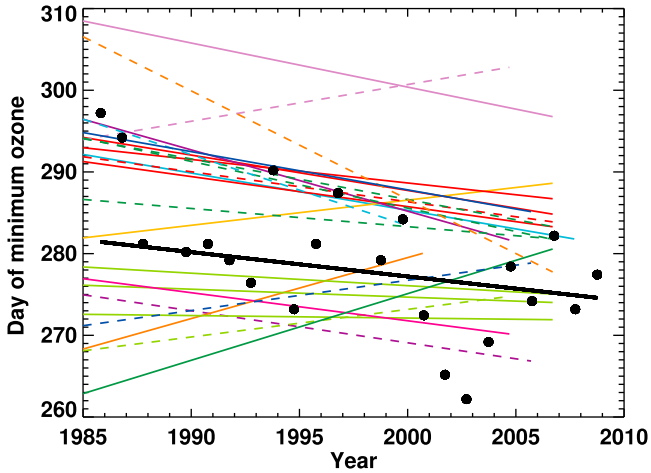

Figure 7. Day that the minimum spring ozone is attained in the REF-B1 simulations. The lines are linear regressions through the individual model years for the period 1985 to the end of each simulation. The observations from the NIWA database are shown in black both as a linear regression and also with individual black points to show the interannual variability. See Figure 8 for the color scheme of the different lines.

thresholds, denoted A_{188} and A_{195}, respectively. The actual PSCs which drive model chemistry will not necessarily follow A_{195}, but as shown by *Austin and Wilson* [2010], this comparison with observations can provide a first test of model performance. However the model PSCs are determined, it is likely that low-temperature regions have a substantial impact on heterogeneous reactions and hence models which do not simulate well the low-temperature areas will likely be deficient.

[24] Figures 10 and 11 show the 50 hPa A_{195} determined for each of the model simulations in REF-B1 and REF-B2, averaged for the period July to September for each year. This is qualitatively very similar to the accumulated sum for each year presented previously by, for example, *Pawson*

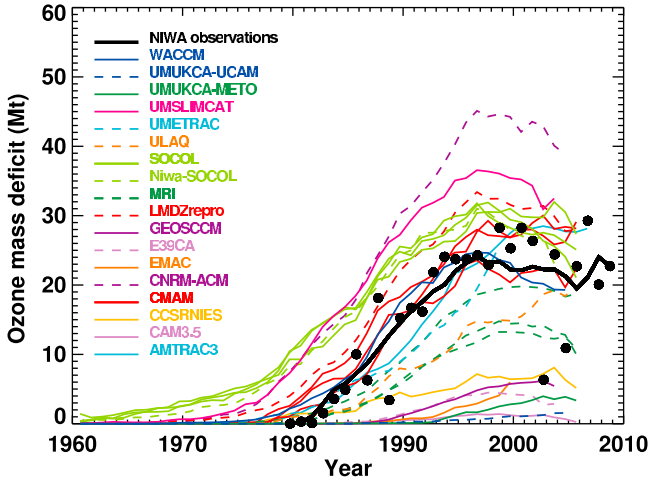

Figure 8. Ozone mass deficit for the period 1960–2008 for each model compared with observations (thick black line). The results were calculated from the REF-B1 simulations.

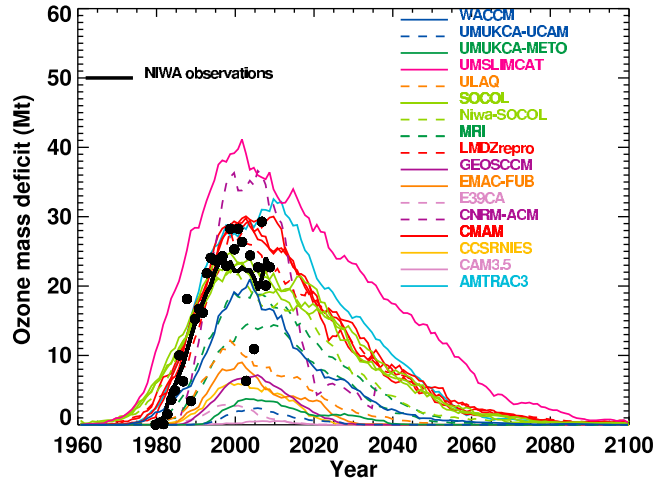

Figure 9. As in Figure 8 but for the experiment REF-B2, and covering the period to 2100.

et al. [1999] and *Austin et al.* [2003], but is here the preferred measure of PSC-related diagnostic, as indicated in section 4.2.

[25] Many models (AMTRAC3, CMAM, LMDZrepro, MRI, Niwa-SOCOL, SOCOL, ULAQ) are within 10% of the observed values of A_{195} (Table 1). E39CA and WACCM are slightly too high while several other models (CNRM-ACM, EMAC, UMETRAC, UMSLIMCAT) are slightly low. Of the remaining models, most simulate A_{195} values that are significantly lower than observed (e.g., UMUKCA [*Morgenstern et al.*, 2009]), although the values of CCSRNIES are 20% higher than observed. Overall, although a large number of models agree reasonably well with observations in this broad view, the timing of PSCs is likely to be slightly different in models than in the observations. Restricting the comparison to the August–September average, the observations increase by about 10% whereas most models remain about the same or decrease slightly. This implies that models tend to simulate more PSCs than observed in the winter when the impact on ozone is less, and, despite their typically late stratospheric warmings [e.g., *Eyring et al.*, 2006, Figure 2], PSCs tend to be underpredicted in the more important (for ozone) spring period. In most cases, the results for REF-B2 agree with the corresponding results for REF-B1. The main exceptions are WACCM which is slightly lower for REF-B2 and, for reasons that are not clear, CAM3.5 is somewhat lower for REF-B2 than for REF-B1. The UMUKCA models, by contrast have slightly higher A_{195} values for the REF-B2 experiment. EMAC-FUB also has slightly higher values for the REF-B2 run than the results of the sister model EMAC for REF-B1. This difference is due to the different vertical formulation of the models, the SSTs or the ozone amount due to the change in the model PSC scheme.

4.2. Ratio Between the Ozone Hole Area and the Low-Temperature Area

[26] In the presence of polar stratospheric clouds, halogen reservoir species are converted to active forms and ozone is depleted in subsequent sunlit conditions [e.g., *Solomon*, 1999]. It is therefore expected that the PSC region deline-

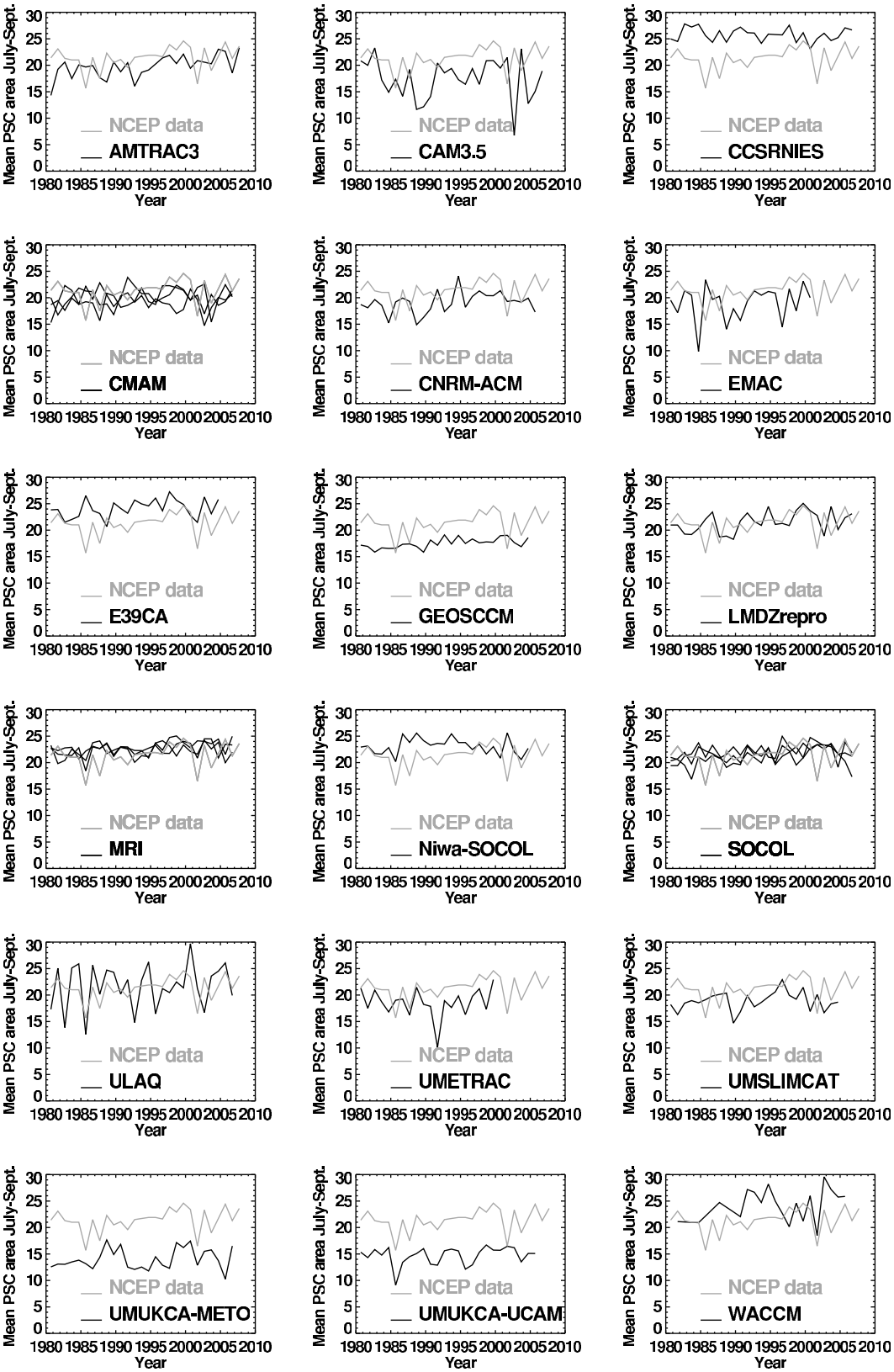

Figure 10. Southern Hemisphere area where the 50 hPa temperature is lower than 195K, averaged for the period July to September for each year from 1980. The simulations are from the REF-B1 experiment, and NCEP observations are shown in gray.

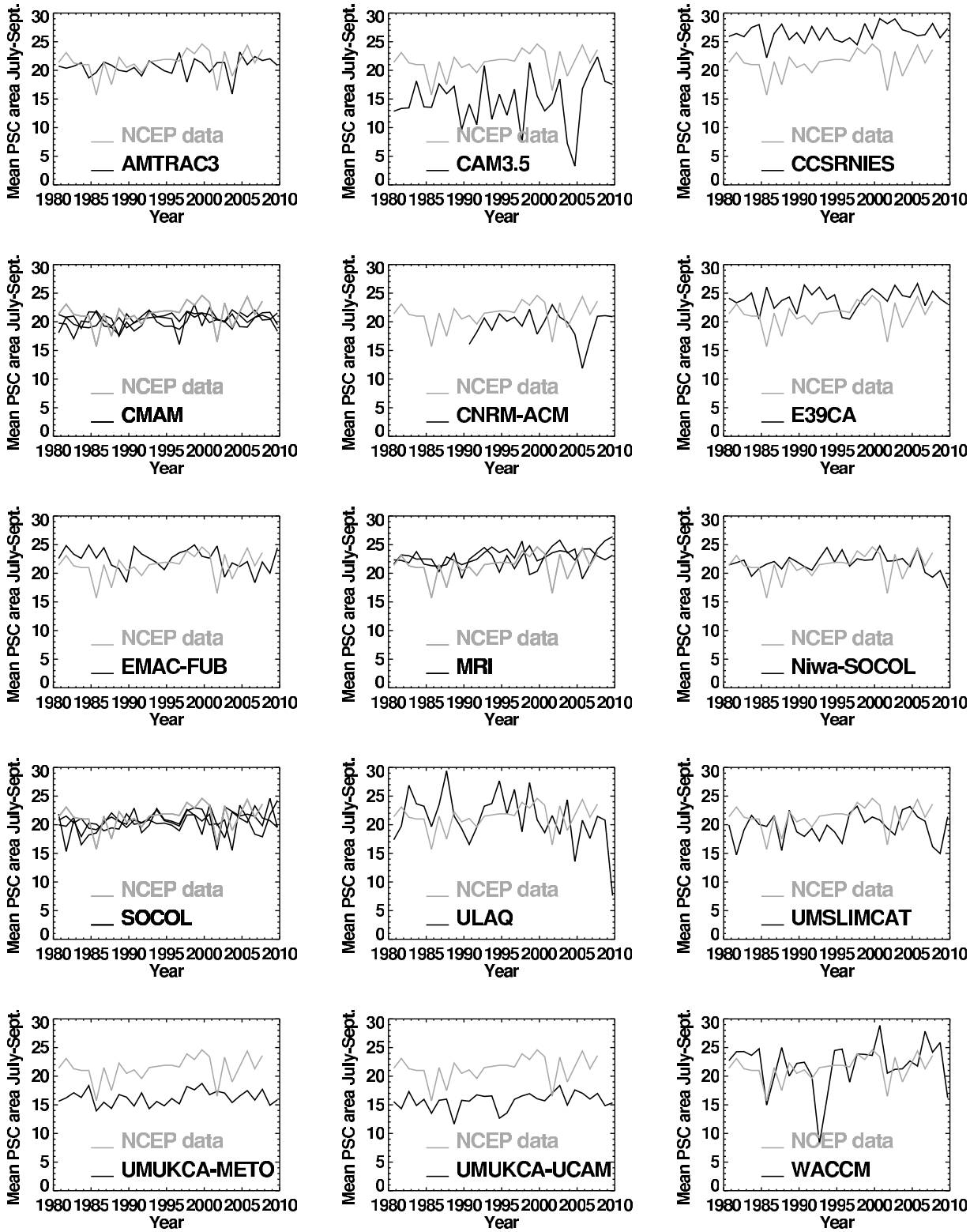

Figure 11. As in Figure 10, using the results of the simulations from experiment REF-B2.

ates the area of ozone destruction, which in turn is dependent on the model vortex structure [*Huck et al.*, 2007; *Tilmes et al.*, 2008; *Struthers et al.*, 2009]. The PSCs control the rate of change of ozone and hence the time integral of the PSCs (or equivalently their mean area) determine the ozone perturbation. Hence mean PSC area should be related to the size of the ozone hole. In practice, obtaining PSC areas from observations is difficult, and so instead, we use the approximate NAT areas indicated by A_{195}. Using NIWA and NCEP data, *Austin and Wilson* [2010] calculated the

ratio, Γ, of the maximum ozone hole area to the value of A_{195} averaged for July to September. Γ increased steadily from the 1980s as halogen amounts increased, and reached an asymptotic limit of about 1.2. Because of almost complete destruction in the lower stratosphere, the ozone hole did not change substantially in size from the 1990s [e.g., *Huck et al.*, 2007]. Figure 12 shows the results obtained for Γ for the REF-B1 simulations. The WACCM and CAM3.5 results were only available every 10 days instead of every day for the other models. Hence, the maximum ozone hole area shown earlier in the paper is biased low. AMTRAC3 results were compared at 10 and 1 day frequency and in the latter case, the results of the maximum were higher on average by about 3.6%. Therefore, in Figure 12 and subsequently WACCM and CAM3.5 results have been increased by 3.6% to provide a more consistent comparison.

[27] Half the models (AMTRAC3, CMAM, LMDZrepro, Niwa-SOCOL, SOCOL, ULAQ, UMETRAC, UMSLIMCAT and WACCM) agree within about 15% of the observations, but the others generally substantially underpredict the value of Γ and one model (CNRM-ACM) simulates a ratio somewhat higher than observed. For those models which supplied data, similar results were obtained for experiment REF-B2 (Figure 13). Small but notable differences are seen in AMTRAC3, which agrees better with observations in REF-B2 and Niwa-SOCOL and WACCM, which do not agree quite as well in REF-B2. These differences may relate to the actual simulation of PSCs, which as noted above depend on the H_2O concentration. In turn this may be determined by the differences in the forcing data set, that is, the SSTs. In comparisons of two of the sets of experiments shown here, *Garny et al.* [2009] and *Austin and Wilson* [2010] indicate a slight sensitivity of the ozone results to the SSTs, and both suggest that changes in the Brewer-Dobson circulation have an impact. Garny et al. attribute their differences to changes in mean SSTs, while the Austin and Wilson results arise from changes which may be related to the Niño 3.4 index.

[28] The results are summarized in Figures 14 and 15 where the cold areas and the ozone hole areas have been averaged for the period 1990 to 2008, or to the end of the simulation. This is a sufficiently long period to ensure enough data for statistical purposes, and starting late enough that there was sufficient chlorine present to produce almost complete ozone destruction in the lower stratosphere each southern spring. There are eight models (AMTRAC3, CMAM, LMDZrepro, Niwa-SOCOL, SOCOL, ULAQ, UMSLIMCAT, and WACCM) which provided good fits to the observations in Figure 12, and these models are seen to be significantly closer to the observations in Figure 14 than the other models. The ratio of the mean ozone hole area to the PSC area for the eight models is 1.14 compared with the observed ratio of 1.21 ± 0.05. Similar results are obtained for REF-B2 experiments for the ratio between ozone hole area and PSC area, which has a mean of 1.10 for seven models (the above 8 less LMDZrepro, for which data are not available for the relevant years).

[29] The other models have a variety of discrepancies from observations. CNRM-ACM has either a large low bias in ozone, or it is possible that the low-temperature region in the model is not representative of the actual PSC area which drives the chemistry. UMETRAC has a smaller ozone hole and a smaller cold area than observed and their ratio is similar to that observed. This suggests that the main problem is dynamical, primarily a 13% underprediction in the cold areas. The remaining models simulate a smaller ozone hole than low-temperature area. The CCSRNIES, E39CA and MRI models yield approximately the observed value of A_{195} but a small ozone hole, suggesting the presence of an ozone high bias. CAM3.5, EMAC, GEOSCCM, UMUKCA-METO and UMUKCA-UCAM have a combination of reduced values of A_{195} and an ozone high bias of varying degrees. The results are consistent with a high ozone bias identified in UMUKCA-METO (or equivalently UMUKCA-UCAM), MRI and CCSRNIES (Figure 1) as well as GEOSCCM [*Pawson et al.*, 2008] and E39CA [*Loyola et al.*, 2009]. Because of the sensitivity of the PSCs to temperature, the model thermal structure could have an important impact on the ozone loss and hence on the vertical extent of the ozone loss. This is considered in section 5.

5. High-Latitude Vertical Distribution of Ozone and Chlorine

5.1. Regional Mean Ozone 60°S–90°S

[30] The vertical distribution of decadally averaged ozone in the high latitudes in October is shown in Figure 16 for the observations [*Randel and Wu*, 2007] and each of the models which contributed to CCMVal-2. In Figure 16, the area between the curve and the ordinate is proportional to the ozone column. A wide range of results is obtained which can be used to put the previous results into context. One group of models provides a realistic simulation of the differences between the decades of the 1970s and 1990s, and agree reasonably with observations (AMTRAC3, CMAM, E39CA, GEOSCCM, LMDZrepro, MRI, UMETRAC, UMSLIMCAT, and WACCM). Another group (CAM3.5, CCSRNIES, EMAC, UMUKCA-METO, and UMUKCA-UCAM) have shallow ozone holes. The remaining models (CNRM-ACM, the SOCOL models and ULAQ) have a mixture of discrepancies with observations, although some features of the observations are reproduced. For example, for Niwa-SOCOL and SOCOL, the 1990s ozone hole is generally simulated, but the 1970s ozone amounts are too low. Other models (CNRM-ACM and ULAQ) are too low below 100 hPa in particular. This would appear to explain in part the excessive ozone holes (measured using 220 DU column) for CNRM-ACM. For the more successful models, the region of loss is generally extended more in the vertical than observed. For example, for AMTRAC3, CMAM, LMDZrepro, UMSLIMCAT and WACCM, the region of loss for the 1990s extends to 20 hPa, compared with about 30 hPa in the observations. Most of these models have less ozone than observed below 100 hPa, while several models (e.g., MRI, ULAQ, UMETRAC) simulate the ozone hole well in the upper levels, but extend the loss to below 100 hPa. Of these model runs, for two of the models (AMTRAC3 and WACCM) the temperatures over the south polar cap have been examined and found to be biased low, leading to a vertical extension of the PSC region.

5.2. Ozone at 90°S

[31] To a good approximation, values at the South Pole are representative of the core of the ozone hole and are shown in Figure 17, which illustrates values for October in

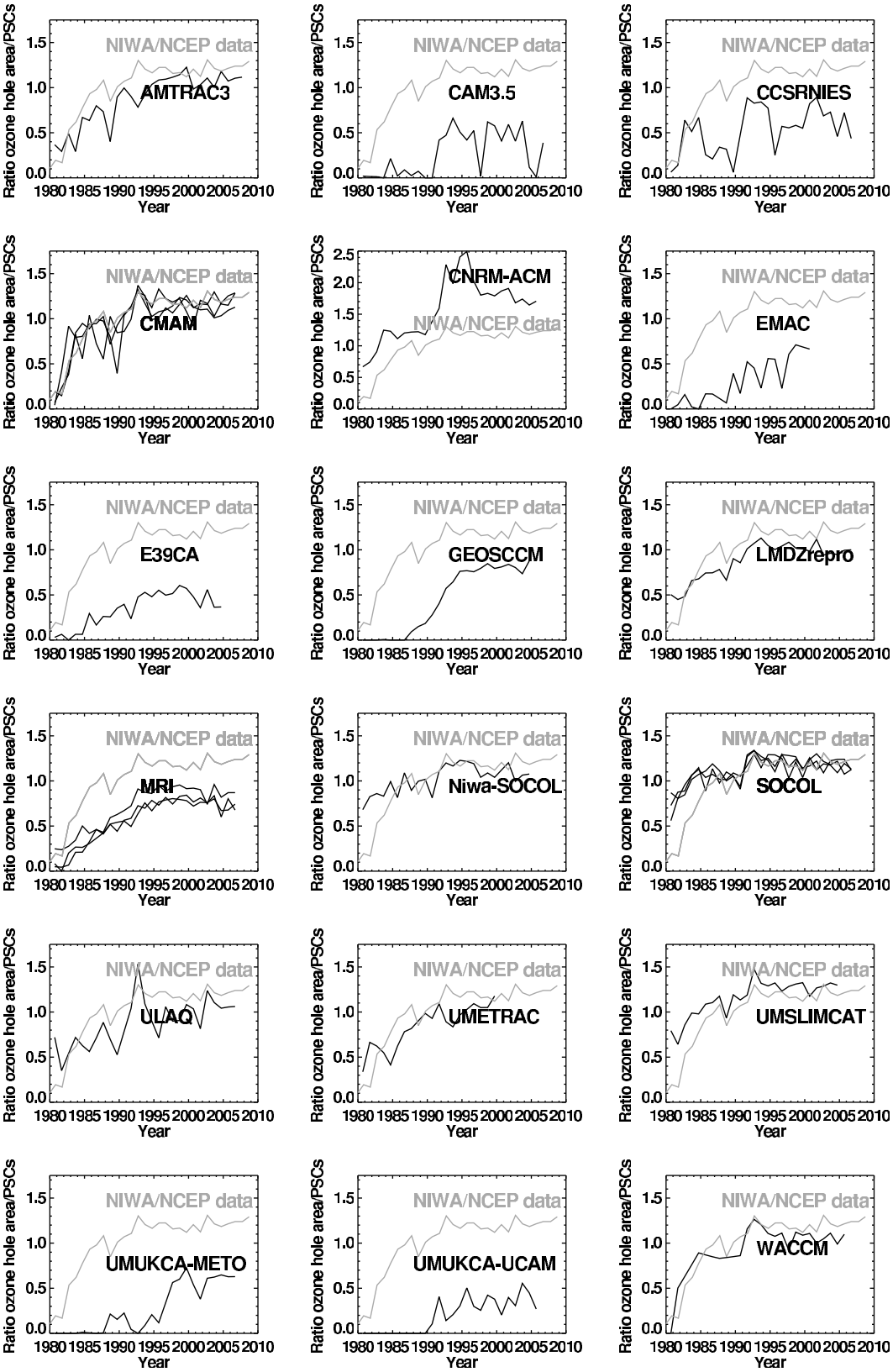

Figure 12. Ratio, Γ, between the maximum size of the ozone hole each year and the mean model $T <$ 195 K areas for July to September on the 50 hPa surface. The results were calculated from the REF-B1 simulations and are compared with values derived from NIWA and NCEP data. The ozone hole areas for WACCM and CAM3.5 have been increased by 3.6% to correct for the 10 day frequency of the output (see text). Note the change in ordinate scale for the results of CNRM-ACM.

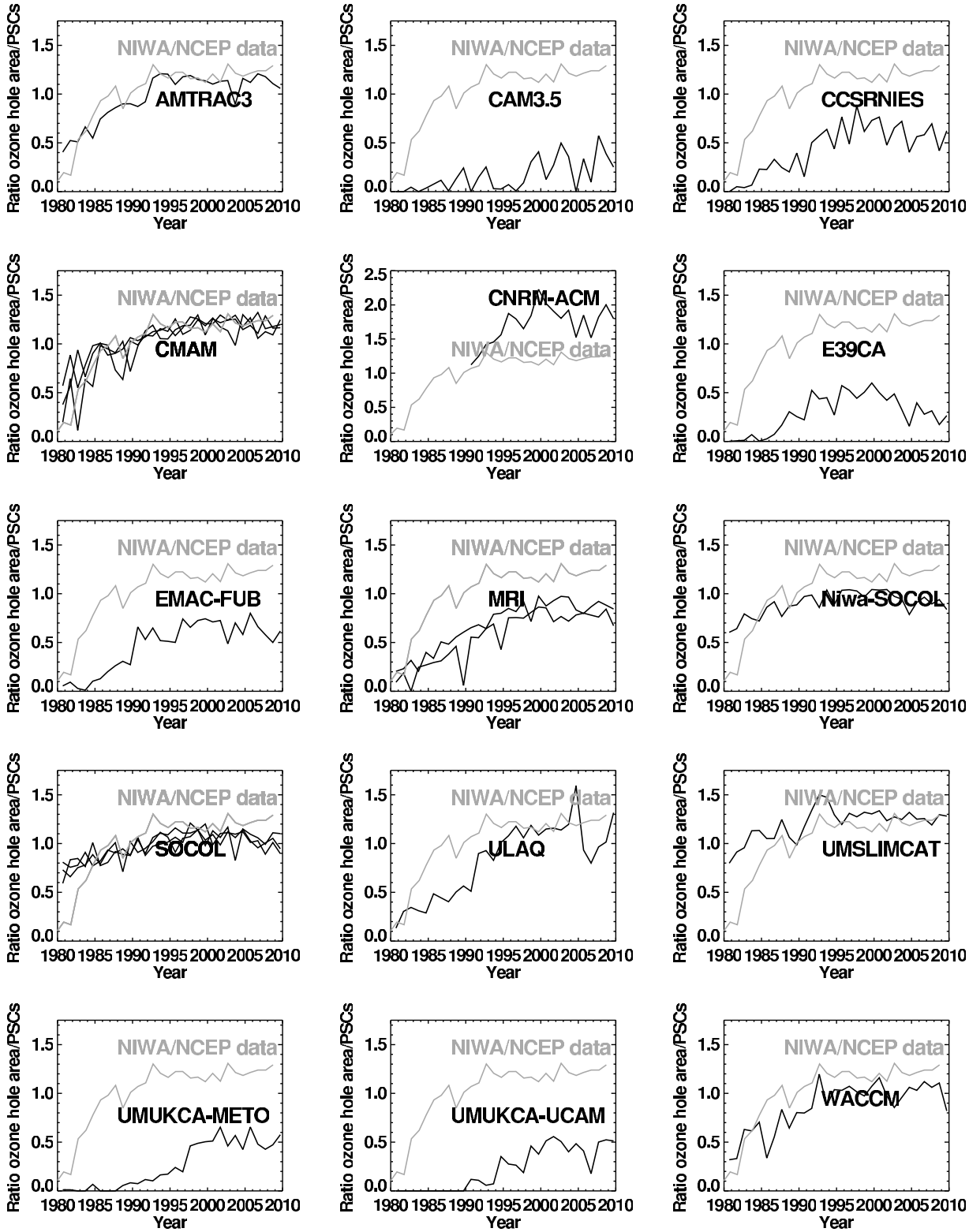

Figure 13. As in Figure 12 but for the REF-B2 simulations.

1979 (solid curves) and 1999 (dashed curves). In many models (shown by the red lines), the ozone hole is clearly established by 1999 with ozone amounts close to zero in the lower stratosphere, consistent with measurements [*Solomon et al.*, 2005]. The ozone profiles for the earlier period also in many cases show some reduction compared with the *Randel and Wu* [2007] climatology, suggesting that many models depleted ozone earlier than was measured or have a significant ozone bias in their nonperturbed ozone chemistry. It should be noted, however, that in the Randel and Wu

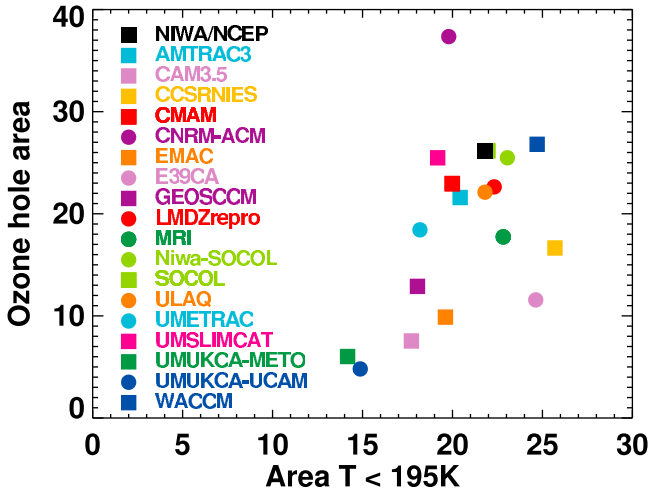

Figure 14. Ozone hole area versus A_{195} (both in units of 10^6 km^2) for each model compared with observations. The results were calculated from the REF-B1 simulations, and are averaged for the period 1990–2008. For those models which performed multiple simulations, the ensemble means are shown. The ozone hole areas for WACCM and CAM3.5 have been increased by 3.6% to correct for the 10 day frequency of the output (see text).

database, the vertical ozone distribution over Antarctica below 25 km is based primarily on Syowa ozonesonde data and as such may underestimate the severity of the ozone depletion deep inside the vortex.

5.3. Chlorine Amounts

[32] The connection between chlorine amounts and ozone are shown for these model results by *Austin et al.* [2010] and *Oman et al.* [2010]. In Figure 18 we show the vertical profiles of the concentrations of the chlorine reservoir species and active chlorine at the same years and location as the ozone profiles in Figure 17. Recognizing the time taken for chlorine to deplete ozone, the chlorine species are shown for one month earlier than the ozone results. Compared with the previous CCMVal experiments [*Eyring et al.*, 2006] there is less variation between the model Cl_y results, although in the region of steep vertical gradients near 100 hPa, the results cover a wide range. Nonetheless, the individual species vary according to the model, reflecting the different amounts of chlorine activation and the different altitudes at which activation takes place. In particular for UMUKCA-METO, there are insufficient PSCs present, and the inorganic chlorine appears largely as $ClONO_2$ peaking at 30 hPa. The main peak in active chlorine occurs just above 100 hPa whereas most models peak between 20 and 50 hPa. Even then there is a factor of 2 range of peak active chlorine both in 1979 and 1999. The models that performed well for the total ozone column (Figure 5) have high chlorine levels for 1999, except SOCOL (and Niwa-SOCOL), which has active chlorine about half that of most of the other models. It would therefore seem that the column ozone behavior shown in the previous figures is influenced less by chlorine chemistry and more by dynamics. This would tend to make this particular model less effective for future predictions.

6. Ozone Recovery

[33] One of the many purposes of long simulations of stratospheric ozone is to determine the timing of ozone recovery. Here, we refer to recovery as the process of ozone increase and the date by which a given ozone column is attained as a return date or recovery date. To put the Antarctic results into global context, we first show in Figure 19 the ozone return date, as a function of latitude and reference year, averaged across all the models which provided the column ozone results for experiment REF-B2. Results are included for all the models except E39CA, which finished too early (2050). In the annual mean, ozone recovery occurs over most of the globe. In the tropics, ozone returns to 1983 values by the middle of the century, but the increased strength of the Brewer-Dobson circulation prevents return to earlier, and higher values [*Shepherd*, 2008; *Waugh et al.*, 2009]. From about 2050 onward, the simulated tropical ozone columns decrease slightly. Although it is not possible to determine an ozone return date from the mean model results, the tropical column ozone change is only about 10 DU. In the middle- and high-latitude Northern Hemisphere the model mean returns to 1980 values as soon as 2025, compared with about 2050 for the stratospheric halogen loading. In the Southern Hemisphere, the high latitudes return to 1980 values on a similar time scale as the halogen loading, while the midlatitudes return about a decade earlier hence a substantial interhemispheric asymmetry is simulated in recovery time scale. This is likely to be due to higher ozone transport into the Northern Hemisphere possibly due to the trend in the Brewer-Dobson circulation [*Austin and Wilson*, 2006; *Eyring et al.*, 2007; *WMO*, 2007, chapter 6; *Shepherd*, 2008; *Waugh et al.*, 2009].

[34] Recovery of ozone back to mid 1960 values occurs in Southern Hemisphere middle latitudes by 2060, but at the south pole, this time scale is extended beyond 2080. There is little seasonal variation in these results except for the results for southern spring (Figure 19, top) which show an even later

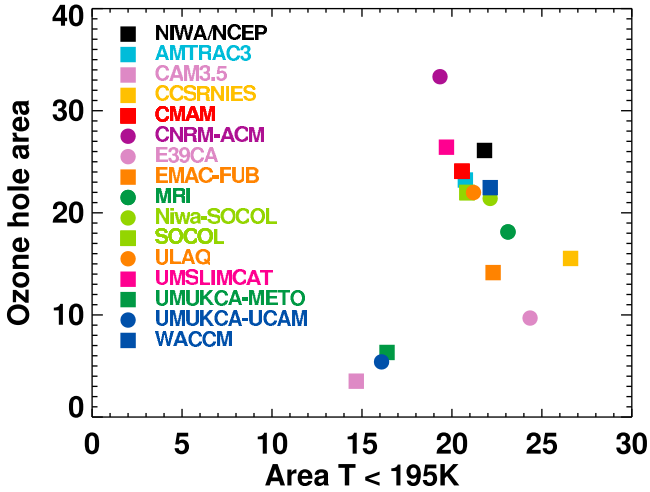

Figure 15. As in Figure 14 but for the results of Experiment REF-B2.

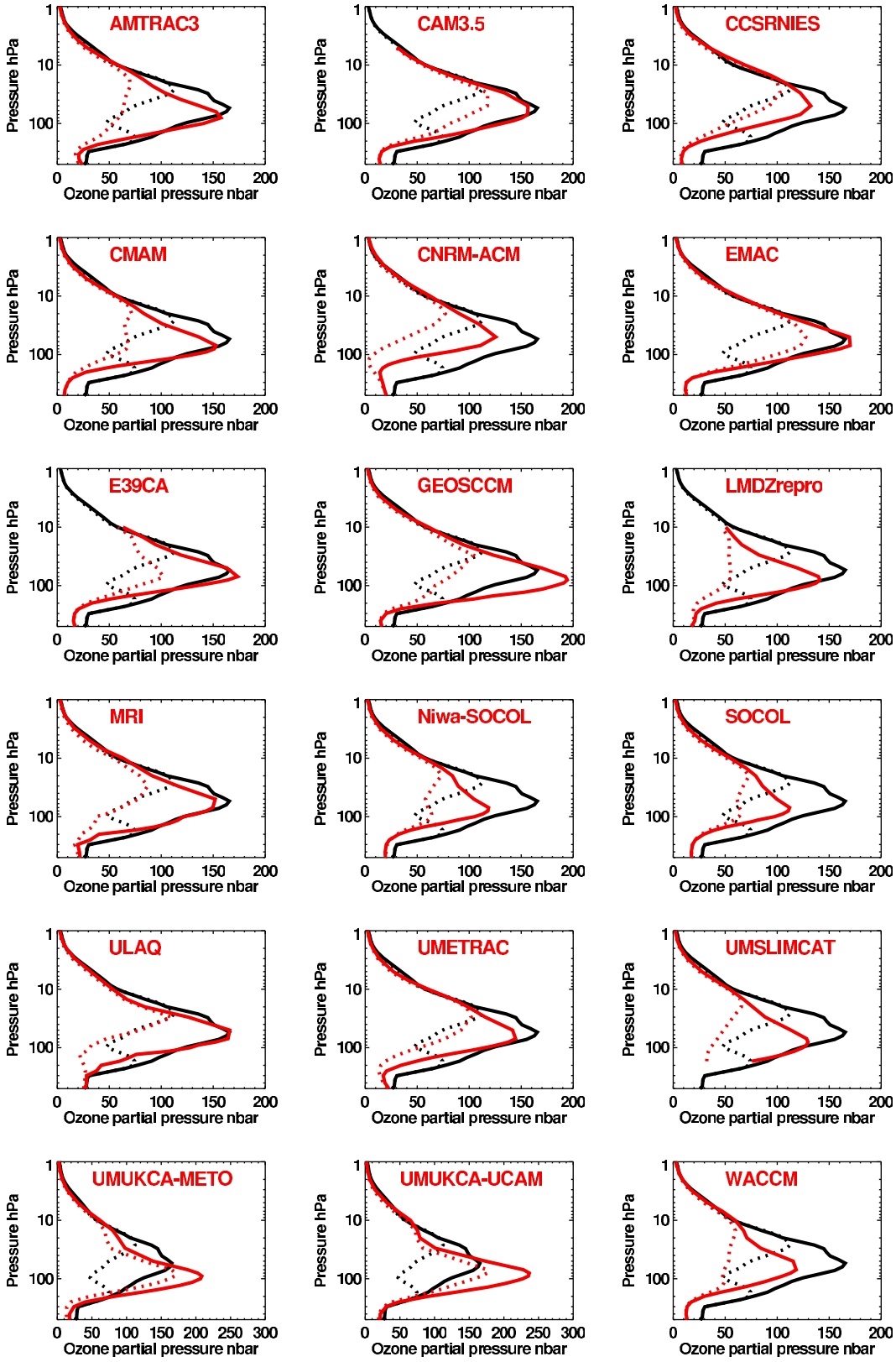

Figure 16. Vertical profile of October mean ozone partial pressure averaged over the latitude range 60°S–90°S for 1971–1979 (solid lines) and 1991–1999 (dashed lines). For those models which supplied more than one realization, the ensemble means are plotted. The black lines are taken from climatology [*Randel and Wu*, 2007], although for the earlier period, data from just the single year 1979 are shown. Note the change in scale on the abscissa for the UMUKCA pair. The results are from the REF-B1 simulations.

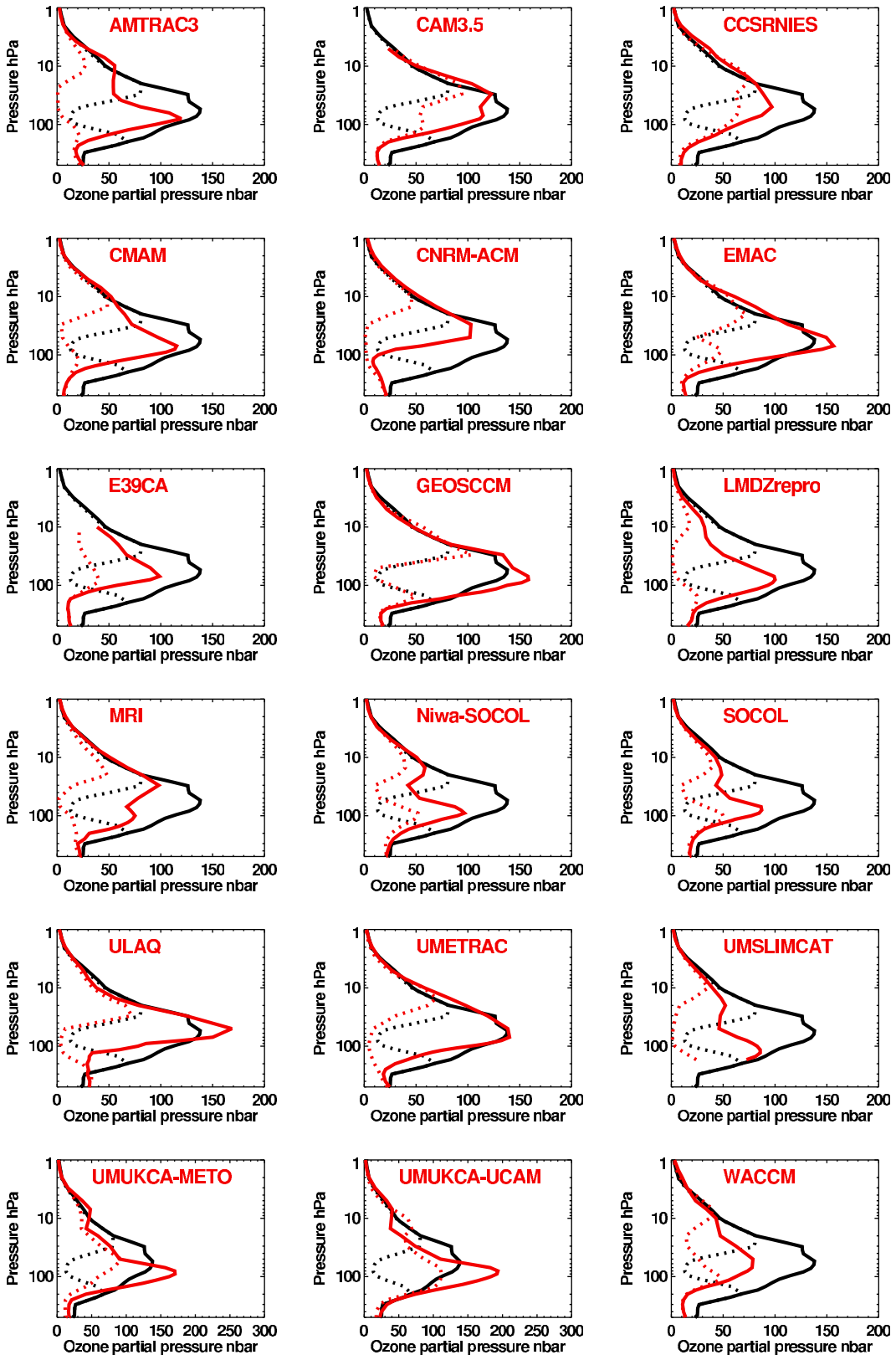

Figure 17. As in Figure 16 but for the South Pole and for the single years 1979 (solid curves) and 1999 (dashed curves). For those models which supplied more than one realization, only the first of the ensemble members is plotted so that the results are representative of the core of the ozone hole.

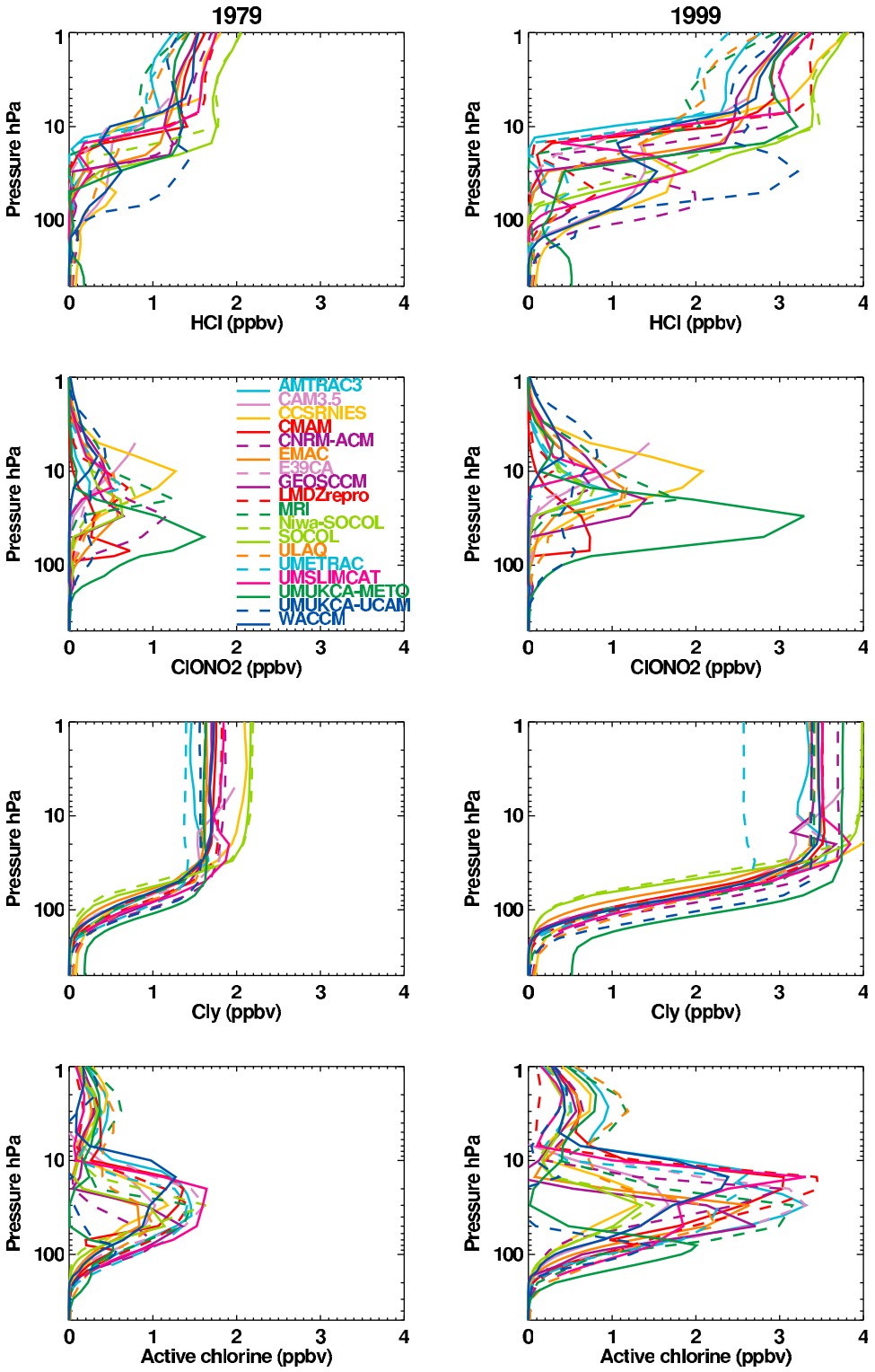

Figure 18. Vertical profiles of HCl, $ClONO_2$, Cl_y, and active chlorine = Cl_y − HCl − $ClONO_2$ for September for the South Pole for the REF-B1 model simulations. Results are plotted for (left) 1979 and (right) 2000. As in Figure 17, CAM3.5, EMAC, and UMETRAC results are shown for 1999. For those models which supplied more than one realization, only the first of the ensemble members is plotted so that the results are representative of the core of the ozone hole.

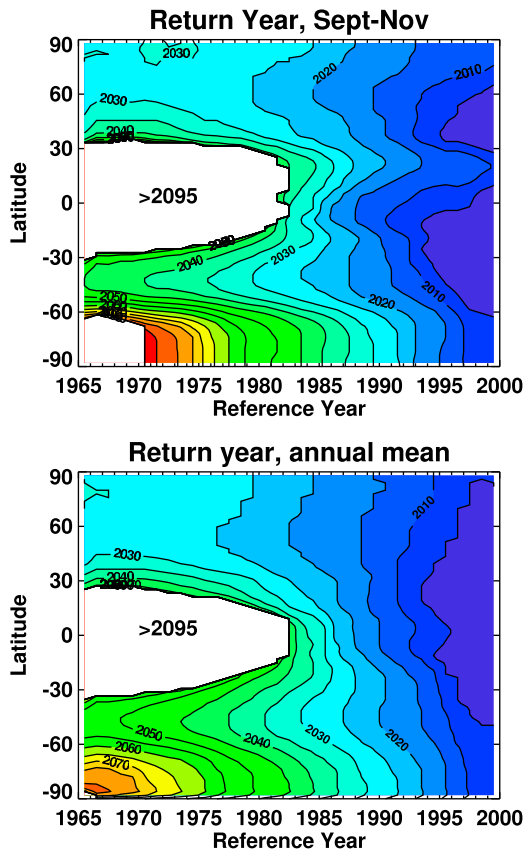

Figure 19. Date of return of total column ozone to the value present at the reference year. Results were taken from the 16 models which contributed results for the REF-B2 simulations at least to the year 2080. The model results were averaged with equal weighting and then smoothed with an 11 year running mean filter. Data prior to 1965 (which limits the definition of the reference year) or after 2094 (which limits the recovery year data) do not exist because of the need for an accurate time smoothed field. The white regions indicate where the mean model has not recovered by the end of the simulations (effectively 2094). (top) Results for the southern spring, and (bottom) results for the annual average.

recovery in Antarctic ozone. This follows from the results shown earlier, in which the ozone hole in some simulations continued until the end of the integrations. Over Antarctica there is also some sensitivity of the results obtained for the annual average, depending on the precise mix of models included.

[35] The individual models show a wide range of results. Figure 20 illustrates the annual mean ozone return year as a function of reference year for each individual model, averaged over the latitude range 60°S–90°S. As in Figure 19, results are presented for the spring (Figure 20, top) and annual average (Figure 20, bottom). CCSRNIES and UMUKCA-UCAM have weak ozone holes (measured by the 220 DU threshold), and recover early. By contrast, GEOSCCM also has an ozone hole which is weaker than average, but it recovers late. Return to 1980 values occurs between 2020 and 2080, but most models lie between 2030 and 2065, and the return to 1970 values is generally simulated to occur about 15 years later. However, some of the models diverge further as the reference year reduces and several models indicate recovery to pre-1975 Antarctic values beyond the end of the simulation, as indicated in Figure 19.

7. Discussion

[36] The model results are sensitive to a large number of factors which contribute to the stratospheric chemistry, radiation and dynamics in the models. Many of these, for example the halogen loadings, GHG concentrations and the SSTs were specified in large part by the experimental design. In the case of the SSTs, some sensitivity of the results could be determined by comparison between the REF-B1 and REF-B2 simulations and this has been described in the above results in a few cases. Several other factors, which we explore here, were not specified and tended to vary widely amongst the models.

7.1. Parameterized Gravity Wave Drag

[37] One of the most challenging aspects of modeling the dynamics of the middle atmosphere remains the

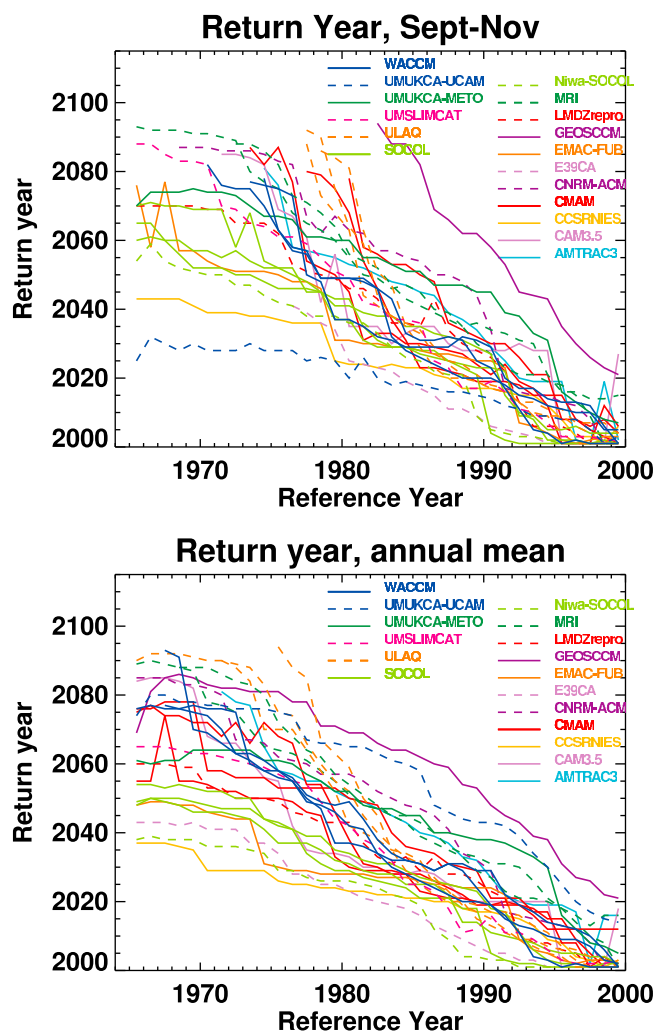

Figure 20. Date of return of total column ozone, averaged over 60°S–90°S, to the value present at the reference year, for each individual model. (top) Results for the southern spring, and (bottom) results for the annual average.

parameterization of gravity wave drag (GWD). These schemes assume that the sources are highly simplified and are columnar in their formulation ignoring any lateral propagation of the waves. While their simplicity means that optimal stratospheric climates can be tuned in a relatively straightforward manner, the resulting climate will also involve a variety of trade-offs. For example, in CMAM it is found that adjusting the orographic GWD parameterization can lead to improved lower stratospheric polar temperatures that satisfy PSC thresholds. However, this improvement comes at the expense of degraded temperatures in the upper stratosphere and troposphere and increased mean sea-level pressure biases [*Scinocca et al.*, 2008].

[38] Even the simplest nonorographic scheme is generally preferred over the use of Rayleigh friction as its proxy. There is now sufficient evidence to conclude that this practice is undesirable in the context of both climate [e.g., *Shaw et al.*, 2009] and weather forecasting [*Orr et al.*, 2010]. So, while GWD parameterization remains a difficult problem, there do exist avenues of immediate improvement in two of the CCMVal models. Nonetheless, in these simpler models it is possible to simulate accurate lower stratospheric temperatures, within certain limits, by adjusting their Rayleigh friction and diffusion parameters.

7.2. Heterogeneous Reactions

[39] The existence of the ozone hole is critically dependent on the presence of surface reactions to activate chlorine and bromine from the reservoir species [*Solomon*, 1999]. The different models of CCMVal use a variety of heterogeneous reaction schemes, although they have common elements as noted in section 2. Whether the surfaces concerned are NAT like particles or liquid aerosols remains an important scientific debate [e.g., *Tilmes et al.*, 2007]. In a Chemistry Transport Model where external parameters (temperature and winds) are explicitly defined by observations, differences in assumptions can have important consequences [e.g., *Krämer et al.*, 2003; *Santee et al.*, 2008]. Nonetheless, in simulating the ozone hole in a climate model, the first objective is to activate the halogens and some parameterizations can be adjusted accordingly (e.g., GWD). Many of the above studies are applicable mainly to the Arctic where PSC formation is more transient. Although the rates of change of ozone may vary according to the heterogeneous rates adopted, in the Antarctic, the rates are generally sufficiently fast in many models to provide high levels of chlorine activation in the presence of low temperatures, as shown in Figure 18. Unfortunately, no sensitivity tests were performed with different PSC or heterogeneous reaction rate assumptions. Nonetheless, similar ozone hole results were simulated by, for example AMTRAC3 and CMAM despite the absence of NAT in the latter model. Many models assume the presence of both NAT and STS [*Morgenstern et al.*, 2010]. *Pitts et al.* [2007] have shown that CALIPSO observations in the Antarctic indicate a PSC area that is significantly smaller than what would be inferred from the commonly used temperature-based proxy T_{NAT}, but which is similar in magnitude to that inferred from T_{STS}. Unfortunately, these values were not archived on the CCMVal database. Hence, although there is undoubtedly some sensitivity of the results to heterogeneous reaction rate assumptions, the accurate simulation of lower stratospheric temperature is currently the biggest challenge for CCMs in simulating the ozone hole.

7.3. Sensitivity of the Results to Water Vapor Amounts

[40] A further sensitivity of the ozone hole could arise from the water vapor amounts prior to the formation of PSCs [*Stenke and Grewe*, 2005]. In the Antarctic stratosphere the coldest time of year occurs in June or July and if there is too much condensation and settling of water vapor at this time, there may be insufficient H_2O to form PSCs in the spring when ozone loss occurs in the atmosphere. There may be other reasons for low H_2O concentrations, such as a tropical tropopause cold bias. The objective of this paper has not been to validate model water vapor concentrations, which is beyond the scope of the current work. However, a brief analysis indicates that for the South Pole in December at the end of the REF-B1 simulations, 16 of the 18 models simulated H_2O concentrations in the range 4.7 to 6.4 ppmv at 10 hPa, before significant condensation will have occurred. All else being equal, including 10 ppbv HNO_3, this range in water vapor would result in a NAT formation temperature in the (small) range 195.5–196.8 K at 50 hPa. By contrast, two of the models, CCSRNIES and LMDZrepro, simulated water vapor amounts at 10 hPa of only 1.9 and 2.1 ppmv, respectively, which correspond to a NAT formation temperature of about 192 K, significantly lower than the above range. While this might explain the small ozone hole for CCSRNIES despite realistic A_{195} values (see Figure 14), LMDZrepro has a robust ozone hole and must therefore be making different approximations in the cloud physics.

8. Conclusion

[41] Simulations of the Antarctic ozone hole have been investigated for a set of experiments completed for the Chemistry-Climate Model Validation project. The results cover a wide range and many models agree broadly with measurements, but typically the ozone hole is too small in area and the ozone mass deficit is too small. Although individual models have in some cases undergone major improvements, overall there have been few improvements in ozone hole statistics since the work by *Eyring et al.* [2006] and *WMO* [2007] in comparison with measurements.

[42] In this paper, simple diagnostics have been presented and compared with observations. Comparison of model results relative to the values which were simulated for the 1960–1965 period identified some models with clear ozone biases. Using the area less than 195K for the 50 hPa temperature (A_{195}, a proxy for the nitric acid trihydrate polar stratospheric cloud areas) it was found that many models underpredicted the observations. In those models which underpredicted the cold areas, the steepest ozone meridional gradients (which signify the edge of the ozone hole) were also typically found to be poleward of the steepest ozone gradients in the observations. For the experiments REF-B1 and REF-B2, which differ primarily in the specification of the sea surface temperatures, the A_{195} values were generally very similar. Those models which performed well regarding the ozone hole area, typically underpredicted the Antarctic ozone minimum, and those models which had a small area also tended to overpredict the minimum. As a result, the ozone mass deficit relative to 220 DU, covered an extremely

wide range in the models. For those models which were most successful in reproducing the observed ozone hole size, additional loss tended to occur above and below the observed ozone hole region. On the basis of the mean of all the model results, ozone will likely recover later over Antarctica than the rest of the atmosphere, apart from the tropics, as halogen levels decrease. Individually, though, the models give a wide range of recovery dates. For the return of the polar cap ozone to 1980 values, typical dates are 2030–2065 and for the return to 1970 values the approximate dates are 2045–2080. However, for several models which agree best with observations, the recovery time scale is at the upper end of the range given, and in the Antarctic spring, return to 1970 values does not occur before the end of the simulations.

[43] The causes of the discrepancies from observations for the ozone hole area could be identified in most models. In several cases the discrepancy is due simply to a small area of low temperatures for PSC formation and subsequent surface reactions. Several other models differ from observations due to their background ozone climatology being too high, while in one case the simulated ozone had a low bias. Whether an individual model has an ozone bias, or low A_{195} value (and hence low PSC amounts) its results can represent a significant deficiency in comparison with measurements. In principle, it is possible to extrapolate the results to allow for discrepancies such as a small vortex size. However, the coupling between chemistry and temperature in the models, which for heterogeneous chemistry is a positive feedback process, would make the procedure subjective.

[44] The results imply the need for better dynamical simulations before the ozone hole can be properly simulated by many models. This is urgent as the ozone hole is seen as a proxy for ozone depletion in general and simulations of ozone recovery will depend on the quality of the ozone hole simulations. The fidelity of the Antarctic vortex is crucial to the accurate simulation of the ozone hole, as shown in previous work [*Huck et al.*, 2007; *Tilmes et al.*, 2007; *Struthers et al.*, 2009]. Until further advances are made in GWD parameterization, discrepancies between observations and model simulations of the Antarctic vortex will almost certainly remain. Currently, those models which by design or fortune simulated accurate lower stratospheric dynamics also typically simulated accurate ozone hole behavior for the current atmosphere. Even then, the fact that the simulated column minima are typically attained 10 days after the observations may have been due to a possible cold bias in the middle and upper stratosphere. Nonetheless, these models would be expected to be the most reliable for predicting the disappearance of the ozone hole as halogen levels subside. Those models suggest that by end of the 21st century an ozone hole of several million km^2 is still expected to be occurring, based on the canonical 220 DU amount.

[45] **Acknowledgments.** CCSRNIES research was supported by the Global Environmental Research Fund of the Ministry of the Environment of Japan (A-071). The MRI and CCSRNIES simulations were completed with the supercomputer at the National Institute for Environmental Studies, Japan. CMAM simulations were supported by CFCAS through the C-SPARC project. The computer time for the EMAC-FUB simulation at ECMWF was provided by the German Weather Service. The Niwa-SOCOL and UMETRAC simulations were supported by the New Zealand Foundation for Research, Science and Technology under contract C01X070. The UMSLIMCAT work was supported by NERC. The contribution of the Met Office Hadley Centre was supported by the Joint DECC and Defra Integrated Climate Programme, DECC/Defra (GA01101). The scientific work of the European CCM groups was supported by the European Commission through the project SCOUT-O3 under the 6th Framework Programme. J.A.'s research was administered by the University Corporation for Atmospheric Research at the NOAA Geophysical Fluid Dynamics Laboratory. John Wilson and Rolando Garcia provided useful comments on the paper. We acknowledge the Chemistry-Climate Model Validation (CCMVal) Activity for WCRP's (World Climate Research Programme) SPARC (Stratospheric Processes and their Role in Climate) project for organizing and coordinating the model data analysis activity and the British Atmospheric Data Center (BADC) for collecting and archiving the CCMVal model output.

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.

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

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.

Austin, J., and R. J. Wilson (2010), Sensitivity of polar ozone to sea surface temperatures and halogen amounts, *J. Geophys. Res.*, *115*, D18303, doi:10.1029/2009JD013292.

Austin, J., et al. (2003), Uncertainties and assessments of chemistry-climate models of the stratosphere, *Atmos. Chem. Phys.*, *3*, 1–27.

Bodeker, G. E., H. A. Struthers, and B. J. Connor (2002), Dynamical containment of Antarctic ozone depletion, *Geophys. Res. Lett.*, *29*(7), 1098, doi:10.1029/2001GL014206.

Bodeker, G. E., H. Shiona, and H. Eskes (2005), Indicators of Antarctic ozone depletion, *Atmos. Chem. Phys.*, *5*, 2603–2615.

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*(D21), 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.

Eyring, V., et al. (2005), A strategy for process-oriented validation of coupled chemistry-climate models, *Bull. Am. Meteorol. Soc.*, *86*, 1117–1133.

Eyring, V., et al. (2006), Assessment of coupled chemistry-climate models: Evaluation of dynamics, transport and ozone, *J. Geophys. Res.*, *111*, D22308, doi:10.1029/2006JD007327.

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

Eyring, V., M. P. Chipperfield, M. A. Giorgetta, D. E. Kinnison, E. Manzini, K. Matthes, P. A. Newman, T. G. Shepherd, and D. W. Waugh (2008), Overview of the new CCMVal reference and sensitivity simulations in support of the upcoming ozone and climate assessments and the planned SPARC CCMVal, *SPARC Newsl.*, *30*, 20–26.

Farman, J. C., B. G. Gardiner, and J. D. Shanklin (1985), Large losses of total ozone in Antarctica reveal seasonal ClO_x/NO_x interaction, *Nature*, *315*, 207–210.

Garcia, R. R., D. R. Marsh, D. E. Kinnison, B. A. Boville, and F. Sassi (2007), Simulation of secular trends in the middle atmosphere, 1950–2003, *J. Geophys. Res.*, *112*, D09301, doi:10.1029/2006JD007485.

Garny, H., M. Dameris, and A. Stenke (2009), Impact of prescribed SSTs on climatologies and long term trends in CCM simulations, *Atmos. Chem. Phys.*, *9*, 6017–6031.

Hanson, D. R., and K. Mauersberger (1988), Laboratory studies of nitric acid trihydrate: Implications for the south polar stratosphere, *Geophys. Res. Lett.*, *15*, 855–858.

Huck, P. E., S. Tilmes, G. E. Bodeker, W. J. Randel, A. J. McDonald, and H. Nakajima (2007), An improved measure of ozone depletion in the Antarctic stratosphere, *J. Geophys. Res.*, *112*, D11104, doi:10.1029/2006JD007860.

Hurwitz, M. M., P. A. Newman, F. Li, L. D. Oman, O. Morgenstern, P. Braesicke, and J. A. Pyle (2010), Assessment of the breakup of the Antarctic polar vortex in two new chemistry-climate models, *J. Geophys. Res.*, *115*, D07105, doi:10.1029/2009JD012788.

Intergovernmental Panel on Climate Change (IPCC) (2001), *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report*, edited by J. T. Houghton et al., 881 pp., Cambridge Univ. Press, Cambridge, U. K.

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.

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.

Krämer, M., et al. (2003), Intercomparison of stratospheric chemistry models under polar vortex conditions, *J. Atmos. Chem.*, *45*, 51–77, doi:10.1023/A:1024056026432.

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.

Loyola, R., M. Coldewey-Egbers, M. Dameris, H. Garny, A. Stenke, M. van Roozendael, C. Lerot, D. Balis, and M. Koukouli (2009), Global long-term monitoring of the ozone layer—A prerequisite for predictions, *Int. J. Remote Sens.*, *30*, 4295–4318, doi:10.1080/01431160902825016.

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 present-generation stratospheric chemistry-climate models and associated external forcings, *J. Geophys. Res.*, *115*, D00M02, doi:10.1029/2009JD013728.

Müller, R., J.-U. Grooß, C. Lemmen, D. Heinze, M. Dameris, and G. Bodeker (2008), Simple measures of ozone depletion in the polar stratosphere, *Atmos. Chem. Phys.*, *8*, 251–264.

Newman, P. A., J. S. Daniel, D. W. Waugh, and E. R. Nash (2007), A new formulation of equivalent effective stratospheric chlorine (EESC), *Atmos. Chem. Phys.*, *7*, 4537–4552.

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

Orr, A., P. Bechtold, J. F. Scinocca, M. Ern, and M. Janiskova (2010), Improved middle atmosphere climate and forecasts in the ECMWF Model through a non-orographic gravity wave drag parametrization, *J. Clim.*, in press.

Pawson, S., K. Krüger, R. Swinbank, M. Bailey, and A. O'Neill (1999), Intercomparison of two stratospheric analyses: Temperatures relevant to polar stratospheric cloud formation, *J. Geophys. Res.*, *104*(D2), 2041–2050, doi:10.1029/98JD02279.

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 between 1950 and 2005, *J. Geophys. Res.*, *113*, D12103, doi:10.1029/2007JD009511.

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*, 414–420.

Pitts, M. C., L. W. Thomason, L. R. Poole, and D. M. Winker (2007), Characterization of polar stratospheric clouds with spaceborne lidar: CALIPSO and the 2006 Antarctic season, *Atmos. Chem. Phys.*, *7*, 5207–5228.

Randel, W. J., and F. Wu (2007), A stratospheric ozone profile data set for 1979–2005: Variability, trends, and comparisons with column ozone data, *J. Geophys. Res.*, *112*, D06313, doi:10.1029/2006JD007339.

Sander, S. P., et al. (2006), Chemical kinetics and photochemical data for use in atmospheric studies: Evaluation No. 15, *JPL Publ.*, *06-2*, 522 pp.

Santee, M. L., et al. (2008), A study of stratospheric chlorine partitioning based on new satellite measurements and modeling, *J. Geophys. Res.*, *113*, D12307, doi:10.1029/2007JD009057.

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. A. Plummer (2008), Technical note: The CCCma third generation AGCM and its extension into the middle atmosphere, *Atmos. Chem. Phys.*, *8*, 7055–7074.

Shaw, T. A., M. Sigmond, T. G. Shepherd, and J. F. Scinocca (2009), Sensitivity of simulated climate to conservation of momentum in gravity wave drag parameterization, *J. Clim.*, *22*, 2726–2742.

Shepherd, T. G. (2008), Dynamics, stratospheric ozone, and climate change, *Atmos. Ocean*, *46*, 371–392.

Shibata, K., and M. Deushi (2008), 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.

Solomon, S. (1999), Stratospheric ozone depletion: A review of concepts and history, *Rev. Geophys.*, *37*, 275–316, doi:10.1029/1999RG900008.

Solomon, S., R. W. Portmann, T. Sasaki, D. J. Hofmann, and D. W. J. Thompson (2005), Four decades of ozonesonde measurements over Antarctica, *J. Geophys. Res.*, *110*, D21311, doi:10.1029/2005JD005917.

Stenke, A., and V. Grewe (2005), Simulation of stratospheric water vapour trends: Impact on stratospheric ozone chemistry, *Atmos. Chem. Phys.*, *5*, 1257–1272.

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

Struthers, H., K. Kreher, J. Austin, R. Schofield, G. Bodeker, P. Johnston, H. Shiona, and A. Thomas (2004), Changes in the rate of increase of NO_2 as predicted by a three-dimensional coupled chemistry-climate model, *Atmos. Chem. Phys.*, *4*, 2227–2239.

Struthers, H., et al. (2009), The simulation of the Antarctic ozone hole by chemistry-climate models, *Atmos. Chem. Phys.*, *9*, 6363–6376.

Teyssèdre, H., et al. (2007), A new tropospheric and stratospheric chemistry and transport model MOCAGE-Climat 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 O_3-climate predictions, *Q. J. Meteorol. Soc.*, *131*, 281–303.

Tilmes, S., D. E. Kinnison, R. R. Garcia, R. Müller, F. Sassi, D. R. Marsh, and B. A. Boville (2007), Evaluation of heterogeneous processes in the polar lower stratosphere in the Whole Atmosphere Community Climate Model, *J. Geophys. Res.*, *112*, D24301, doi:10.1029/2006JD008334.

Tilmes, S., et al. (2008), Chemical ozone loss in the Arctic winter 1991–1992, *Atmos. Chem. Phys.*, *8*, 1897–1910.

Waugh, D. W., L. Oman, S. R. Kawa, R. S. Stolarski, S. Pawson, A. R. Douglass, P. A. Newman, and J. E. Nielsen (2009), Impact of climate change on stratospheric ozone recovery, *Geophys. Res. Lett.*, *36*, L03805, doi:10.1029/2008GL036223.

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

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

J. Austin, Geophysical Fluid Dynamics Laboratory, NOAA, Princeton, NJ 08542-0308, USA. (john.austin@noaa.gov)

A. J. G. Baumgaertner, C. Brühl, and P. Jöckel, Max-Planck-Institut für Chemie, D-55020 Mainz, Germany.

S. Bekki, D. Cugnet, and M. Marchand, LATMOS, IPSL, UVSQ, UPMC, CNRS, INSU, F-75231 Paris, France.

G. E. Bodeker, Bodeker Scientific, 42 Young Ln., RD1, Alexandra, New Zealand.

P. Braesicke and J. Pyle, NCAS-Climate-Chemistry, Centre for Atmospheric Science, Department of Chemistry, Cambridge University, Cambridge CB2 IEW, UK.

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

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

M. Dameris and H. Garny, Deutsches Zentrum für Luft- und Raumfahrt, Institut für Physik der Atmosphäre, Wessling D-82234, Germany.

S. Frith and J. E. Nielsen, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA.

A. Gettelman, D. Kinnison, and J. F. Lamarque, NCAR, Boulder, CO 80305, USA.

A. Kubin and U. Langematz, Institut für Meteorologie, Freie Universität, D-12165 Berlin, Germany.

E. Mancini and G. Pitari, Dipartimento di Fisica, University of L'Aquila, I-167100 L'Aquila, Italy.

M. Michou and H. Teyssèdre, GAME, CNRM, Meteo-France, F-31057 Toulouse, France.

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

D. A. Plummer, Science and Technology Branch, Environment Canada, Toronto, ON M3H 5T4, Canada.

E. Rozanov, Physical-Meteorological Observatory, World Radiation Center, CH-7260 Davos, Switzerland.

J. Scinocca, CCCMA, University of Victoria, Victoria, BC V8W 3V6, Canada.

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

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

H. Struthers, Department of Applied Environmental Science, University of Stockholm, SE-10691 Stockholm, Sweden.