JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 111, D22308, doi:10.1029/2006JD007327, 2006

Assessment of temperature, trace species, and ozone in chemistry-climate model simulations of the recent past

V. Eyring,[1] N. Butchart,[2] D. W. Waugh,[3] H. Akiyoshi,[4] J. Austin,[5] S. Bekki,[6] G. E. Bodeker,[7] B. A. Boville,[8] C. Brühl,[9] M. P. Chipperfield,[10] E. Cordero,[11] M. Dameris,[1] M. Deushi,[12] V. E. Fioletov,[13] S. M. Frith,[14] R. R. Garcia,[8] A. Gettelman,[8] M. A. Giorgetta,[15] V. Grewe,[1] L. Jourdain,[6] D. E. Kinnison,[8] E. Mancini,[16] E. Manzini,[17] M. Marchand,[6] D. R. Marsh,[8] T. Nagashima,[4] P. A. Newman,[18] J. E. Nielsen,[14] S. Pawson,[18] G. Pitari,[16] D. A. Plummer,[13] E. Rozanov,[19] M. Schraner,[20] T. G. Shepherd,[21] K. Shibata,[12] R. S. Stolarski,[18] H. Struthers,[7] W. Tian,[10] and M. Yoshiki[4]

Received 21 March 2006; revised 28 July 2006; accepted 10 August 2006; published 23 November 2006.

[1] Simulations of the stratosphere from thirteen coupled chemistry-climate models (CCMs) are evaluated to provide guidance for the interpretation of ozone predictions made by the same CCMs. The focus of the evaluation is on how well the fields and processes that are important for determining the ozone distribution are represented in the simulations of the recent past. The core period of the evaluation is from 1980 to 1999 but long-term trends are compared for an extended period (1960–2004). Comparisons of polar high-latitude temperatures show that most CCMs have only small biases in the Northern Hemisphere in winter and spring, but still have cold biases in the Southern Hemisphere spring below 10 hPa. Most CCMs display the correct stratospheric response of polar temperatures to wave forcing in the Northern, but not in the Southern Hemisphere. Global long-term stratospheric temperature trends are in reasonable agreement with satellite and radiosonde observations. Comparisons of simulations of methane, mean age of air, and propagation of the annual cycle in water vapor show a wide spread in the results, indicating differences in transport. However, for around half the models there is reasonable agreement with observations. In these models the mean age of air and the water vapor tape recorder signal are generally better than reported in previous model intercomparisons. Comparisons of the water vapor and inorganic chlorine (Cl_y) fields also show a large intermodel spread. Differences in tropical water vapor mixing ratios in the lower stratosphere are primarily related to biases in the simulated tropical tropopause temperatures and not transport. The spread in Cl_y, which is largest in the polar lower stratosphere, appears to be primarily related to transport differences. In general the amplitude and phase of the annual cycle in total ozone is well simulated apart from the southern high latitudes. Most CCMs show reasonable agreement with observed total

[1]Institut für Physik der Atmosphäre, Deutsches Zentrum für Luft- und Raumfahrt, Oberpfaffenhofen, Wessling, Germany.
[2]Climate Research Division, Met Office, Exeter, UK.
[3]Department of Earth and Planetary Sciences, Johns Hopkins University, Baltimore, Maryland, USA.
[4]National Institute for Environmental Studies, Tsukuba, Japan.
[5]Geophysical Fluid Dynamics Laboratory, NOAA, Princeton, New Jersey, USA.
[6]Service d'Aéronomie du Centre National de la Recherche Scientifique, Paris, France.
[7]National Institute of Water and Atmospheric Research, Lauder, New Zealand.
[8]National Center for Atmospheric Research, Boulder, Colorado, USA.
[9]Max Planck Institut für Chemie, Mainz, Germany.
[10]Institute for Atmospheric Science, University of Leeds, Leeds, UK.
[11]Department of Meteorology, San Jose State University, San Jose, California, USA.
[12]Meteorological Research Institute, Tsukuba, Japan.
[13]Environment Canada, Toronto, Ontario, Canada.
[14]Science Systems and Applications, Inc., Lanham, Maryland, USA.
[15]Max Planck Institut für Meteorologie, Hamburg, Germany.
[16]Dipartimento di Fisica, Università L'Aquila, L'Aquila, Italy.
[17]Istituto Nazionale di Geofisica e Vulcanologia, Bologna, Italy.
[18]NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.
[19]Physical-Meteorological Observatory/World Radiation Center, Davos, Switzerland.
[20]Institute for Atmospheric and Climate Science, Eidgenössische Technische Hochschule, Zurich, Switzerland.
[21]Department of Physics, University of Toronto, Toronto, Ontario, Canada.

0148-0227/06/2006JD007327

ozone trends and variability on a global scale, but a greater spread in the ozone trends in polar regions in spring, especially in the Arctic. In conclusion, despite the wide range of skills in representing different processes assessed here, there is sufficient agreement between the majority of the CCMs and the observations that some confidence can be placed in their predictions.

Citation: 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.

1. Introduction

[2] The future evolution of the stratosphere is of particular interest for the timing of ozone recovery and chemistry-climate interactions. An important tool for understanding past changes and predicting the future evolution of the stratosphere is the coupled chemistry-climate model (CCM). CCMs include full representations of dynamical, radiative, and chemical processes in the atmosphere and their interactions. In particular, CCMs include feedbacks of the chemical tendencies on the dynamics, and hence the transport of chemicals. This feedback is a major difference between CCMs and chemical transport models, and is required to simulate the evolution of ozone in a changing climate. Ozone is a major radiative agent in the stratosphere and is also strongly affected by dynamics and transport, so the ozone radiative-dynamical feedback in CCMs is of particular importance for the representation of chemistry-climate coupling [*World Meteorological Organization/United Nations Environment Programme (WMO/UNEP)*, 2003].

[3] CCMs have been used to simulate the evolution of the stratosphere in the 21st century, and to predict the recovery of the ozone layer [*Austin et al.*, 2003; *Austin and Wilson*, 2006; *Dameris et al.*, 2006]. To interpret and assess these predictions the capabilities and limitations of the models first need to be determined by comparing their simulated meteorology and trace gas distributions with meteorological analyses and observations. Such comparisons not only point to model deficiencies and uncertainties, but can also help in understanding processes, mechanisms and feedbacks within the atmosphere and identifying deficiencies in our understanding. It is also of interest to compare models with each other, to determine the consistency between the models and to relate intermodel spread to model formulation and experimental setup.

[4] An intercomparison and assessment of CCMs was performed by *Austin et al.* [2003]. Since then a number of new CCMs with a focus on the middle atmosphere have been developed, and changes have also been made to the CCMs considered by *Austin et al.* [2003]. Some of the new and updated CCMs have been tested and compared to observations [*Austin et al.*, 2006; *Dameris et al.*, 2005; *Egorova et al.*, 2005; *Steinbrecht et al.*, 2006a, 2006b; *Struthers et al.*, 2004; *Tian and Chipperfield*, 2005; R. R. Garcia et al., Simulations of secular trends in the middle atmosphere, 1950–2003, submitted to *Journal of Geophysical Research,* 2006, hereinafter referred to as Garcia et al., submitted manuscript, 2006], but there is need for a new multimodel assessment with a common focus and experimental setup.

[5] In this study we evaluate simulations from thirteen CCMs. The CCMs examined have performed simulations of the past evolution of the stratosphere and have been used to predict the future evolution of stratospheric ozone in the 21st century in a follow up study and in support of the 2006 WMO/UNEP Scientific Assessment of Ozone Depletion. The primary objective of this paper is to evaluate the capabilities of the CCMs to simulate processes and fields that play an important role in determining the ozone distribution. Such an evaluation will provide guidance for the interpretation of predictions of future ozone evolution made by the CCMs.

[6] To evaluate the CCM results we compare the simulated distributions of temperature, ozone, and other trace gases with those derived from meteorological analyses and trace gas observations. We focus on quantities that are important for the simulation of ozone distribution and can be validated against observations. This includes temperature, water vapor, hydrogen chloride (a principal reservoir of inorganic chlorine), and observationally based transport diagnostics. Both, the climatological mean distributions as well as decadal-scale variations and trends during the 1960 to 2004 period are examined.

[7] This study extends both the *Austin et al.* [2003] assessment of CCMs and the *Pawson et al.* [2000] assessment of middle atmospheric general circulation models (GCMs). First, a larger number of CCMs is considered. Second, in contrast to these previous studies, the CCM simulations defined as part of the Chemistry-Climate Model Validation Activity for SPARC (CCMVal) [*Eyring et al.*, 2005a] used here are all transient simulations and have almost identical forcings (e.g., sea surface temperatures (SSTs), greenhouse gases (GHGs), and halocarbons). This eliminates many of the uncertainties in the conclusions of the earlier assessments that resulted from the differences in experimental setup of individual models. Finally, and perhaps most importantly, this study is the first multi-CCM assessment to evaluate transport and distributions of important trace gases. This includes an evaluation of the simulations of inorganic chlorine, which is of particular importance for simulations of the evolution of ozone and ozone recovery.

[8] The models and simulations as well as the observational data sets that are used in this study to evaluate the CCMs are described in section 2. In section 3 the simulated stratospheric temperatures in the CCMs and associated wave forcing is evaluated and compared with similar diagnostics calculated from meteorological analyses. Section 4 focuses on the evaluation of transport characteristics and long-lived tracers, by comparing methane, water vapor and mean age of air distributions to observations. Inaccuracies in dynamics and transport affect modeled inorganic

Table 1. CCMs Used in This Study[a]

Model	Investigators	Underlying GCM	Domain/Resolution	Radiative Feedbacks	Tracer Advection Scheme	O-GWD	NonO-GWD	Reference
AMTRAC	J. Austin, R. J. Wilson	AM2 [*Anderson et al.*, 2004]	2° × 2.5°, 48 L, 0.0017 hPa	O_3, H_2O	finite-volume [*Lin*, 2004]	*Stern and Pierrehumbert* [1988]	*Alexander and Dunkerton* [1999]	*Austin et al.* [2006]; *Austin and Wilson* [2006]
CCSRNIES	H. Akiyoshi, T. Nagashima, M. Yoshiki	CCSRNIES [*Numaguti*, 1993]	2.8° × 2.8° (T42), 34 L, 0.01 hPa	O_3, H_2O, CH_4, N_2O, CFCs	spectral in the horizontal, finite difference for the vertical	*McFarlane* [1987]	*Hines* [1997]	*Akiyoshi et al.* [2004], *Kurokawa et al.* [2005]
CMAM	J. McConnell, N. McFarlane, D. Plummer, J. Scinocca, T. Shepherd	CCCma AGCM3 [*Scinocca and McFarlane*, 2004]	3.75° × 3.75° (T32), 71 L, 0.0006 hPa	O_3, H_2O	spectral in the horizontal, finite elements in the vertical	*Scinocca and McFarlane* [2000]	*Scinocca* [2003]	*Beagley et al.* [1997], *de Grandpré et al.* [2000]
E39C	M. Dameris, V. Eyring, V. Grewe, M. Ponater	ECHAM4 [*Roeckner et al.*, 1996]	3.75° × 3.75° (T30), 39 L, 10 hPa	O_3, CH_4, N_2O, H_2O, CFCs	semi-Lagrange [*Williamson and Rasch*, 1994]	*Miller et al.* [1989]	none	*Dameris et al.* [2005, 2006]
GEOSCCM	A. Douglass, P. Newman, S. Pawson, R. Stolarski	GEOS-4 [*Bloom et al.*, 2005]	2° × 2.5°, 55 L, 0.01 hP	O_3, H_2O, CH_4, N_2O, CFC-11, CFC-12	finite-volume [*Lin*, 2004]	*Kiehl et al.* [1998]	adapted from *Garcia and Boville* [1994]	*Bloom et al.* [2005], *Stolarski et al.* [2006]
LMDZrepro	S. Bekki, F. Lott, F. Lefevre, M. Marchand	LMDz4 [*Lott et al.*, 2005]	2.5° × 3.75°, 50 L, 0.07 hPa	O_3, CH_4, N_2O, H_2O, CFC-11, CFC-12	finite-volume [*Hourdin and Armengaud*, 1999]	*Lott and Miller* [1997]	*Hines* [1997]	chemistry part: *Lefevre et al.* [1994]
MAECHAM4CHEM	C. Brühl, M. Giorgetta, E. Manzini, B. Steil	MAECHAM4 [*Manzini et al.*, 1997]	3.75° × 3.75° (T30), 39 L, 0.01 hPa	O_3, H_2O, CH_4	flux-form semi-Lagrange SPITFIRE [*Steil et al.*, 2003]	*McFarlane* [1987]	*Hines* [1997]	*Manzini et al.* [2003], *Steil et al.* [2003]
MRI	K. Shibata, M. Deushi	MRI/JMA98 [*Shibata et al.*, 1999]	2.8° × 2.8° (T42), 68 L, 0.01 hPa	O_3, CH_4, N_2O	hybrid semi-Lagrange [*Shibata et al.*, 2005]	*Iwasaki et al.* [1989], only short-wavelength GW	*Hines* [1997]	*Shibata and Deushi* [2005], *Shibata et al.* [2005]
SOCOL	E. Rozanov, M. Schraner	MAECHAM4 [*Manzini et al.*, 1997]	3.75° × 3.75° (T30), 39 L, 0.1 hPa	O_3, CH_4, N_2O, H_2O, CFCs	hybrid advection scheme [*Zubov et al.*, 1999]	*McFarlane* [1987]	*Hines* [1997]	*Egorova et al.* [2005], *Rozanov et al.* [2005]
ULAQ	E. Mancini, G. Pitari	ULAQ-GCM [*Pitari et al.*, 1992]	10° × 22.5°, 26 L, 0.04 hPa	O_3, H_2O, CH_4, N_2O, CFCs, HCFCs, aerosols.	flux-form Eulerian fully explicit scheme [e.g., *Pitari et al.*, 2002]		Rayleigh friction [*Smith and Lyjak*, 1985]	*Pitari et al.* [2002]
UMETRAC	N. Butchart, H. Struthers	UM [*Pope et al.*, 2000]	2.5° × 3.75°, 64 L, 0.01 hPa	O_3, H_2O	quintic-mono [*Gregory and West*, 2002]	*Gregory et al.* [1998]	*Warner and McIntyre* [1996]	*Austin* [2002], *Austin and Butchart* [2003], *Struthers et al.* [2004]
UMSLIMCAT	M. Chipperfield, W. Tian	UM [*Pope et al.*, 2000]	2.5° × 3.75°, 64 L, 0.01 hPa	O_3, CH_4, N_2O, H_2O	quintic-mono [*Gregory and West*, 2002]	*Gregory et al.* [1998]	*Warner and McIntyre* [1996]	*Tian and Chipperfield* [2005]
WACCM (v.3)	B. Boville, R. Garcia, A. Gettelman, D. Kinnison, D. Marsh, F. Sassi	CAM [*Collins et al.*, 2004]	4° × 5°, 66 L, 4.5 × 10^{-6} hPa	O_3, H_2O, N_2O, CH_4, CFC-11, CFC-12, NO, CO_2, O_2	finite-volume [*Lin*, 2004]	*McFarlane* [1987]	based on *Lindzen* [1981], *Holton* [1982], *Garcia and Solomon* [1985]	Garcia et al. (submitted manuscript, 2006)

[a]The models are listed alphabetically by name. The horizontal resolution is given in degrees latitude × longitude (grid point models). For spectral models the resolution is also given as T30, T32, and T42 corresponding to the triangular truncation of the spectral domain with 30, 32, and 42 wave numbers, respectively. All CCMs have a comprehensive range of chemical reactions. The time varying concentrations of CO_2 as well as other GHGs are considered in the radiation schemes of all models.

chlorine as well as ozone distributions, which are discussed in sections 5 and 6. In section 7 a summary is presented.

2. Models and Observational Data Sets

2.1. Model Descriptions

[9] Thirteen coupled chemistry-climate models have participated in this model intercomparison. Their main features are summarized in Table 1 and detailed descriptions of the CCMs are given in the cited literature.

[10] Some of the CCMs can be grouped together as they have been built on the same underlying GCM, although important differences exist among the models in each group (see Table 1). For example, UMETRAC and UMSLIMCAT are both based on the U.K. Unified Model (UM) [*Pope et al.*, 2000]. MAECHAM4CHEM and SOCOL are based on the same middle atmosphere GCM [*Manzini et al.*, 1997], but each model uses a different forcing for the *Hines* [1997] nonorographic gravity wave drag (NonO-GWD) scheme [see *Egorova et al.*, 2005]. E39C can also be grouped with MAECHAM4CHEM and SOCOL, because their underlying GCMs are all derived from ECHAM4 [*Roeckner et al.*, 1996]. In addition, the chemistry schemes of E39C and MAECHAM4CHEM are both based on *Steil et al.* [1998], while those of UMETRAC and AMTRAC are based on the chemistry scheme described by *Austin et al.* [2006]. Other groups of models are not based on the same GCM but do share some similar characteristics in their dynamics and transport. For example, AMTRAC, GEOSCCM, and WACCM all use finite-volume dynamical and advection schemes and GEOSCCM and WACCM have similar physical parameterizations. However, in all cases there are either completely separate components in the related CCMs (e.g., the base GCM is the same, but the chemistry differs) or significant modifications have been made to the originally common components.

[11] The horizontal resolution varies from $10^\circ \times 20^\circ$ (ULAQ) to $2^\circ \times 2.5^\circ$ (AMTRAC, GEOSCCM), and the location of the upper boundary from being centered at 10 hPa (E39C) to 4.5×10^{-6} hPa (WACCM). The number of vertical levels ranges from 26 (ULAQ) to 71 (CMAM). The numerical methods used for the dynamical core and advection schemes also vary (see Table 1). The dynamics in all CCMs is determined by solving the "primitive" equations, except for ULAQ which is a quasi-geostrophic model [*Pitari et al.*, 2002].

[12] A major issue with GCMs of the middle atmosphere is the treatment of gravity waves. Given the limited spatial resolution of these models, gravity waves need to be parameterized, and these parameterizations vary significantly among the models. Orographic gravity wave drag schemes are employed in all models except ULAQ. To represent the effects of gravity waves of nonorographic origin, parameterizations of gravity wave spectra are used in all models (see Table 1), except in E39C (the upper boundary is located below the region where these effects are important) and in ULAQ (Rayleigh friction). Currently a limitation of the gravity wave schemes in the models used in this study is that their source spectrum is specified externally and does not evolve in time in response to a changing climate.

[13] All CCMs have a comprehensive range of chemical reactions and include a number of chemical species from the O_x, HO_x, ClO_x, BrO_x, and NO_x families (apart from E39C and MAECHAM4CHEM, which do not include bromine species) and source gases. All models include both gas-phase chemistry and heterogeneous chemistry on aerosols, and on polar stratospheric clouds (PSCs), but different PSC and denitrification schemes are used. The formation of NAT particles is not included in CMAM and dehydration/denitrification by gravitational settling of PSC particles is not included in CMAM and UMSLIMCAT.

[14] The 13 CCMs differ in how they couple radiation and composition. All models use the simulated O_3 and H_2O fields in their radiation calculations, but the number of other simulated greenhouse gases used varies (see Table 1).

[15] In contrast to the other models, the UMSLIMCAT model carries two separate water vapor fields, a tropospheric humidity field that is used for tropospheric physics but is not coupled with the chemistry, and a stratospheric H_2O tracer used in the stratospheric chemistry module. This second water vapor field has a constant value of 3.3 ppmv in the troposphere and a chemical source in the stratosphere from methane oxidation. In all other models there is a single water vapor field that is used in the tropospheric physics and stratospheric chemistry.

[16] A subset of CCMs (CCSRNIES, CMAM, E39C, MAECHAM4CHEM, ULAQ, and UMETRAC) have already contributed to an earlier assessment of CCMs [*Austin et al.*, 2003]. However, all of these models have been improved since the last assessment. The horizontal resolution in CCSRNIES has been increased from T21 to T42, and a nonorographic gravity wave drag parameterization has been included. The chemistry scheme now includes bromine chemistry and the methane oxidation process [*Akiyoshi et al.*, 2004], and heterogeneous chemistry has changed [*Sessler et al.*, 1996]. The plane-parallel radiative transfer scheme has been replaced with a pseudospherical approximation [*Kurokawa et al.*, 2005]. The number of vertical levels in CMAM has been increased from 65 to 71 to improve resolution through the tropopause region, and the parameterization of both orographic and nonorographic gravity wave drag has also been changed (see Table 1). The surface drag and vertical diffusion coefficients have been decreased in the ULAQ model and cirrus ice particles have been included in the upper troposphere. The underlying GCM has not changed in E39C, MAECHAM4CHEM and UMETRAC since *Austin et al.* [2003] but changes have been made to the chemistry schemes: Two more heterogeneous reactions on stratospheric aerosol have been added to the chemistry module in E39C and MAECHAM4CHEM and photolysis for twilight conditions with a solar zenith angle up to 93° is included [*Lamago et al.*, 2003; *Steil et al.*, 2003]. UMETRAC uses the same chemistry scheme as AMTRAC, including among other changes an explicit treatment of long-lived tracers [*Austin et al.*, 2006; *Austin and Wilson*, 2006] and coupling of the underlying climate model water vapor to the stratospheric chemistry.

2.2. Model Simulations

[17] The model simulations considered here are transient simulations of the last decades of the 20th century. The specifications of the simulations follow or are similar to the "reference simulation 1" (REF1) of CCMVal [*Eyring et al.*, 2005b] and include anthropogenic and natural forcings

Table 2. External Forcings and Representation of QBO Used in the Model Simulations[a]

		External Forcings				
				Volcanoes		
Model	Length of Simulation	SSTs	Solar Variability	Chemical Effects	Direct Radiative Effects	QBO
AMTRAC	1960–2004	J. Hurrell (personal communication, 2005)	intensity of the 10.7 cm radiation of the sun from *Lean* [2000]	SADs from G. Stenchikov (personal communication, 2005)	calculated on line with climate model radiation scheme	no
CCSRNIES	1980–2004	HadISST1 [*Rayner et al.*, 2003]	intensity of the 10.7 cm radiation of the sun from *Lean et al.* [1997]	SADs from D. Considine (personal communication, 2005)	calculated on line using aerosol data from *Sato et al.* [1993]	forced
CMAM	1960–2004	HadISST1 [*Rayner et al.*, 2003]	no	SADs from D. Considine (personal communication, 2005)	calculated on line with climate model radiation scheme	no
E39C	1960–1999	HadISST1 [*Rayner et al.*, 2003]	intensity of the 10.7 cm radiation of the sun from *Lean et al.* [1997]	SADs from *Jackman et al.* [1996]	precalculated net heating rates [*Kirchner et al.*, 1999]	forced
GEOSCCM	1960–2003	HadISST1 [*Rayner et al.*, 2003]	no	no	no	no
LMDZrepro	1979–1999	AMIP II [*Taylor et al.*, 2000]	no	Modeled SADs	no	no
MAECHAM4CHEM	1980–1999	HadISST1 [*Rayner et al.*, 2003]	intensity of the 10.7 cm radiation of the sun from *Lean et al.* [1997]	SADs from *Jackman et al.* [1996]	precalculated net heating rates [*Kirchner et al.*, 1999]	forced
MRI	1980–2004	HadISST1 [*Rayner et al.*, 2003]	intensity of the 10.7 cm radiation of the sun from *Lean et al.* [1997]	SADs from D. Considine (personal communication, 2005)	calculated on line from the optical depth and effective radius based on the updated data set of *Sato et al.* [1993]	internally generated
SOCOL	1980–2004	HadISST1 [*Rayner et al.*, 2003]	*Lean* [2000]	SADs from *Thomason and Peter* [2006]	calculated on line using aerosol data from *Thomason and Peter* [2006]	forced
ULAQ	1970–2004	HadISST1 [*Rayner et al.*, 2003]	no	SADs from D. Considine (personal communication, 2005)	calculated on line [*Pitari*, 1993]	no
UMETRAC	1980–1999	HadISST1 [*Rayner et al.*, 2003]	intensity of the 10.7 cm radiation of the sun from *Lean et al.* [1997]	SADs from D. Considine (personal communication, 2005)	precalculated heating rates (G. Stenchikov, personal communication, 2005)	internally generated
UMSLIMCAT	1980–1999	AMIP II [*Gates et al.*, 1999]	no	SADs from D. Considine (personal communication, 2005)	no	internally generated
WACCM (v.3)	1950–2003	J. Hurrell (personal communication, 2005)	daily values of 10.7 cm radio flux from http://www.sec.noaa.gov	SADs from D. Considine (personal communication, 2005)	no	no

[a]Changes in halocarbons are prescribed as monthly means based on Table 4B-2 of *WMO/UNEP* [2003]. GHGs are based on *IPCC* [2001] in all simulations. Both data sets are extended through 2004 on the basis of observations.

based on changes in SSTs, trace gases, solar variability, and aerosol effects (from major volcanic eruptions). Full details of the REF1 specifications are summarized in *Eyring et al.* [2005b] and only the key aspects are given here. All forcings for the simulations can be downloaded from the CCMVal Project website at http://www.pa.op.dlr.de/CCMVal.

[18] In REF1 simulations:

[19] 1. The SSTs are prescribed as monthly means following the global sea ice and sea surface temperature (HadISST1) data set provided by the UK Met Office Hadley Centre [*Rayner et al.*, 2003]. This data set is based on blended satellite and in situ observations.

[20] 2. The surface concentrations of GHGs are based on *Intergovernmental Panel on Climate Change (IPCC)* [2001].

[21] 3. The surface halogens are based on Table 4B-2 of *WMO/UNEP* [2003] and are extended through 2004 (S. Montzka, personal communication, 2005).

[22] 4. Chemical kinetics is taken from *NASA Panel for Data Evaluation* [2006].

[23] 5. The influence of the 11-year solar cycle on photolysis and heating rates is parameterized according to the intensity of the 10.7 cm radiation of the sun [*Lean et al.*, 1997].

[24] 6. Both chemical and direct radiative effects of enhanced stratospheric aerosol abundance from large volcanic eruptions are considered (see below).

[25] 7. A quasi-biennial oscillation (QBO) is either internally generated or forced (see below).

[26] The three major volcanic eruptions (Agung in 1963, El Chichón in 1982, and Pinatubo in 1991) are taken into account in REF1 simulations by prescribing observed sulfate aerosol surface area densities (SADs) and additional modeled heating rates (see Table 2). SADs are specified from a monthly climatology based on satellite data, similar to that used by *Jackman et al.* [1996] and updated by D. B. Considine (NASA Langley Research Center). The heating rates are monthly means from January 1950 to December 1999 for all-sky condition, and were calculated using volcanic aerosol parameters from *Sato et al.* [1993], *Hansen et al.* [2002] and GISS ModelE radiative routines and climatology [*Schmidt et al.*, 2006; G. Stenchikov and L. Oman, personal communication, 2005].

[27] The QBO is included in one of two ways in REF1 simulations. Three of the models (MRI, UMETRAC, and UMSLIMCAT) are able to internally generate a QBO driven by resolved and parameterized wave mean-flow interaction [*Scaife et al.*, 2002; *Shibata and Deushi*, 2005]. The QBO in these models is an internal mode and hence is not synchronized with the observed QBO, but has the correct mean period. The other CCMs do not generate an in situ QBO and, instead, simulate permanent tropical easterlies. In these models the QBO is forced externally by relaxation (''nudging'') of the stratospheric equatorial zonal wind to observed equatorial zonal wind profiles [*Giorgetta and Bengtsson*, 1999]. The externally driven QBO is therefore synchronized with the observed QBO.

[28] The forcings employed in each of the 13 transient CCM simulations used in this intercomparison are summarized in Table 2. Not all the model simulations examined here followed the suggested REF1 specifications exactly. The SSTs used are all based on observations, but in AMTRAC, LMDZrepro, UMSLIMCAT and WACCM other SSTs than the HadISST1 data set have been used (see Table 2). Furthermore, not all models included a QBO or changes in solar forcing and the amount of sulfate aerosols, and the method for accounting for the radiative effects of volcanic aerosols varied and is not included in all simulations.

2.3. Observational Data Sets

[29] For the intercomparisons shown in the following sections different sets of observations have been employed. Meteorological fields from several different analysis systems are used to assess the models. The differences between the climatologies derived from these fields are an indicator of the uncertainties in the meteorological analyses. As many of the CCM simulations started in January 1980 and ended in December 1999, we use the same time period for the ERA-40 reanalyses [*Uppala et al.*, 2005] and the NCEP Climate Prediction Center (CPC) stratospheric analyses [*Gelman et al.*, 1996]. The production of UK Met Office assimilated stratospheric analyses UKMO [*Swinbank and O'Neill*, 1994] only started in October 1991, and to produce a 10 year climatology representative for the 1990s we use the assimilated fields from January 1992 to December 2001. Modeled temperature time series are compared to ERA-40 reanalysis data and radiosonde data derived from the Radiosonde Atmospheric Temperature Products for Assessing Climate (RATPAC) climatology [*Free et al.*, 2005].

[30] To evaluate transport as well as hydrogen chloride and ozone distributions, we compare modeled tracer distributions with data of the Halogen Occultation Experiment (HALOE) on board the Upper Atmosphere Research Satellite (UARS). Methane, water vapor, hydrogen chloride and ozone retrieved from HALOE are available since 1991 [*Russell et al.*, 1993]. Model climatologies of the 1990s are compared to a climatology derived from 1991 to 2002 HALOE data, where data since September 2002 have not been included because of the unusual major warming in the Antarctic in 2002 and because the observations have been less frequent after 2002 [*Grooß and Russell*, 2005]. This climatology provides a consistent data set and is easily accessible. Uncertainties of single profile HALOE retrievals have been estimated in several studies [e.g., *Brühl et al.*, 1996; *Harries et al.*, 1996; *Park et al.*, 1996; *Russell et al.*, 1996]. For all measured species the accuracy of the HALOE retrievals decreases near the tropopause. In addition, sparse coverage of the polar regions increases the uncertainty in the HALOE climatologies there. In all intercomparisons the HALOE climatological mean and the interannual standard deviation (1σ) are shown.

[31] Not enough observations are available to form a global climatology for mean age of air and inorganic chlorine (Cl_y). However, measurements in single years exist that can be used for assessing the simulations. These data are described in sections 4.3 and 5.

[32] Climatological means of calculated total column ozone as well as variability and trends are compared against four observational data sets: ground-based, merged satellite data, the National Institute of Water and Atmospheric Research (NIWA) assimilated database and Solar Backscatter Ultraviolet (SBUV, SBUV/2) retrievals. The ground-based zonal mean data set used in this study includes long-

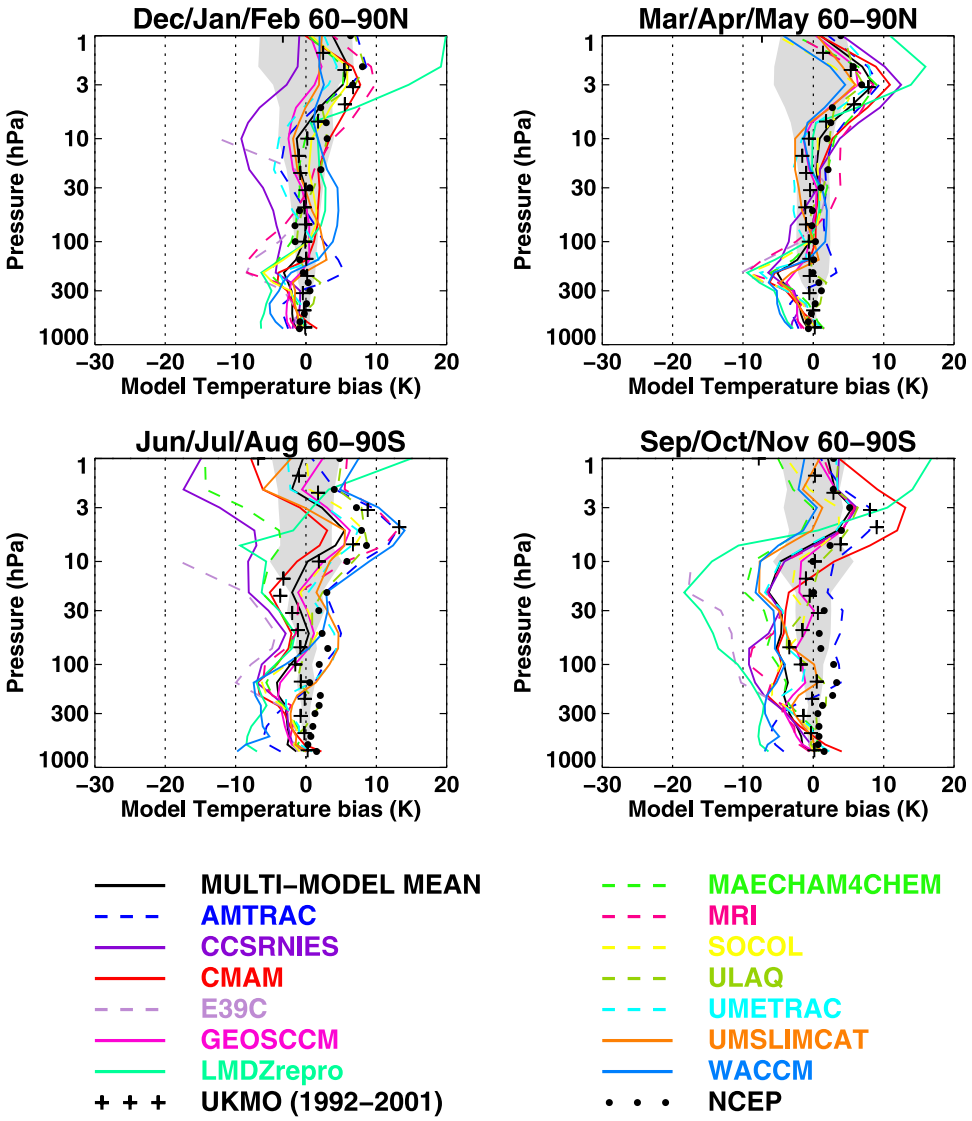

Figure 1. Climatological mean temperature biases for (top) 60–90°N and (bottom) 60–90°S for the (left) winter and (right) spring seasons. The climatological means for the CCMs and NCEP data from 1980 to 1999 and for UKMO from 1992 to 2001 are included. Biases are calculated relative to ERA-40 reanalyses. The grey area shows ERA-40 plus and minus 1 standard deviation (σ) about the climatological mean.

term measurements from Dobson, Brewer and filter ozonometer instruments and is updated from *Fioletov et al.* [2002]. The merged satellite data set consists of monthly mean zonal and gridded average data sets constructed from individual Total Ozone Mapping Spectrometer (TOMS) and SBUV/2 satellite data sets [*Stolarski and Frith*, 2006]. All data in the present version of the merged data set have been derived using the TOMS Version 8 and SBUV Version 8 algorithms. The NIWA data set is the same as that described by *Bodeker et al.* [2005] except that daily total column ozone fields from the Ozone Monitoring Instrument (OMI) are used, when they are available, in preference to the Earth Probe TOMS data. Differences between OMI overpass measurements and ground-based Dobson spectrophotometer measurements are small (−2.8 ± 5 DU) and are corrected before the data are combined with the other data sources. The SBUV data set is based on zonally averaged SBUV and SBUV/2 data (version 8) for the period from 1979 to 2004 (updated from *Miller et al.* [2002]).

3. Stratospheric Temperatures

[33] Similar to previous assessments [*Pawson et al.*, 2000; *Austin et al.*, 2003] we evaluate stratospheric temperatures in terms of the climatological means and the response to wave forcings from the troposphere. In addition we examine the CCMs' abilities to reproduce observed past temperature trends and variability.

3.1. Climatological Mean Temperatures

[34] Because of the temperature dependence of many of the chemical reactions determining the ozone distribution it is important to accurately model stratospheric temperatures. A measure of the accuracy of the underlying radiative heating in the models is given by the simulated global

annual mean temperatures. In the lower stratosphere where there is very good agreement among the three analyses, a large subset of models (9 out of 13) have relatively small biases (not shown). Below 10 hPa there are biases only in four models (CCSRNIES ±6 K; UMETRAC −2 K, AMTRAC +4 K between 50 and 10 hPa, and WACCM +3 K between 50 and 20 hPa). In the upper stratosphere, where radiosondes are generally not available, the only constraint on the analyses comes from deep-layer satellite radiances with poorly known biases and from the characteristics of the underlying assimilation model. Consequently the differences between the analyses are quite large (up to 5 K) and 11 of the 13 CCMs lie in the same range. Only CCSRNIES and CMAM lie outside the range of analyzed global mean temperatures with warm biases of 5–10 K. Note that CMAM unlike most models includes near-infrared CO_2 solar heating, which warms the upper stratosphere by 1–1.5 K [*Fomichev et al.*, 2004]. However, despite these issues, overall this is a significant improvement on the persistent stratospheric problem of negative global, annual mean temperature biases found in most of the middle atmosphere GCMs evaluated by *Pawson et al.* [2000, Figure 1]. This is due in part to improvements in the radiation schemes and, in the upper troposphere, in dynamics and/or vertical resolution, but could also be related to chemistry. In the upper stratosphere, using interactive rather than prescribed ozone as in most of the *Pawson et al.* [2000] models, has a self-correcting effect on temperature, because of the negative ozone-temperature feedback [*de Grandpré et al.*, 2000]: a warm bias acts to decrease ozone chemically, which leads to less radiative heating and thus reduces the warm bias.

[35] High-latitude temperatures in winter and spring are particularly important for correctly modeling PSC induced polar ozone depletion. The mean winter and springtime temperature biases poleward of 60° are shown in Figure 1. In the upper stratosphere there are large variations between the analyses (±9 K) and most models except LMDZrepro and CCSRNIES lie within this range. In the Northern Hemisphere below 10 hPa the meteorological analyses from ERA-40, NCEP, and UKMO agree very well. Between 100 and 10 hPa in spring all CCMs and in winter all CCMs except CCSRNIES, E39C (only at 10 hPa) and WACCM have rather small biases (1–3 K) and lie within the ERA-40 interannual standard deviations. Compared to the *Austin et al.* [2003] assessment the vertical profiles in particular in Northern Hemisphere spring have more similar characteristics and span a smaller range. However, it should be noted that in this study the results include 20 years of data and are averaged over 60° to 90° in contrast to *Austin et al.* [2003], who compared UKMO data from 1992 to 2001 to 10 years of transient simulations and time slice experiments at a single latitude (80°). For CMAM, the elimination of a significant warm bias in the Arctic compared to previously reported results is attributed to a reduction in the critical inverse Froude number used to define wave breaking in the orographic gravity wave drag parameterization (following independent work by S. Webster and B. Boville, private communication, 2004).

[36] In the Southern Hemisphere winter and spring upper stratosphere, again there are large differences between the meteorological analyses and all except five models

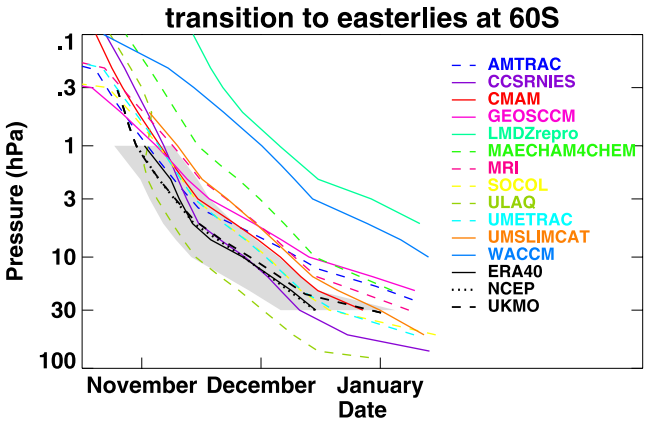

Figure 2. Descent of the zero zonal mean wind lines at 60°S based on the climatological mean annual cycle calculated from the monthly mean zonal mean winds. The grey area indicates the variation of the timing in the transition from westerlies to easterlies for ERA-40 due to a plus or minus one interannual standard deviation in the mean annual cycle. Tick marks refer to the first of the month, and climatological means are calculated as in Figure 1.

(CCRSNIES, E39C, MAECHAM4CHEM and LMDZrepro in winter; CMAM and LMDZrepro in spring) lie within this range. Compared to ERA-40 the mean bias is slightly positive, but compared to NCEP and UKMO it is slightly negative in most models. Overall it is difficult to assess the models' abilities to simulate temperatures in the upper stratosphere because of the large uncertainties in the analyses there. In the Southern Hemisphere winter lower stratosphere, the biases in individual models are in the range of ±5 K which is larger than the interannual standard deviation and differences between analyses, though the mean bias with respect to ERA-40 analyses is close to zero. In the Southern Hemisphere spring all the models apart from AMTRAC have definite cold biases which are most pronounced in LMDZrepro and E39C (around 15 K). This "cold-pole" problem is a long-standing problem of stratospheric GCMs and CCMs that can be attributed in part to missing wave drag [*Garcia and Boville*, 1994]. Again there is less spread in model biases in spring compared to the previous CCM assessment partly because the poorest performing models are no longer included or have improved. However, again, as noted above, a direct comparison is not possible.

[37] The "cold-pole" problem in the models is associated with a stronger vortex and a later than observed breakup. This affects the simulation of the Antarctic ozone hole and can be quantified by looking at zonal mean zonal winds. On the basis of the mean seasonal cycle in the monthly and zonally averaged zonal winds at 60°S the transition from westerlies to easterlies occurs too late in many models (Figure 2). In two models (LMDZrepro and WACCM) the easterlies do not descent below 10 hPa before the middle of January, and in one model (E39C) the transition from westerlies to easterlies does not occur at all as a consequence of the model's low upper boundary. The poor representation of this aspect of the seasonal cycle appears to be a persistent feature for many models with important implications for tracking the seasonal longevity of Antarctic

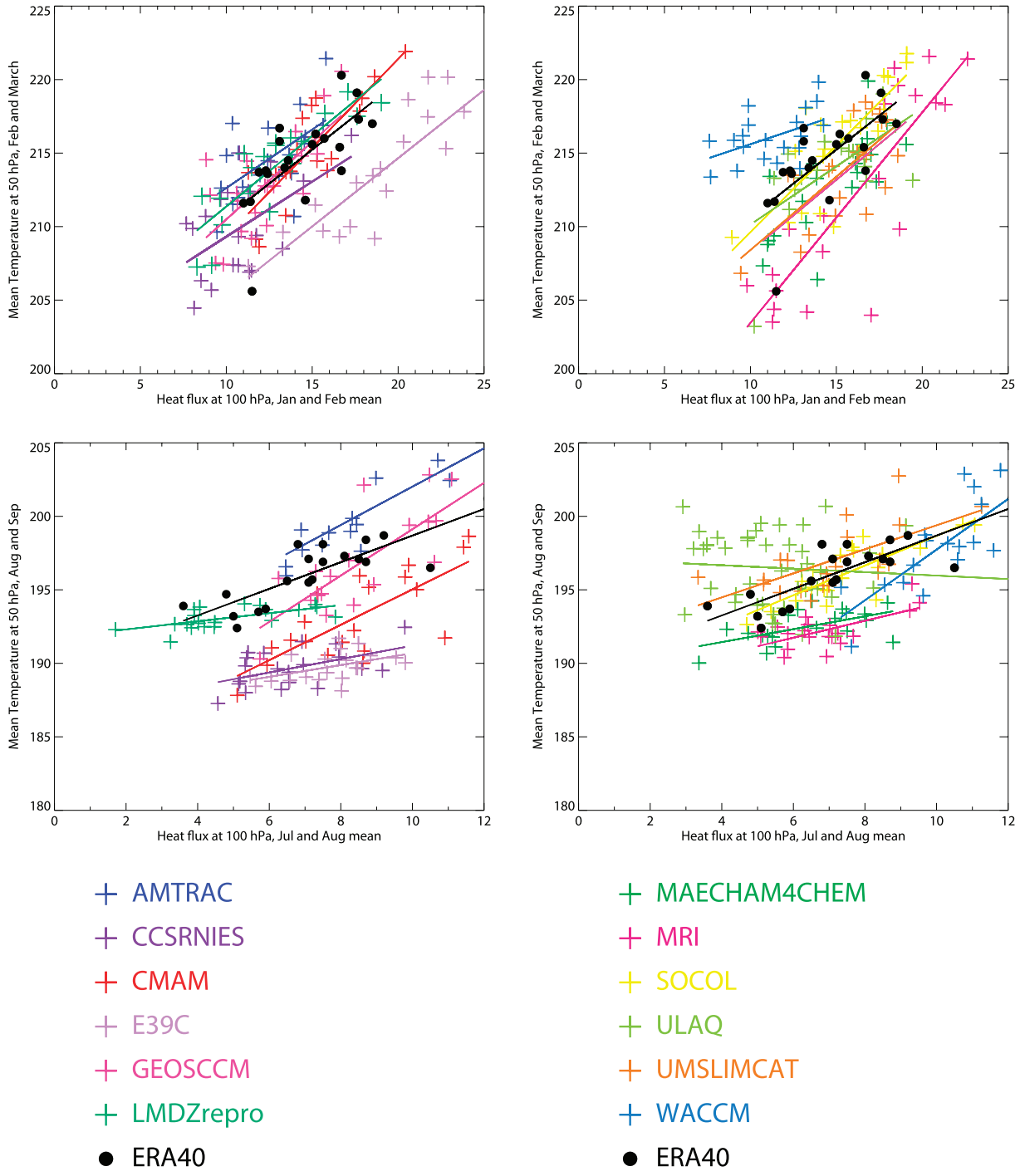

Figure 3. (top) Heat fluxes ($\overline{v'T'}$) at 100 hPa (averaged over 40°N to 80°N for January and February) versus temperatures at 50 hPa (averaged over 60°N to 90°N for February and March). Shown are the 20 years from 1980 to 1999 for each model simulation compared to observations from ERA-40 reanalyses. (bottom) Same for Southern Hemisphere, but heat fluxes ($\overline{v'T'}$) at 100 hPa averaged over 40°S to 80°S for July and August versus temperatures at 50 hPa averaged over 60°S to 90°S for August and September.

ozone depletion. In particular, the late breakdown of the polar vortex means that the Antarctic ozone hole persists longer than in reality in most models (see section 6), though the ozone depletion itself may also contribute to the late breakdown. A comparison of UKMO (1992–2001) with ERA-40 (1980–2000) and NCEP (1980–2000) results in Figure 2 suggests that when the ozone depletion was most severe in the 1990s the vortex persisted longer on average in the lower stratosphere than in the 1980s.

[38] The vertical propagation of planetary waves from the troposphere into the stratosphere plays a significant role in determining the temperatures of the polar stratosphere during winter and spring and their interannual variability. The wave forcing can be quantified in terms of the meridional

heat flux ($\overline{v'T'}$) at 100 hPa, which is proportional to the vertical component of the Eliassen-Palm (EP) flux entering the stratosphere. The observed interannual variations of 50 hPa polar cap average temperatures are well correlated with the meridional heat flux at 100 hPa [*Newman et al.*, 2001].

[39] Figure 3 shows scatterplots of polar cap temperatures at 50 hPa versus heat fluxes at 100 hPa for the northern (Figure 3, top) and southern (Figure 3, bottom) hemispheres. The correlations calculated from ERA-40 reanalyses give almost identical linear relationships of the ones calculated from the NCEP/NCAR reanalyses [*Kalnay et al.*, 1996] (not shown), with some scatter among individual years. The displacement of the model results above or below the meteorological analyses is a reflection of the temperature bias (see Figure 1). In the Northern Hemisphere the models generally reproduce the slope of the linear fit between temperature and heat flux, though in some models (especially ULAQ and WACCM) the correlation between the variables is lower than in the ERA-40 and NCEP/NCAR reanalyses. The slope of the linear fit is quite accurate in most of the models which suggests that they should be able to predict the correct temperature response to a change in the wave forcing. Compared to *Austin et al.* [2003], there are in general no significant improvements. However, the results from two of the models (CCSRNIES and ULAQ) are in better agreement with meteorological reanalyses than previously reported. Higher horizontal resolution (T42 rather than T21) in the CCSRNIES model has led to, on average, larger and more realistic heat fluxes though most of this increase can be attributed to the use of daily fields rather than the 10 day averages used in the previous assessment. In ULAQ the decrease of eddy friction in the troposphere could be the main reason for the improvement in the heat flux versus temperature diagnostic; previously, the slope in the Northern Hemisphere was much too shallow, indicating too weak a sensitivity to the wave forcing.

[40] In the Southern Hemisphere (Figure 3, bottom) there is larger variation among the models, and on average the models are in poorer agreement with the ERA-40 and NCEP/NCAR reanalyses than in the north. Relative to the meteorological reanalyses six of the models (CCSRNIES, E39C, MAECHAM4CHEM, MRI, LMDZrepro, and ULAQ) have temperatures that are rather unresponsive to changes in the heat fluxes (i.e., the slope of the linear fit is much flatter than observed), with the ULAQ model having the incorrect sign. These same six models also have the lowest correlations between heat fluxes and temperatures indicating problems in their dynamical response to changes in planetary wave forcing in the Southern Hemisphere in winter. In general, in most of the models with the exception of AMTRAC or LMDZrepro, the mean heat flux or wave activity entering the stratosphere appears to be broadly correct and therefore the deficiencies in the heat flux temperature relationship are probably due more to weaknesses in the representation of the stratosphere. As in the work by *Austin et al.* [2003] it is not possible to identify any direct link between the deficiencies in the Southern Hemisphere behavior and model resolution.

3.2. Past Variability and Trends in Temperature

[41] Figure 4 shows time series of zonally averaged temperature anomalies at 50 hPa between 1960 and 2005 from the CCM simulations, ERA-40 and radiosonde data. The RATPAC data are used for comparisons with the global anomalies, and the ERA-40 data are only included since 1980, when satellite observations became available. The temperature anomalies are calculated with respect to a mean reference period between 1980 and 1989 using 3-month averages for February to April in the polar Northern Hemisphere (60–90°N), September to November in the polar Southern Hemisphere (60–90°S) and annual averages for the global anomalies. A linear temperature trend in K/decade is calculated for each model using results between 1980 and 1999, and the statistical significance of that trend being nonzero is evaluated using a Student t-test at the 95% level.

[42] In the Northern Hemisphere (60–90°N) February to April means the year-to-year temperature anomalies range between ±5 K. Overall, 11 out of 13 models show a cooling trend between 1980 and 2000, although in only one of these models is the trend statistically significant (assuming each year is independent). The ERA-40 data indicate a statistically significant trend of −2.1 K/decade, a larger cooling trend than all but three of the models (GEOSCCM, LMDZrepro, and MRI). The relatively large interannual variability of the high-latitude winter hemisphere and small magnitude of the trend makes it difficult to determine accurate trends over only 20 years of data.

[43] Compared to the north, in the high-latitude Southern Hemisphere (60–90°S) September to November means (Figure 4, middle) the trend is stronger in ERA-40 data (−2.4 K/decade; not statistically significant) and also in the CCMs (between 0.97 to −4.1 K/decade). The trends in five out of 13 of the models are statistically significant. The stronger cooling trend in the Southern Hemisphere high latitudes is expected because of the larger Southern Hemisphere polar ozone loss. Over the length of the model simulations, the interannual variability in the Southern Hemisphere appears larger than in the Northern Hemisphere, although at higher altitudes this relationship does not hold (10 hPa) or reverses (1 hPa).

[44] The computed global trends (Figure 4, bottom) all show cooling between 1980 and 1999 and are in good agreement with *Shine et al.* [2003], who compare various model simulated trends (with imposed ozone changes) to observations. At 50 hPa, the CCM average model trends (using only those models with statistically significant trends) between 1980 and 1999 are −0.6 K/decade, while the similar average shown in Figure 5 of *Shine et al.* [2003] is −0.55 K/decade. Both sets of models underpredict the RATPAC observations and are closer to the satellite (MSU) observation which show a cooling of −0.8 K/decade. It is noted that the RATPAC climatology may have cold biases compared to the MSU observations in the tropical stratosphere [*Randel and Wu*, 2006]. The six models that extend back to 1960 all show a statistically significant 1960 to 1980 trend between −0.02 and −0.48 K/decade compared to RATPAC at −0.45 K/decade. These results are also in agreement with coupled atmosphere ocean models, where statistically significant global mean trends (−0.4 to −0.8 K/decade between 1980 and 2000) at 50 hPa are found in models that include changes in stratospheric ozone [*Cordero and Forster*, 2006].

[45] Superimposed on the overall observed cooling trend are large stratospheric warming signals due to the three

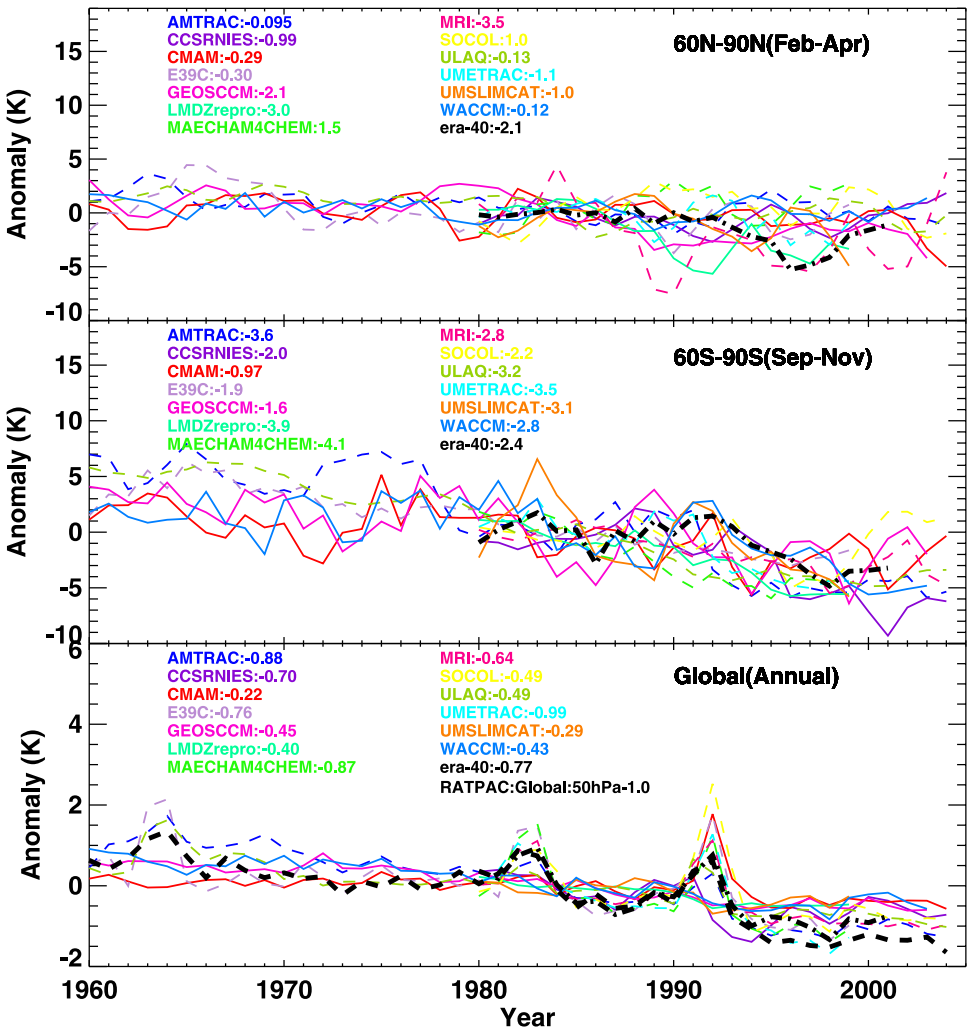

Figure 4. Modeled and observed time series of monthly mean temperature anomalies at 50 hPa from the CCMs, ERA-40 reanalyses, and RATPAC. The temperature anomalies are calculated with respect to a mean reference period between 1980 and 1989 using 3-month averages for (top) February to April in the polar Northern Hemisphere (60–90°N) and (middle) September to November in the polar Southern Hemisphere (60–90°S) and (bottom) annual averages for the global anomalies. For the polar plots a 3-year smoothing window has been applied. AMTRAC, E39C, MAECHAM4CHEM, MRI, SOCOL, ULAQ, and UMETRAC are shown with dashed lines, and all others CCMs are shown with solid lines. A linear temperature trend in K/decade is calculated for each model using data between 1980 and 1999. The temperature trend is given next to the name of each participating model.

major volcanic eruptions (Agung in 1963, El Chichón in 1982, and Pinatubo in 1991) which occur because volcanic aerosols absorb both incoming solar radiation and outgoing thermal radiation [*Ramaswamy et al.*, 2001]. From the 12 models that include volcanic aerosols a subset of nine models (see Table 2) considered both chemical and direct radiative effects of enhanced stratospheric aerosol abundance. The volcanic eruptions of El Chichón and Mt. Pinatubo are clearly visible in the temperature anomalies in these models, with a tendency to overestimate the magnitude of the temperature response compared to ERA-40 and RATPAC. The heating rates have not been uniformly specified in different simulations which is responsible for some of the intermodel differences in the warming signals. The observed "step-like" behavior of the long-term evolution of 50 hPa temperature, i.e., obviously lower temperature after the disappearance of the aerosol signal and more or less no trend in the following years, is only reproduced by those models which include the solar cycle effects as well as chemical and direct radiative effects from volcanic eruptions (see Table 2). Overall, the long-term behavior of the lower-stratospheric temperatures, a linear cooling trend caused by greenhouse gas increases and stratospheric ozone depletion, modulated by the solar cycle which causes higher temperatures in solar maximum than in solar minimum [*Santer et al.*, 2006], is well reproduced by the CCMs.

[46] At higher altitudes the CCMs indicate an average global temperature trend of −0.7 K/decade at 10 hPa and −1.6 K/decade at 1 hPa (not shown), compared to *Shine et al.* [2003] who calculated −0.6 K/decade at 10 hPa and −1.9 K/decade at 1 hPa. In the Southern Hemisphere polar region, the trend is considerably smaller at both 10 hPa

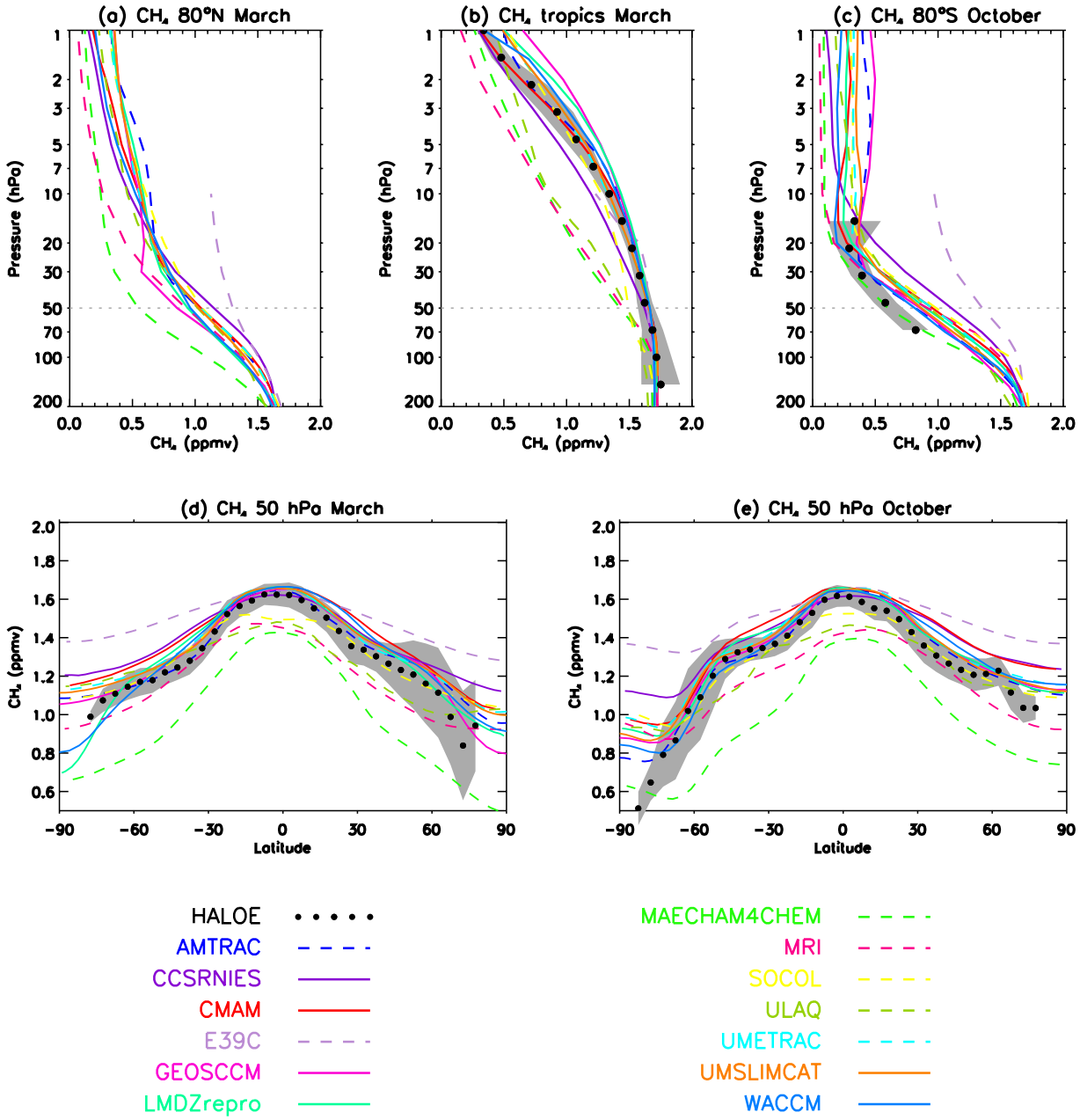

Figure 5. Climatological zonal mean CH_4 mixing ratios from the CCMs and HALOE in ppmv. Vertical profiles at (a) 80°N in March, (b) 0° in March, and (c) 80°S in October. Latitudinal profiles at 50 hPa in (d) March and (e) October. The grey area shows HALOE plus and minus 1 standard deviation (σ) about the climatological zonal mean.

(−0.6 K/decade) and 1 hPa (−0.7 K/decade) compared to 50 hPa (−2.7 K/decade). In the Northern Hemisphere polar region, there is a smaller difference between the trends at 10 hPa (−0.3 K/decade) and at 1 hPa (−1.2 K/decade) compared to 50 hPa (−0.8 K/decade).

4. Stratospheric Transport Characteristics

[47] Transport in the stratosphere involves both meridional overturning (the residual circulation) and mixing. Horizontal mixing is highly inhomogeneous, with transport barriers in the subtropics and at the edge of the wintertime polar vortices [e.g., *Sankey and Shepherd*, 2003]. Possible errors in transport characteristics have important impacts on the distribution of chemical families and individual species that affect ozone chemistry (NO_y, Cl_y, H_2O, and CH_4) and consequently the ozone distribution itself.

[48] Useful information on transport characteristics can be obtained from examining the distribution of long-lived tracers, such as methane and nitrous oxide. Also, information on the tropical ascent, vertical diffusion and tropical-extratropical mixing can be obtained from the vertical propagation of the annual cycle in water vapor (the so-

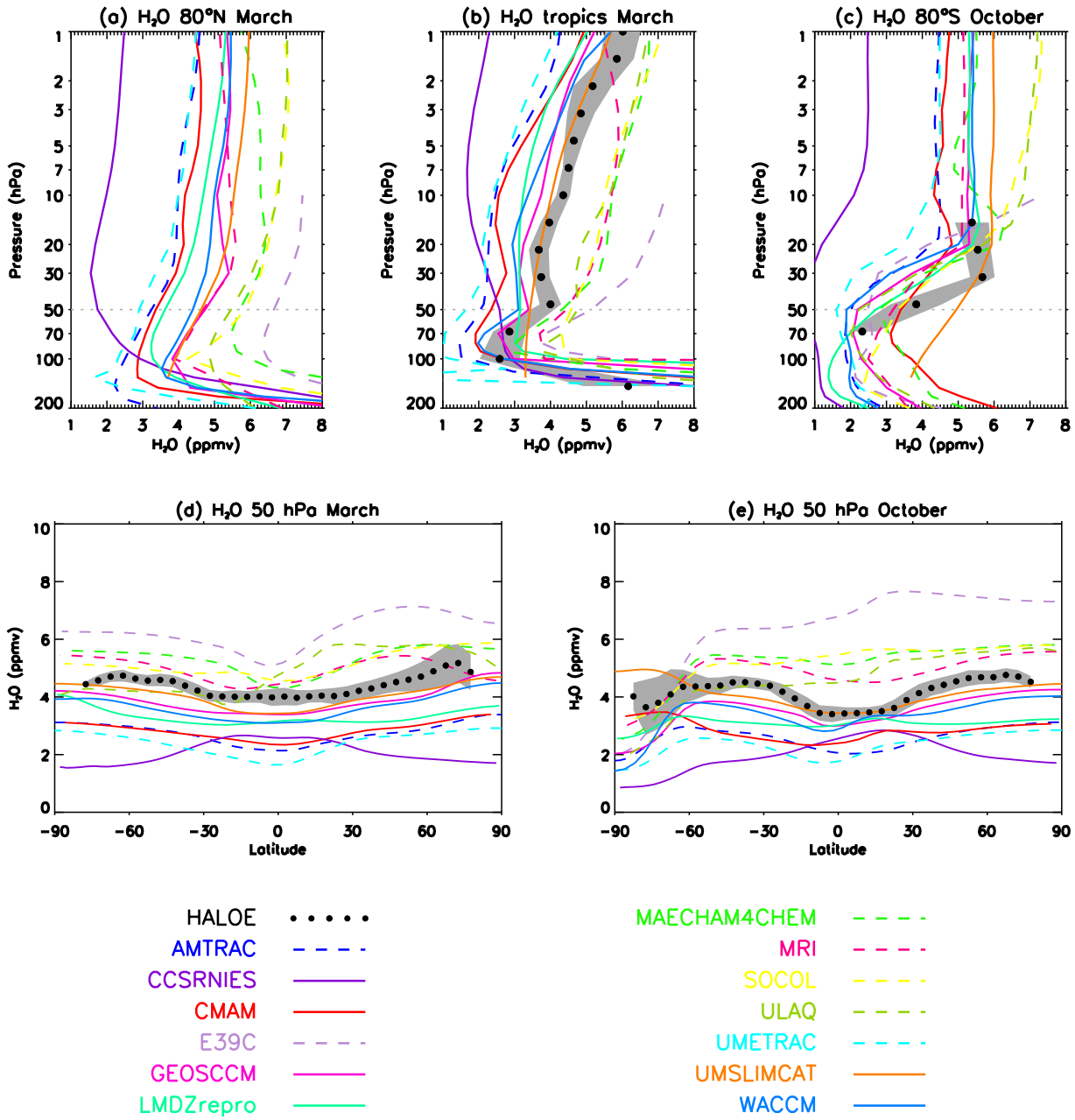

Figure 6. (a–e) As in Figure 5 but for H_2O in ppmv.

called "tape recorder"), while the mean age of air provides information on integrated transport within the stratosphere.

4.1. Methane

[49] As all CCMs simulate methane and there is a long observational record from HALOE we focus on this tracer. The concentration of methane entering the stratosphere is largely controlled by surface emissions and hydroxyl-initiated oxidation in the troposphere. Once methane enters the stratosphere the concentration is controlled by methane oxidation (primarily in the middle and upper stratosphere) and transport. In the simulations the surface concentration of CH_4 is prescribed with the same time varying values and it is expected that the oxidation in the stratosphere is very similar in all the models, so differences in the CH_4 distribution can be mainly attributed to differences in transport.

[50] Figure 5 shows climatological mean vertical profiles at different latitudes (Figures 5a–5c) and latitudinal variation (Figures 5d and 5e) at 50 hPa of monthly mean zonal mean CH_4 mixing ratios from the CCMs and HALOE. There are large variations in CH_4 mixing ratios among the different CCMs, with some large differences from observations. For example, in the lower stratosphere (50 hPa) tropical CH_4 mixing ratios from 4 CCMs (MAECHAM4-CHEM, MRI, SOCOL and ULAQ) are much smaller than those observed or calculated by other models, while CH_4 from E39C is much higher than observed and than in other models in the extratropical lower stratosphere. There is a particularly large spread in the simulated mean CH_4 mixing ratios in the polar lower stratosphere, with values varying from around 0.6 ppmv to around 1.4 ppmv. These large variations in CH_4 indicate large differences in transport.

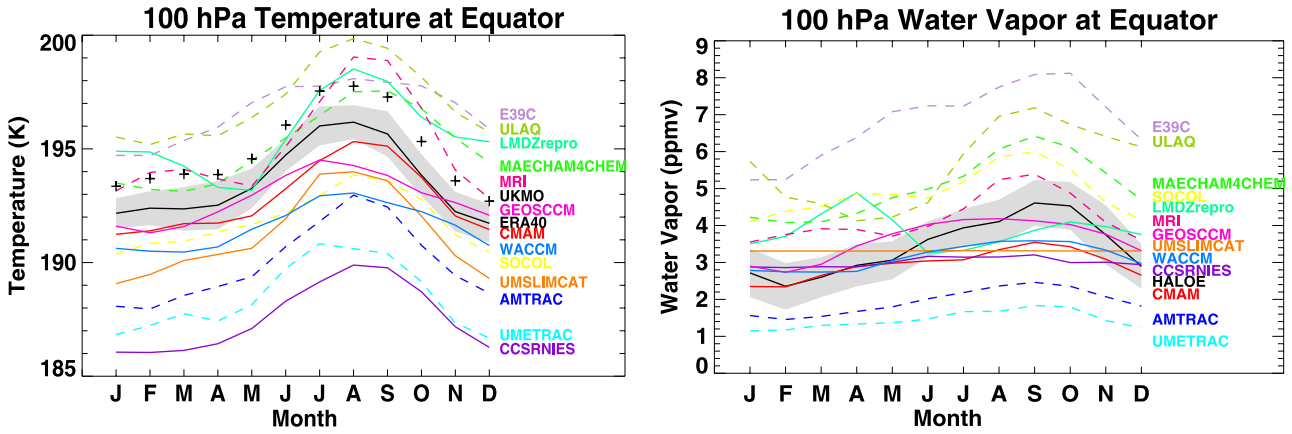

Figure 7. Seasonal variation of climatological means at 100 hPa at the equator for (left) temperature and (right) water vapor. Modeled fields for the 1990s are compared to the 1991–2002 HALOE water vapor climatology and the 1992–2001 temperature climatology from UKMO and ERA-40.

However, in most regions, the large range of model values is caused by the few CCMs mentioned above, and for a large subset of the models (8 out of the 13 models) the spread in CH_4 mixing ratios is much smaller and the simulated mean CH_4 mixing ratio is, in general, within 1σ of the HALOE mean. The one exception is in polar regions in spring where there is larger variation among the 8 CCMs, and all are higher than HALOE observations at 80°S. However, there is large interannual, including decadal, variability in polar transport (and tracers), and this could contribute to some of the model spread.

4.2. Water Vapor

[51] Future changes in stratospheric water vapor could have a major impact on stratospheric ozone, both through changes in the radiative balance, changes in HO_x, and through the formation of PSCs. It is therefore important to assess how well the CCMs simulate the observed water vapor distribution. Furthermore, the vertical propagation of the annual cycle in tropical water vapor (the water vapor "tape recorder" [*Mote et al.*, 1996]) can be used to assess the ascent rate and the mixing with midlatitudes in models. HALOE water vapor observations [*Mote et al.*, 1996; *Randel et al.*, 2004; *Grooß and Russell*, 2005] have been used in previous model-observation comparisons, including the NASA Models and Measurements II ("MMII") study [*Hall et al.*, 1999; *Sage et al.*, 1999] and NASA GMI model evaluation [*Douglass et al.*, 1999].

[52] There is an extremely large spread in the simulated water vapor fields (Figure 6) and, unlike the case for CH_4, this is not due to only a small subset of the models. The water vapor mixing ratios at 50 hPa in the models vary from around 2 ppmv in some models to over 6 ppmv in others. Most models have a minimum water vapor mixing ratio ("hygropause" [*Kley et al.*, 1979]) near the tropical cold point (see Figure 6b). However, there is a large spread in the value of the hygropause, with the minimum tropical mean March water vapor mixing ratio varying between 1 and 4.5 ppmv. The observed minimum from HALOE is around 3 ppmv and the uncertainty in the absolute value of the HALOE measurements is around 10% [*Kley et al.*, 2000].

[53] Figure 7 shows the annual cycle of equatorial temperatures and water vapor mixing ratios at 100 hPa. In general the lower-stratospheric water vapor mixing ratio is a function of the model temperature near the tropical tropopause at 100 hPa, and model-model variations in tropical tropopause temperatures can explain much of the variation in tropical lower-stratospheric water vapor. One exception is the UMSLIMCAT which lacks any seasonal cycle in water vapor, because the chemical water vapor field in the simulation is not coupled to the dynamical core of the model. However, the cold point is not necessarily at 100 hPa which causes some deviations, but for consistency with *Pawson et al.* [2000], we show the 100 hPa values. Compared to the models examined by *Pawson et al.* [2000], the annual cycle of tropical temperature at 100 hPa has improved markedly and most models reproduce the phase of the seasonal cycle of temperature and water vapor amount reasonably well. The models assessed here nearly all have the correct annual cycle, and the spread of values is only 10 K rather than 20 K.

[54] Dehydration can also cause differences in the simulated water vapor amount in polar regions (Figures 6a and 6c). There is a wide spread in simulated stratospheric water vapor mixing ratios in polar lower stratosphere in spring of both hemispheres (varying between 1 and 7 ppmv), which is linked to the large differences in polar temperatures (see Figure 1). Note that CMAM and UMSLIMCAT do not simulate the process of dehydration in these runs.

[55] The annual variations in water vapor mixing ratios shown in Figure 7 propagate vertically into the tropical stratosphere. This is illustrated for the CCMs and HALOE in Figure 8 which shows the time-height sections of water vapor mixing ratio calculated as the deviation (in parts per million by volume) from the time mean profile, averaged between 10°S and 10°N. Note that UMSLIMCAT is not included in the figure because, for the reason mentioned above, it does not have a tape recorder signal. In addition to the model-model and model-data differences in the amplitude around 100 hPa there are also differences in the attenuation of the amplitude and the speed of propagation ("phase propagation") of the water vapor signal. These differences are due to differences in tropical transport, and comparison of these quantities provides information on the transport differences among the models [e.g., *Hall et al.*, 1999]. The

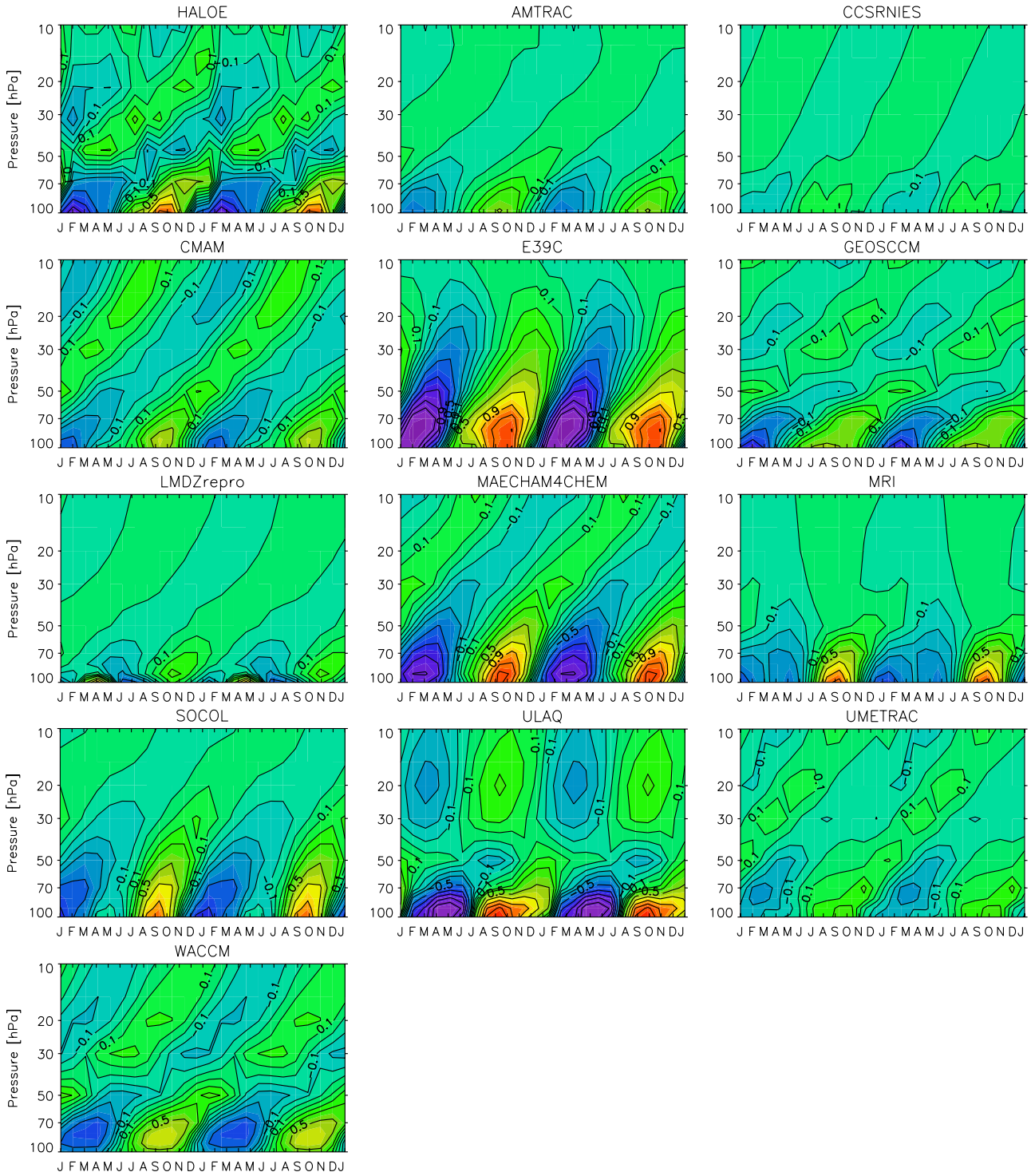

Figure 8. Time-height sections of water vapor mixing ratio shown as the deviation (in parts per million by volume) from the time mean profile, averaged between 10°S and 10°N ("tape recorder") for all CCMs and HALOE data. Two consecutive cycles are shown.

vertical variation of the normalized amplitude and phase propagation of the water vapor signals in the models and in the HALOE observations is shown in Figure 9.

[56] To focus on the decay of the amplitude with height due to mixing rather than temperature-induced differences in the amplitude near 100 hPa, the amplitudes shown in Figure 9a are normalized to unity at a reference level. In previous comparisons [e.g., *Hall et al.*, 1999; *Douglass et al.*, 1999] this reference level is 16 km. However, several CCMs have maximum amplitude above 16 km. Because of this we use the level where the maximum value occurs (between 16 and 20 km) as the reference level. Figure 9a shows that the decay of amplitude with height (attenuation) in most models is in reasonable agreement with HALOE.

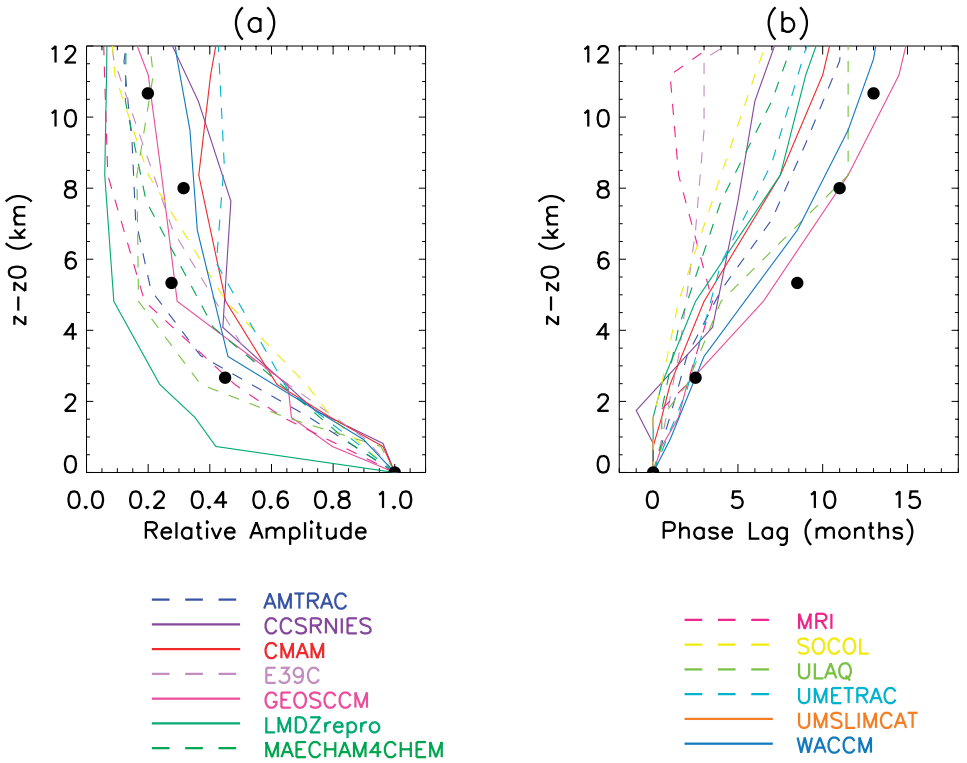

Figure 9. Vertical variation of (a) amplitude and (b) phase lag of annual cycle of water vapor averaged between 10°S and 10°N. The amplitude is normalized to unity and phase lag is set to zero at the level where the amplitude is maximum (between 16 and 20 km), which varies between the CCMs. The vertical axis in both plots is the distance from level of maximum amplitude. Solid circles are HALOE observations.

There is too rapid decay in the lower stratosphere in LMDZrepro and in the middle stratosphere (above 20 km) in MRI, and too little decay in CCSRNIES, CMAM and UMETRAC. However, in the other models the agreement with HALOE is good.

[57] The phase propagation of the water vapor tape recorder signal (Figure 9b) is calculated as the average propagation of the maximum and minimum phases of the signal, and the phase lag is taken zero at the level where the amplitude is maximum. The phase propagation in GEOSCCM and WACCM is in good agreement with the HALOE observations, but the propagation is too fast in the remaining models. For several models the propagation time over 10 km is only a few months less than HALOE, but for around half the models the propagation time is around or less than half that of HALOE (CCSRNIES, E39C, MAECHAM4CHEM, MRI, and SOCOL). Interestingly, differences from the observed CH_4 distribution were also found in these models (see Figure 5).

[58] Differences from observations in both CH_4 and the propagation/attenuation of the water vapor annual cycle indicate deficiencies in the transport in these models. However, in most cases it is not known what aspects of the transport are causing the differences in the tracer fields. Further analysis of the water vapor signal in the models may yield information on the relative roles of ascent, mixing with midlatitudes, and vertical diffusion in the models that can be compared with similar inferences from observations [e.g., *Hall and Waugh*, 1997; *Mote et al.*, 1998; *Hall et al.*, 1999]. However, this is beyond the scope of this paper. The results shown here are in much better agreement with HALOE than the majority of the models in the MMII study [*Hall et al.*, 1999], where nearly all the models overattenuated the signal. Note that although the attenuation with height is reasonable in many models, the attenuation per wavelength (as considered by *Hall et al.* [1999]) will be too strong in many of these models because of the too rapid propagation of the signal.

4.3. Mean Age of Air

[59] The mean age of air is defined as the mean time that a stratospheric air mass has been out of contact with the well-mixed troposphere. This diagnostic can be inferred from observations of conserved tracers with an approximately linear increase in concentration with time (e.g., CO_2, SF_6, total chlorine), and has been used in previous model-data comparisons [e.g., *Hall et al.*, 1999; *Sage et al.*, 1999]. These previous comparisons showed that most models, independent of whether 2- or 3-dimensional, had mean ages that were significantly lower than observations. The cause of the low bias in the models considered in these studies could not in general be attributed to any single aspect of the models.

[60] Figure 10 shows the latitudinal variation of the annual mean age of air at 0.5 hPa (Figure 10a), 10 hPa (Figure 10b), and 50 hPa (Figure 10c) from several CCMs, together with observations (symbols). Not all CCMs included a tracer that enables the mean age to be calculated, and the

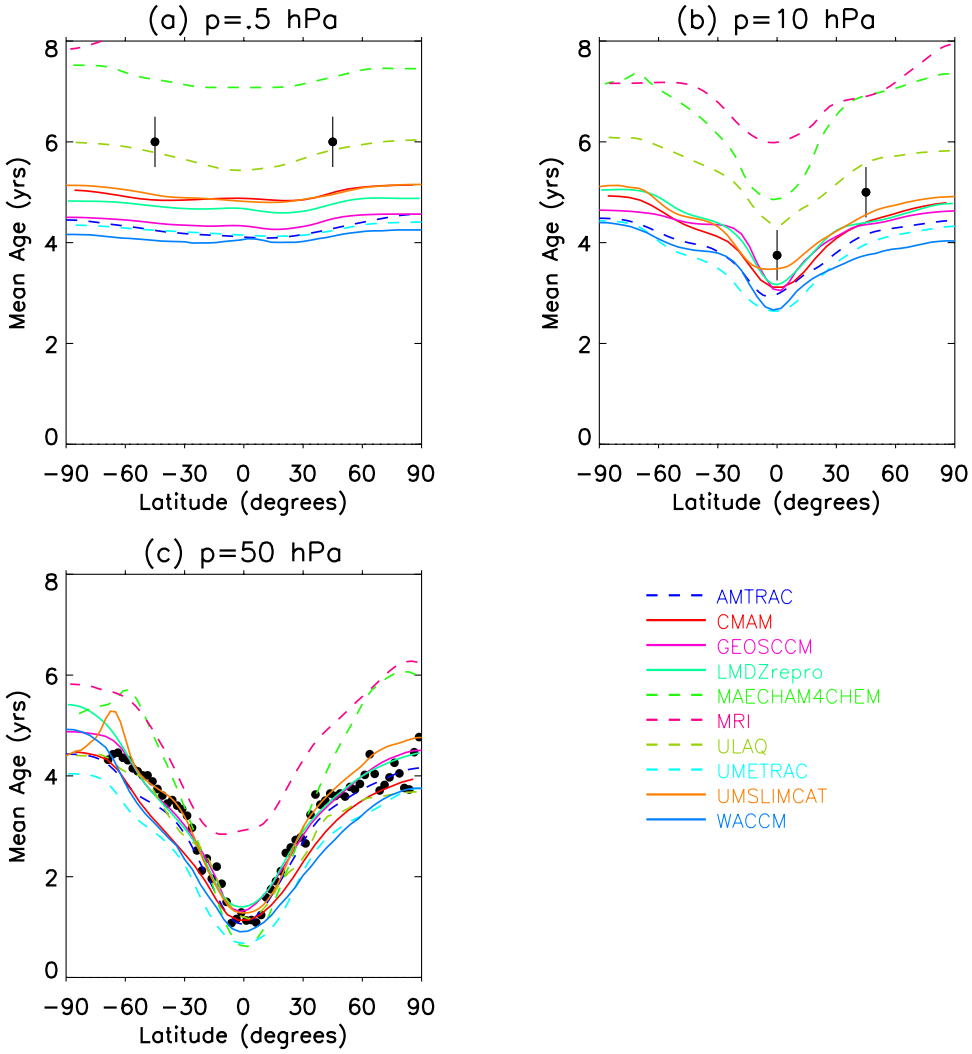

Figure 10. Mean age of air at (a) 0.5, (b) 10, and (c) 50 hPa. Symbols correspond to observations; see text for details.

"age tracers" also varied between the CCMs (e.g., idealized linearly increasing tracer, CO_2 tracer, total chlorine, or tracer with internal source). In some cases two age tracers were available, and produced almost identical results. The age of air observations in Figure 10 are based on (1) ER-2 aircraft measurements of CO_2 from many different years [*Andrews et al.*, 2001], (2) much more limited balloon CO_2 measurements made in northern midlatitudes [see *Waugh and Hall*, 2002, Table 1; also *Aoki et al.*, 2003], and (3) satellite measurements of HF and HCl from HALOE [*Anderson et al.*, 2000]. Note that some age of air estimates derived from SF_6 observations indicate ages significantly higher than 6 years, but these are now thought to be biased high because of mesospheric loss processes [see *Waugh and Hall*, 2002, and references therein].

[61] There is very good agreement among the CCMs and observations at 50 hPa, with the exception of MRI where the age of air is 1–2 years higher than observed at all latitudes, MAECHAM4CHEM where it is 1–2 years higher at high latitudes, and UMETRAC where it is 0.5 years lower at all latitudes. The agreement is not as good for higher altitudes, where most models have lower age of air than observed, and MRI and MAECHAM4CHEM have higher age of air (the MRI age at 0.5 hPa is greater than 8 years). The one exception is the ULAQ model where the age of air agrees with the observations at 0.5 hPa, and is slightly larger at 10 hPa.

[62] As with the tape recorder comparisons, the model-observation agreement in Figure 10 is much better than in the NASA MMII multimodel assessments [e.g., *Hall et al.*, 1999]. In the *Hall et al.* [1999] study the mean ages from all but one of the over twenty 2-D and 3-D models examined were lower than observed at 50 hPa (and most were significantly lower), and all but three had mean ages less than 4 years at 10 hPa in northern midlatitudes. So, although the mean age of air in the CCMs assessed here is generally lower than observed in the middle and upper stratosphere, the differences are generally not as large as previously reported.

4.4. Discussion

[63] Some general conclusions on the transport in the CCMs can be drawn from the above comparisons of

methane, mean age, and water vapor tape recorder signal with observations. In around half of the models the simulated tracer fields are in reasonable agreement with the observations, for all or the majority of the diagnostics. Furthermore, for these models the agreement for mean age and the tape recorder signal is much better than it was in nearly all models reported in the MMII study [*Hall et al.*, 1999]. This indicates that the transport in the models assessed here is reasonable, and some of the previously reported transport problems may no longer apply. Note that the transport analysis has focused on the tropics and middle latitudes, and some of these models still may have major problems in polar transport.

[64] On the other hand, for several models (CCSRNIES, E39C, MAECHAM4CHEM, MRI, SOCOL, and ULAQ) there are large differences from observations in several (if not all) of the transport diagnostics considered here. The cause of the problems in these models is generally not known, and will require further analysis which is beyond the scope of this study. However, in some cases there are indications that the problems are related to model setup. For example, in E39C the most likely cause of the high CH_4 mixing ratios in the extratropics is the low upper boundary centered at 10 hPa. Although the downward transport at high latitudes is fairly well captured, the poleward transport in the model near the upper boundary is too strong, which leads to high CH_4 values in the extratropics [*Grewe*, 2006]. In ULAQ the transport deficiencies are likely related to very low horizontal resolution (10° by 22.5°), which is too low to reproduce observed barriers to transport (and steep tracer gradients) at the edge of the tropics and polar vortices.

[65] In general, the major issue with identifying the problems is that tracer distributions depend on the balance of advective and diffusive processes and it is not straightforward to link a difference in a tracer field to a particular transport process when evaluating models. As an illustration consider the tropical transport in MAECHAM4CHEM, MRI, and SOCOL. The rapid phase propagation of the tape recorder (Figure 9b) in these models suggests too rapid vertical motion (the phase is in general only weakly influenced by tropical-midlatitude mixing [*Hall et al.*, 1999]). However, the vertical gradients of CH_4 (Figure 5b) tend to be too strong particularly in MRI and MAECHAM4CHEM, which is the opposite of what would be expected for too rapid ascent. The vertical gradient of the age of air in MRI and MAECHAM4CHEM is too large, which is also not a symptom of too rapid ascent. Therefore other transport processes must be playing a role. One possibility is that the errors in the tropical CH_4 and mean age of air are due to overly strong tropical-extratropical mixing (this mixing has a larger impact on these tracers than the phase of the tape recorder). While this could be the case in some models (e.g., MRI which has weak subtropical gradients of CH_4 and rapid attenuation of the tape recorder above 20 km) it does not appear to be the case for MAECHAM4CHEM, which has strong subtropical gradients of CH_4 similar to what is observed and reasonable attenuation of the tape recorder amplitude (although the attenuation per wavelength is too strong). Thus, although we have identified problems in the transport in these models in most cases more detailed analysis of the models and output fields is needed to identify the cause.

5. Hydrogen Chloride (HCl) and Inorganic Chlorine (Cl_y)

[66] The accumulation of halogenated compounds in the stratosphere over the past 40 years has been the primary driver of stratospheric ozone depletion. It is therefore important to assess how realistic the simulated inorganic chlorine (Cl_y) is in the CCMs with a particular focus on polar spring Cl_y in the Southern Hemisphere. There are only limited observational data available for Cl_y (i.e., not enough to form a climatology). We therefore compare simulated hydrogen chloride (HCl) with HALOE measurements. HCl is one of the principal components of Cl_y, and overall the model-model differences in HCl are similar to those in Cl_y (not shown).

[67] Figure 11 shows the climatological mean HCl mixing ratios from the CCMs and HALOE. In these plots November and April, rather than March and October as for other species in previous figures, are shown as the HCl/Cl_y ratio is generally higher. As with the other trace gases considered above there is a large spread among the models. All the CCMs used the same, or very similar, halogen boundary conditions so the differences are not likely due to differences in tropospheric concentrations. A more likely cause is differences in transport, and, consistent with this, the CCMs that have problems simulating the CH_4 distribution also generally have problems simulating HCl. For example, in the tropical lower stratosphere HCl mixing ratios in E39C are biased low compared to HALOE while HCl mixing ratios in MRI, MAECHAM4CHEM and ULAQ are biased high (Figures 11b, 11d, and 11e).

[68] However, transport differences cannot explain all of the model-model differences, and other factors play a role. For example, lower-stratospheric HCl (Figures 11d and 11e) and Cl_y mixing ratios (not shown) in AMTRAC and UMETRAC are generally higher than in other models, yet CH_4 (Figure 5) and mean age of air (Figure 10), and hence transport in these two models is similar to that in several other models. The photolysis rates of organic chlorine species are adjusted in AMTRAC and UMETRAC by about 25% to account for the low bias in the simulated mean age of air so that the Cl_y is in closer agreement with the observed upper stratospheric Cl_y. Hence these models simulate higher Cl_y and HCl values than models with similar mean ages, but they also have very high HCl relative to HALOE throughout the extrapolar lower stratosphere.

[69] The extent to which high Cl_y and HCl values in polar spring reach down to the lower altitudes reflects the degree of unmixed descent from the upper stratosphere within the polar vortex (Figures 11a, 11c, and 12a). In AMTRAC and UMETRAC the high Cl_y and HCl values reach deeper into the lower stratosphere than in all other models. On the other hand, E39C Cl_y (HCl) profiles are generally shifted upward which leads to low Cl_y (HCl) mixing ratios at nearly all levels. A first test with a fully Lagrangian advection scheme instead of the semi-Lagrangian scheme employed indicates significant improvements with regard to the structure of the profiles (A. Stenke and V. Grewe, Lagrangian transport of water vapor and cloud water in the ECHAM4 GCM and its

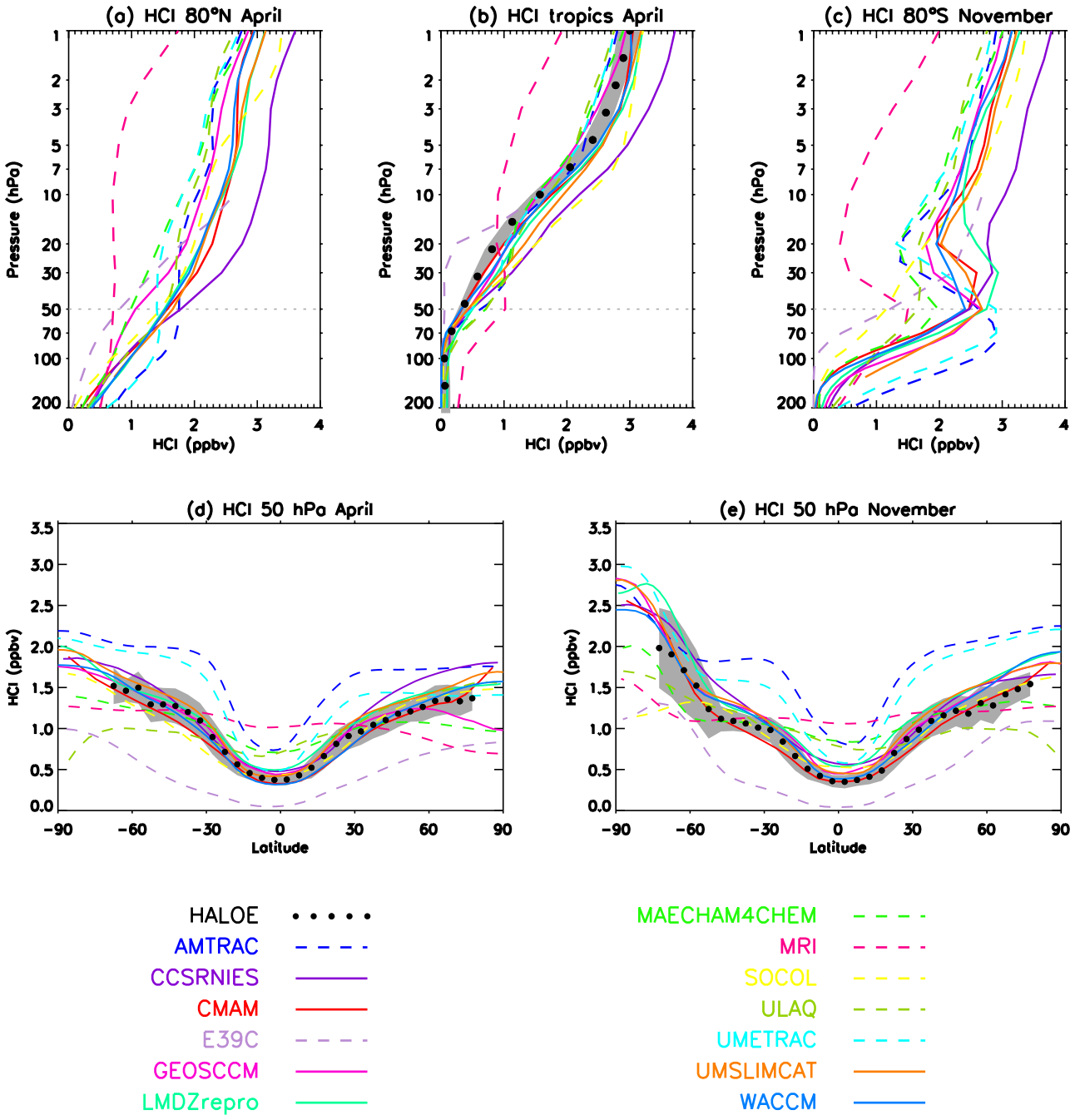

Figure 11. As in Figure 5 but for HCl in ppbv and vertical profiles are shown at (a) 80°N in April, (b) 0° in April, and (c) 80°S in November. Latitudinal profiles at 50 hPa in (d) April and (e) November.

impact on the cold bias, submitted to *Climate Dynamics*, 2006). SOCOL has the lowest Cl_y values among all models at 50 hPa in the polar Southern Hemisphere in spring.

[70] The vertical profile of Cl_y also indicates some serious problems in the upper stratosphere in CCSRNIES and SOCOL with values of Cl_y that are larger than the maximum total chlorine entering the stratosphere, which is physically impossible. The cause for this in the two models is unknown, and is currently being investigated. At the other extreme the MRI model simulates unrealistically low values of Cl_y in the upper stratosphere. These low values are consistent with low CH_4 (Figure 5) and very high mean age (Figure 10) in this model, as a high mean age would delay the buildup of chlorine in the upper stratosphere, but even ages higher than 8 years cannot explain the low Cl_y in the upper stratosphere of the MRI model. Analysis of total chlorine (sum of organic (CCl_y) and inorganic (Cl_y) chlorine) from this model (not shown) indicates that the total chlorine in the upper stratosphere is not only shifted in time from the troposphere but also increases at a much slower rate, so that the time lag from the troposphere increases with time.

[71] All models show Cl_y increasing rapidly during the 1980s and early 1990s in the Southern Hemisphere polar lower stratosphere in October (Figure 12b). However,

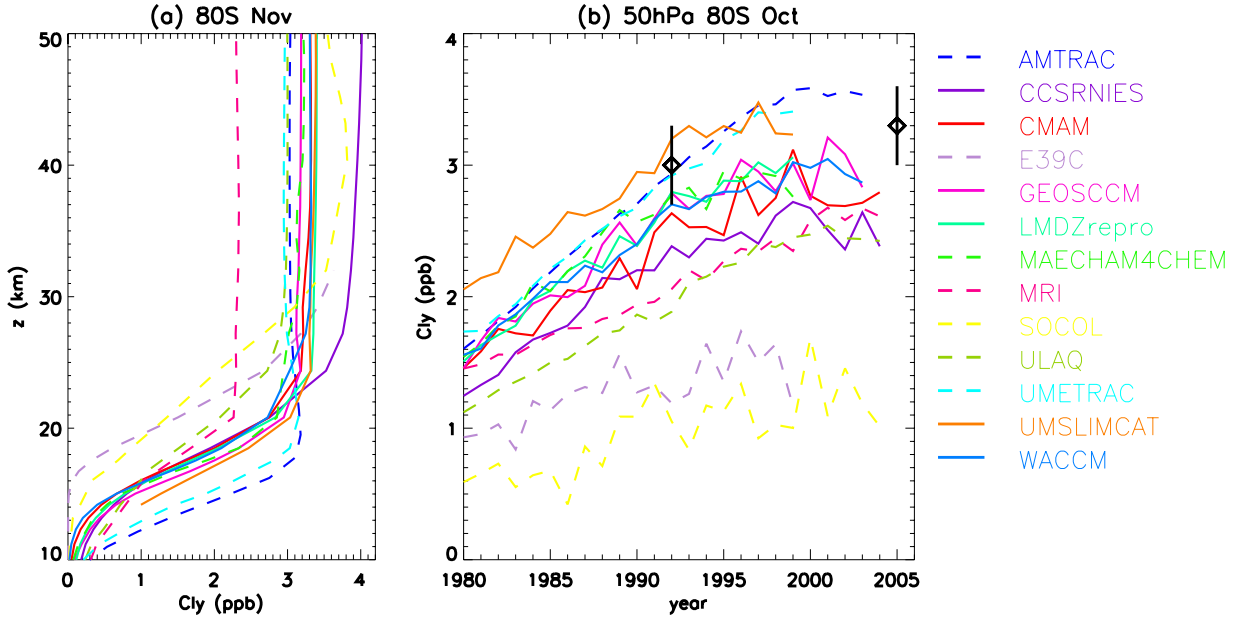

Figure 12. (a) Climatological mean vertical profiles (1990 to 1999) at 80°S in November for Cl_y in ppbv. (b) Time series of October mean Antarctic Cl_y at 80°S from CCM model simulations. Estimates of Cl_y from HALOE HCl measurements in 1992 [*Douglass et al.*, 1995; *Santee et al.*, 1996] and Aura MLS HCl in 2005 (M. Santee, personal communication, 2006) are shown in addition.

consistent with the above analysis, there are large variations in the simulated peak Cl_y (from around 1.2 ppbv to over 3.5 ppbv). Note that the UMSLIMCAT run had not fully spun up by 1980 and did not capture the correct stratospheric lag compared to the troposphere until around 1985. Estimates of Cl_y from measurements in the southern polar lower stratosphere in spring can be made from measurements of HCl in this region, as $Cl_y \sim$ HCl when there are very low values of O_3 [e.g., *Douglass et al.*, 1995]. Measurements of HCl by UARS HALOE in 1992 yield a value around 3 ppbv [*Douglass et al.*, 1995; *Santee et al.*, 1996], whereas measurements by Aura MLS in 2005 yield a mean value around 3.3 ppbv (M. Santee, personal communication, 2006), see symbols in Figure 12b. There are large uncertainties (10–15%) in these values because of the limited coverage in the polar region (for HALOE) and possible biases in the HCl measurements [*Froidevaux et al.*, 2006]. However, even with this uncertainty it is clear that peak Cl_y values around or less than 2.5 ppbv, as simulated in several CCMs (especially in SOCOL and E39C) are too low.

6. Ozone

[72] Differences between observations and some of the models as outlined in the previous sections affect the capability of the models to simulate realistic ozone distributions. Temperature biases affect the temperature-dependent chemical reaction rates globally and lead to significant differences in simulated PSCs in polar regions [*Austin et al.*, 2003]. Deficiencies in transport lead to errors in ozone transport as well as in transport of the chemical species that react with ozone. The ability of CCMs to reproduce past stratospheric chlorine concentration is important to track observed ozone evolution globally, but is of particular importance in the Antarctic. In section 6.1 we compare modeled vertical and horizontal ozone profiles in different seasons to HALOE data and total ozone climatologies to two different data sets. In section 6.2 polar and global mean total ozone time series are compared to four different data sets. The data sets are described in more detail in section 2.3.

6.1. Climatological Mean Ozone

[73] Figure 13 compares climatological mean vertical profiles and latitudinal cross sections of ozone derived from the models and HALOE. In the tropical lower stratosphere between 100 and 30 hPa most models (9 out of 13 CCMs) agree well with HALOE observations and lie within 1σ of the HALOE mean. MAECHAM4CHEM, MRI and SOCOL are biased high, which is consistent with low CH_4 (see Figures 5a–5c and discussion in section 4.1). At higher altitudes (above 7 hPa) these four models lie within the 1σ HALOE mean, whereas peak values of ozone in the tropics are biased high in WACCM and CCSRNIES. The tropical stratosphere in WACCM is dynamically isolated and the mean age of air is too low near the ozone volume mixing ratio maximum (WACCM has lowest age of air of all models at 10 hPa, see Figure 10) leading to NO_x mixing ratios that are too low by approximately 10 to 20%. In this region, the photochemical lifetime of ozone is between 5 and 10 days [*Solomon et al.*, 1985] and the primary odd-oxygen catalytic loss cycle is via NO_x, consistent with the positive bias. CCSRNIES overestimates peak values of ozone not only in the tropics but also in both polar regions, likely because of an overestimation of O_2 photolysis rates at this altitude.

[74] In the southern hemispheric polar spring at 80°S (Figure 13c) most models are biased high compared to HALOE observations below 50 hPa and biased low between 50 and 20 hPa. The most pronounced ozone bias is evident in LMDZrepro, which simulates very low ozone mixing ratios at 80°S in October even though Cl_y, HCl,

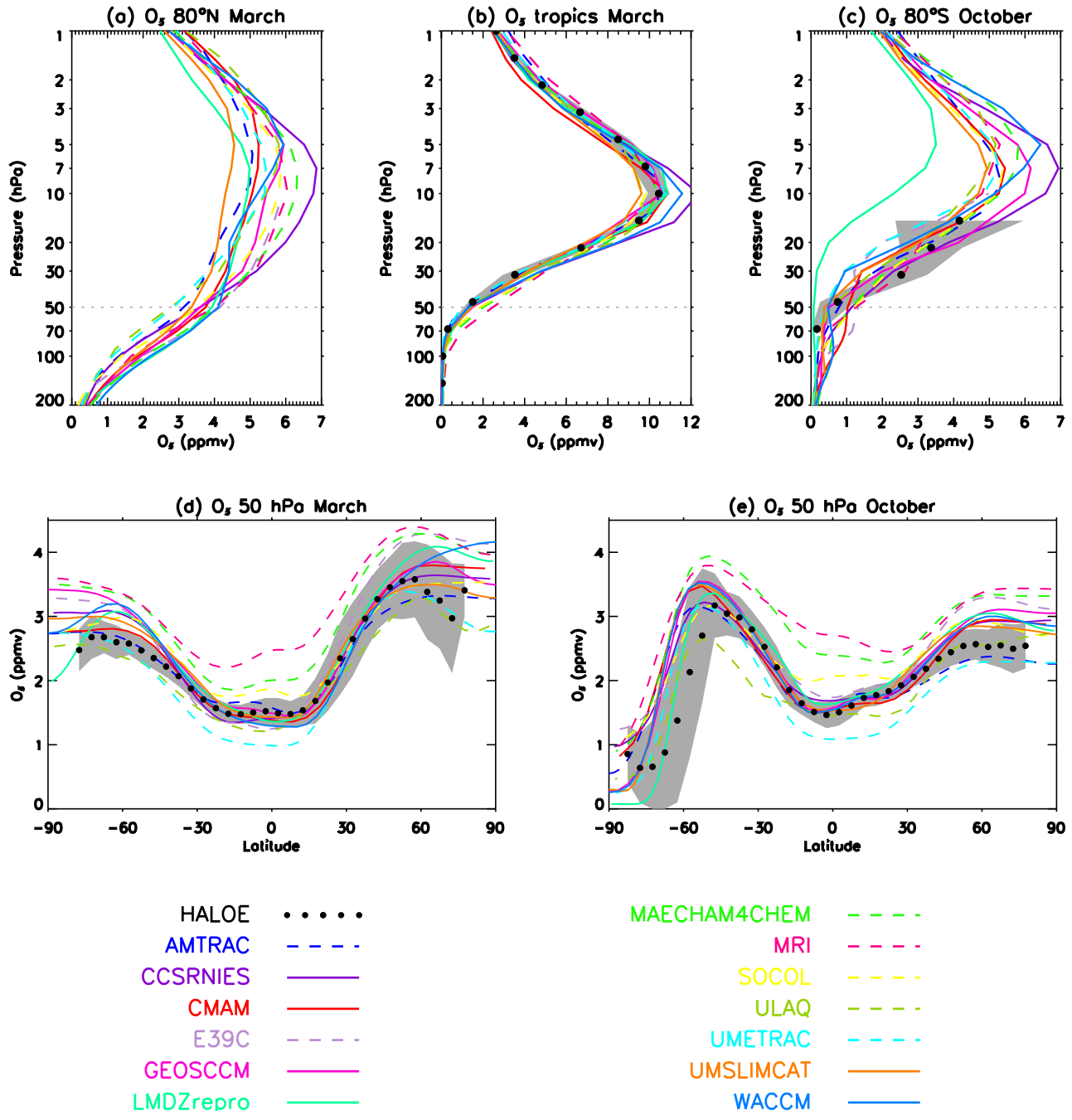

Figure 13. (a–e) As in Figure 5 but for O_3 in ppmv.

CH_4, and H_2O are not anomalous. This bias originates from the 20 K bias in temperature. The LMDZrepro temperatures in October and November stay well below the threshold temperature for PSC formation in the simulation whereas in reality the temperature should increase above the PSC formation temperature. Interestingly, this large negative ozone bias reinforces the temperature bias, as the temperature bias is less pronounced in the underlying GCM simulation [*Lott et al.*, 2005]. The two models that show lowest Cl_y values (E39C and SOCOL) in the lower stratosphere (Figure 12) are also the ones that simulate highest ozone values in the Southern Hemisphere polar regions in October at 50 hPa (Figure 13e).

[75] Figure 14 compares the 20-year mean climatological total column ozone from 1980 to 1999 from models, merged satellite data and NIWA assimilated data. Detrended mean annual cycles for the 1980s are also shown for the polar regions and globally in Figure 15. The well-known features of highest ozone values in northern spring, low ozone values in the tropics with a small seasonal cycle, a relative ozone maximum in the midlatitudes of the Southern Hemisphere in late winter/early spring, and a minimum ozone column above the Antarctic are represented in all models. With the exception of the Southern Hemisphere high latitudes, the amplitude and phase of the annual cycle in individual latitude bands is generally well simulated by the

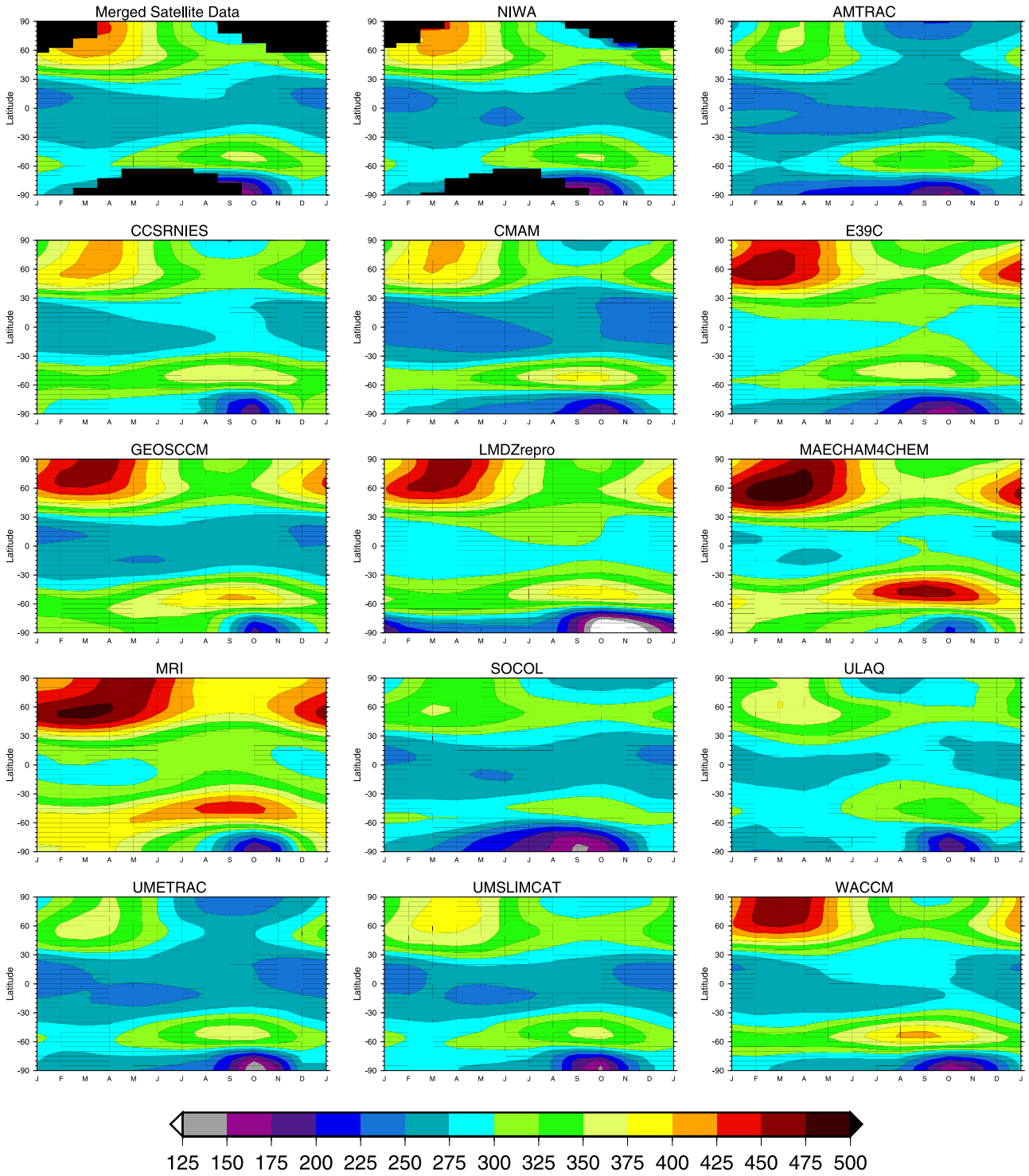

Figure 14. Modeled total column ozone climatologies (1980 to 1999) compared to merged satellite and NIWA assimilated data.

models. However, most models exhibit significant annual mean biases which depend on latitude. The Northern Hemisphere high-latitude spring values are too high in a number of models (E39C, GEOSCCM, LMDZrepro, MAE-CHAM4CHEM, MRI, and WACCM) and in these models the midlatitude Southern Hemisphere maxima in late winter/early spring are also too high. For some of these models this is likely an indication of a generally too strong Brewer-Dobson circulation (see also the meridional ozone gradients in Figures 13d and 13e), although it should be noted that some of these models exhibit a high bias in global mean total ozone (see Figure 15, right plots), especially MAE-CHAM4CHEM and MRI. Moreover, the high Southern Hemisphere midlatitude maximum is also consistent with the general Southern Hemisphere cold pole bias, which implies an overly strong transport barrier at the edge of the

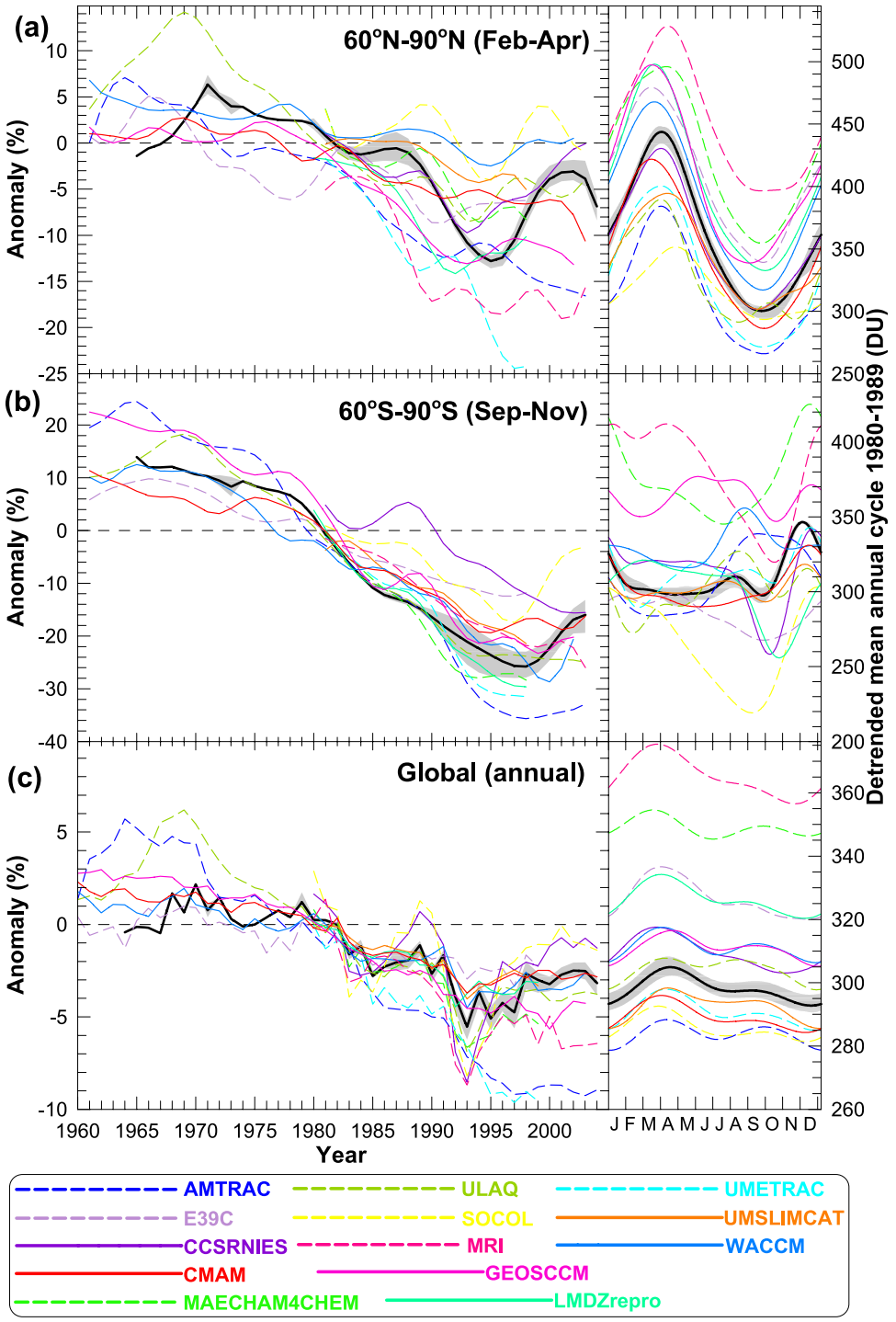

Figure 15. (a) Seasonal (February to April) total column ozone anomalies for the Arctic (60–90°N), (b) seasonal (September to November) total column ozone anomalies for the Antarctic (90–60°S), and (c) annual total column ozone anomalies sets for the whole globe (90°S–90°N) from CCMs and four observational data sets. The CCM results are shown with colored lines while the mean and range of the four observational data sets are shown as a thick black line and grey shaded area respectively. The anomalies are calculated with the method described in Appendix A. The seasonal anomaly time series shown in Figures 15a and 15b have been smoothed by applying a 1:2:1 filter iteratively five times. The filter width is reduced to one at the ends of the time series. The annual global anomalies shown in Figure 15c are unsmoothed. The plots on the right show detrended mean annual cycles for each model (1980 to 1989) and the mean and range of the observations.

Table 3. Mean Antarctic Ozone Depletion Indices for the 1990s Plus Interannual Standard Deviations for Observations and CCMs[a]

Model	Minimum SH ± SDs, DU	OHA ± SDs, million km^2	OMD ± SDs, million tonnes
NIWA database	94.3 ± 11.2	25.4 ± 2.0	23.8 ± 5.4
AMTRAC	80.5 ± 27.0	19.9 ± 5.0	22.0 ± 9.6
CCSRNIES	133.4 ± 29.2	19.4 ± 4.7	10.8 ± 6.9
CMAM	115.7 ± 38.2	17.3 ± 4.5	13.4 ± 8.8
E39C	119.8 ± 10.7	15.8 ± 1.6	8.7 ± 2.1
GEOSCCM	118.5 ± 21.8	16.7 ± 3.7	8.9 ± 5.2
LMDZrepro	48.1 ± 4.4	25.1 ± 2.4	36.3 ± 7.0
MAECHAM4CHEM	133.4 ± 16.0	13.6 ± 2.1	4.2 ± 1.8
MRI	110.9 ± 16.9	15.2 ± 2.9	8.0 ± 3.5
SOCOL	82.4 ± 26.3	35.2 ± 1.6	37.5 ± 7.5
ULAQ	85.1 ± 10.6	24.6 ± 4.6	20.1 ± 7.5
UMETRAC	70.4 ± 22.0	24.2 ± 5.0	28.6 ± 1.0
UMSLIMCAT	100.4 ± 6.7	16.6 ± 2.2	17.2 ± 5.1
WACCM (v.3)	125.9 ± 36.9	11.5 ± 4.9	8.4 ± 5.2

[a]The minimum total ozone column poleward of 60°S, the maximum Antarctic ozone hole area (OHA) between September and October defined by the 220 DU contour, and the ozone mass deficit calculated as the mean of daily values between September and October based on a 220 DU threshold (OMD) have been calculated for each year and then averaged over the 10 years from 1990 to 1999.

vortex. In contrast to many models, SOCOL and ULAQ have relatively low Southern Hemisphere midlatitude maxima, which are likely associated with too rapid horizontal mixing with the tropics (in particular for ULAQ), as reflected in the weak latitudinal gradients of ozone and other tracers (see section 4).

[76] There are also significant differences in the evolution and the size of the Antarctic ozone hole among the models themselves and in comparison with observations (Figure 14). Several ozone depletion indices have been used in the past to assess the severity of polar ozone depletion [e.g., *Austin et al.*, 2003; *Bodeker et al.*, 2005; *WMO/UNEP*, 2003], and these indices have been calculated from the CCMs here. In particular, the minimum total ozone column poleward of 60°S, the maximum Antarctic ozone hole area (OHA) between September and October defined by the 220 DU contour, and the ozone mass deficit calculated as the mean of daily values between September and October based on a 220 DU threshold (OMD) have been calculated for each year. Table 3 compares the 1990 to 1999 average of these ozone depletion indices plus interannual standard deviations of the 10 years from models and the NIWA assimilated database. A 10-year rather than a 20-year period is used as there were large trends in the Southern Hemisphere indices from 1980 to 1990. The modeled minimum Antarctic ozone values are in good agreement with observations and lie within 25 DU of the observed minimum in most of the models, but not in LMDZrepro. This model severely underestimates column ozone throughout Southern Hemispheric spring and even winter (Figure 14) because of a cold pole problem (Figure 1) and accordingly also underestimates the minimum ozone and interannual variability. The interannual variability in most of the other CCMs is too high and is significantly overestimated in CMAM and WACCM. In WACCM the high variability is due mainly to the behavior in a single year (1991), when ozone depletion was almost completely absent because of the occurrence of several large amplitude wave forcing events. In most models the simulated ozone hole is too small, both in terms of ozone hole area and ozone mass deficit, and the date of maximum ozone hole area occurs later than observed by 10 to 20 (up to around 40) days (not shown). The later occurrence of ozone hole maximum area is likely related to the cold pole problem and late vortex breakup (see section 3 and Figure 2). The reasons for the smaller ozone hole are in general not known, but it is important to note that indices calculated from a fixed threshold (e.g., 220 DU) do not account for the general high bias in the total ozone fields that occurs in several models (see Figure 14 and right plots in Figure 15). This high bias however often originates from layers that are not interesting for chemical ozone depletion and in most models this is a global and not purely southern hemispheric bias. For example, the seasonal cycle in the Antarctic could be well represented in a model indicating a correct simulation of the processes of ozone depletion, whereas the indices could be underestimated because of a persistent global high bias (e.g., in MAECHAM4CHEM). This points to a weakness in the definition of the ozone indices based on a fixed threshold. In addition, the diagnostics are very sensitive to the detailed structure of ozone at the edge of the vortex, which depends on both transport and chemistry and varies considerably among the models. In this respect it is also noteworthy that the annual cycle of total ozone over 60–90°S varies considerably among the models (Figure 15b). *Tegtmeier and Shepherd* [2006] have shown how the annual cycle of total ozone in this region can vary tremendously between different versions of the CMAM, even while the basic photochemical relationship between springtime and summertime ozone anomalies is maintained.

6.2. Total Ozone Variations and Trends

[77] Changes in stratospheric ozone are closely coupled to changes in temperatures, which have been shown to be in reasonable agreement with satellite and radiosonde observations in section 3. In this section we examine the CCMs' ability to track past global and polar temporal changes in ozone. To calculate the total column ozone anomalies a regression model is used to recreate the "detrended mean annual cycle," i.e., the average annual cycle over the period 1980 to 1989 that would occur in the absence of trends. This detrended mean annual cycle is subtracted from the total ozone time series (see Appendix A).

[78] Figure 15 shows the area weighted global annual (90°S to 90°N) as well as spring time polar (60° to 90°) monthly mean total column ozone anomalies from the CCMs and observations (left plots) and the detrended mean annual cycles calculated using the method described in

Appendix A (right plots). The measurement precision for the individual total column ozone measurements making up the observational data sets is typically a few percent, and spatial and temporal averaging reduces this further. The features already discussed for the 1980–1999 climatology in section 6.1 are also valid for the 1980s: Apart from differences in the mean value among the models and the observations, most models reproduce the observed amplitude and phasing of the observed mean annual cycle. The MRI model shows the greatest difference in global mean annual cycle in comparison to observations (mean of 366 DU compared to observed mean of 299 DU, see Figure 15c, right plot). This bias is consistent with low CH_4 (section 4.1), older mean age of air (section 4.3) and very low Cl_y and HCl concentrations in the upper stratosphere (section 5). The high global mean annual cycle for MAECHAM4CHEM is also consistent with low CH_4.

[79] Over the Antarctic all models show negative trends in ozone, but with some models (UMETRAC and AMTRAC) showing negative anomalies larger than observed. Others (e.g., CCSRNIES, E39C, and SOCOL) show anomalies about half those observed. Current understanding of Antarctic ozone loss suggests that the observed record over decadal timescales is mainly driven by halogen loading [*WMO/UNEP*, 2003]. Results of AMTRAC [*Austin and Wilson*, 2006] show a strong correlation between halogen loading and Antarctic column ozone over the period 1960 to 2004 (and beyond). Accordingly, the models with the largest (UMETRAC and AMTRAC) and smallest (SOCOL and E39C) Cl_y trends (see Figure 12b and section 5) exhibit respectively too strong and too weak ozone trends.

[80] Over the Arctic, the February to April means of total column ozone anomalies show a large range of ozone trends and large differences in comparisons with observations with some models (SOCOL and WACCM) showing little or no change in Arctic ozone since 1980, while other models (e.g., in particular UMETRAC, and to a lesser extent AMTRAC and MRI) show large negative trends. The observed record of Arctic springtime ozone over decadal timescales varies according to the combined effects of chlorine loading and dynamical variability [*WMO/UNEP*, 2003], so differences in model behavior should reflect differences in one or both of these processes. Similar to the Antarctic, SOCOL with the smallest trend and AMTRAC and UMETRAC with the largest trends in Cl_y have Arctic ozone trends that are respectively too small and too large. Consideration of Arctic late winter/spring temperature trends (Figure 4, section 3.2) indicates that MRI exhibits the most excessive cooling of all models, and accordingly this model has an Arctic ozone trend that is too large.

[81] The CCMs show a wide range of behavior with significant differences in global mean total ozone trends and variability over the period 1980 to 2000, which is partly expected as not all models include the parameterization of the solar cycle and volcanic eruption effects (see Table 2). However, the majority of models represent the observed time evolution between 1980 and 2000 reasonably well, and only AMTRAC and UMETRAC and to a lesser extent MRI show a much larger than observed trend after 1980, whereas in SOCOL, CCSRNIES and E39C the trend is slightly underestimated (see discussion above). Although the magnitude differs, most of the models that include volcanic eruption effects capture the early 1990s minimum seen in the observations (CCSRNIES, CMAM, E39C, LMDZrepro, MRI, MAECHAM4CHEM, SOCOL, UMSLIMCAT, ULAQ, and WACCM), while others (AMTRAC and UMETRAC) do not. The differences in the ability of the models to capture the mid-1990s minimum are likely dependent on their ability to capture the response in global ozone to the Mt. Pinatubo volcanic eruption [*Gleason et al.*, 1993]. Only in GEOSCCM are the effects of volcanic eruptions not included (see Table 2). In the six model simulations that extend back to 1960 (AMTRAC, CMAM, E39C, GEOSCCM, ULAQ, and WACCM), two models (E39C and WACCM) show little or no changes prior to 1980, as expected on the basis of chlorine loading and as indicated by observations. However, two models (AMTRAC and ULAQ) show positive ozone anomalies prior to 1980.

7. Summary

[82] In this paper the current generation of coupled chemistry-climate models (CCMs) has been evaluated by comparing the simulated stratospheric thermal structure and distribution of trace gases of the recent past with meteorological analyses and trace gas observations. This study extends the previous multi-CCM assessment [*Austin et al.*, 2003] by considering transient simulations with almost identical forcings (e.g., sea surface temperatures, greenhouse gases, and halocarbons) from a larger number of CCMs. More importantly, transport and distributions of important trace gases including inorganic chlorine are evaluated for the first time in a multi-CCM study.

[83] The models considered in this study reproduce the vertical profile of global annual mean temperature in the stratosphere fairly accurately. However, temperature biases still occur in the high latitudes in winter and spring. For the Northern Hemisphere the errors are rather small and on average the temperature biases in the lower stratosphere are small. In addition most of the models exhibit the correct sensitivity of polar temperature to variations in wave fluxes from the troposphere. In the lower Southern Hemisphere stratosphere the ''cold-pole'' problem remains and most models still have a cold bias with the polar vortex breaking down much later than observed (most severe in LMDZrepro). Many models also still have problems correctly simulating the stratospheric temperature response to tropospheric wave forcing in the Southern Hemisphere in winter and spring (e.g., CCSRNIES, E39C, MAECHAM4CHEM, MRI, LMDZrepro, and ULAQ, with the ULAQ model having the incorrect sign). Modeled global temperature trends are in reasonable agreement with observations and compare well with previous CCM assessments and results from coupled atmosphere ocean models that incorporated prescribed ozone changes. They show a significant cooling trend from 1980 to present day at 50 hPa, with perturbations due to volcanic eruptions. Most models reproduce the observed phase of the annual cycle of equatorial temperatures, but there are variations in the amplitude and, more importantly, the annual mean value, which causes a large spread in the water vapor distributions in the CCMs.

[84] The transport in the CCMs has been assessed by examining the distributions of methane and mean age of air, as well as the upward propagation of the annual cycle in

tropical water vapor. There is in general a large spread in the simulated tracer fields among the models, indicating large variations in model transport. However, for each diagnostic a large fraction of the spread is due to a subset of models (CCSRNIES, E39C, MAECHAM4CHEM, MRI, SOCOL, and ULAQ) where large deviations from observations are apparent in several of the transport diagnostics. The cause of significant biases in the tracer fields in these models is generally not known, but in some cases it can be attributed to specific model features, e.g., very low horizontal resolution (ULAQ) or a low upper boundary (E39C). For the remaining CCMs there is a much smaller spread for all tracers and the tracer fields are in general in good agreement with observations. Previous model-data comparisons have identified serious model deficiencies in simulating the mean age of air and propagation of the water vapor tape recorder signal [*Hall et al.*, 1999], but better agreement is found for many of the CCMs examined here.

[85] There are also substantial quantitative differences and large deviations from observations in the simulated HCl and Cl_y fields. In the polar Southern Hemisphere lower stratosphere most CCMs underestimate peak Cl_y, but two models simulate values that are much too low (SOCOL and E39C). A likely cause of these differences is transport and, consistent with this, the CCMs that have problems simulating the CH_4 distribution also generally have problems simulating Cl_y and HCl. However, transport differences cannot explain all the model-model differences, and other factors play a role in determining the model-model differences in Cl_y. For example, higher than observed Cl_y in the extrapolar lower stratosphere in AMTRAC and UMETRAC results from photolysis rates of organic chlorine species being artificially increased by about 25% so that the Cl_y in the upper stratosphere is in close agreement with observed Cl_y.

[86] The CCMs are generally able to reproduce the observed amplitude and phasing of the mean annual cycle in total column ozone in different latitude bands, except in southern high latitudes. However, most models exhibit large annual mean biases, with the majority overestimating total ozone (most severe in MAECHAM4CHEM). Related to this, the simulated ozone hole is too small in terms of both ozone hole area and ozone mass deficit in many models. Most CCMs show reasonable agreement with observed total ozone trends and variability on a global scale, but a greater spread in the ozone trends in polar regions in spring, especially in the Arctic. Although it is not possible to trace all these differences in the simulated ozone fields to deficiencies in the simulated temperature and tracers, in some cases a link can be made. For example, the models with low climatological mean extrapolar methane values have high ozone there (MAECHAM4CHEM, MRI, and SOCOL), and the model with the largest cold bias in the Antarctic lower stratosphere in spring (LMDZrepro) simulates very low ozone. In the Antarctic, the models with the largest (AMTRAC and UMETRAC) and smallest (SOCOL and E39C) Cl_y trends exhibit respectively too strong and too weak ozone trends.

[87] The results shown here provide insight into the ability of the different CCMs to model key processes and the observed state of the stratosphere. This in turn provides guidance on how to interpret CCM predictions of the future atmosphere. The CCMs vary in their skill in representing different processes and characteristics of the atmosphere. The importance placed on the models ability to represent a particular process or characteristic will, to some degree, depend on the objective of a model prediction. For predicting the recovery of ozone due to declining concentrations of ozone depleting substances, a realistic simulation of the evolution of chlorine (and bromine) in the stratosphere is important. Thus the evaluation of the abilities of CCMs to simulate stratospheric inorganic chlorine presented here will be valuable for interpreting their predictions of ozone recovery. Furthermore, as temperatures and ozone are closely coupled, the temperature biases in the southern lower stratosphere in spring and associated late break up of the vortex in many of the models discussed here will need to be considered when interpreting the CCM future predictions of polar ozone evolution. The predictions from the suite of CCMs are currently being examined and will be described in a follow up study.

Appendix A

[88] The following approach has been adopted to calculate total column ozone anomalies. A regression model of the form

$$O_3 = \sum_{i=0}^{4} A_{2i}\sin(it') + A_{2i+1}\cos(it') + t\cdot \sum_{i=0}^{2} B_{2i}\sin(it') + B_{2i+1}\cos(it')$$

where $t' = 2\pi(t - 0.5)/12$ and t is the number of months after the start of 1980 has been used to fit the modeled monthly mean total column ozone using the standard linear least squares regression method described by *Press et al.* [1989]. The fit generates the coefficients A_i and B_i. The A_1 term accounts for the baseline offset at 1 January 1980 while the $B_1 \times t$ term accounts for the subsequent trend on that offset. The terms with coefficients A_2 to A_9 account for the stationary part of annual cycle, while the terms with coefficients B_2 to B_5 account for seasonally dependent trends in that stationary annual cycle. The coefficients A_i can then be used to recreate the "detrended mean annual cycle," i.e., the average annual cycle over the period 1980–1989 that would occur in the absence of trends. These functional fits, evaluated monthly, are subtracted from the raw data to produce anomaly time series, which are divided by $A_1/100$ to produce percentage anomalies.

[89] **Acknowledgments.** We dedicate this paper to one of our coauthors, Byron Boville, who passed away recently. Byron was a pioneer in the development of global atmospheric models and a great colleague and friend for many of us. Coordination of this study was supported by the Chemistry-Climate Model Validation Activity (CCMVal) for WCRP's (World Climate Research Programme) SPARC (Stratospheric Processes and their Role in Climate) project. We thank the British Atmospheric Data Centre, the SPARC data center, and the UK Met Office for providing the facilities for central data archives. Special thanks go to Georgiy Stenchikov for providing the aerosol heating rates, David Considine for providing sulfate aerosol surface area densities, Steven Montzka for his help extending GHG and halogen time series, and A. R. Ravishankara and Stanley Sander for providing chemical kinetics from JPL 2006 for the CCMVal reference simulations. We also wish to thank Eric Nash for calculating ERA-40 heat flux data and highly appreciate comments on the manuscript by Ulrich Schumann and three anonymous reviewers. JA's research was supported by the Visiting Scientist Program at the NOAA Geophysical Fluid Dynamics Laboratory, administered by the University Corporation for Atmospheric Research. CCSRNIES' research was supported by the Global Environmen-

tal Research Fund (GERF) of the Ministry of the Environment (MOE) of Japan (A-1). CMAM is supported by CFCAS and NSERC. The European groups acknowledge support of the SCOUT-O3 Integrated Project which is funded by the European Commission.

References

Akiyoshi, H., T. Sugita, H. Kanzawa, and N. Kawamoto (2004), Ozone perturbations in the Arctic summer lower stratosphere as a reflection of NO_x chemistry and planetary scale wave activity, *J. Geophys. Res.*, *109*, D03304, doi:10.1029/2003JD003632.

Alexander, M. J., and T. J. Dunkerton (1999), A spectral parameterization of mean-flow forcing due to breaking gravity waves, *J. Atmos. Sci.*, *56*, 4167–4182.

Anderson, J., J. M. Russell III, S. Solomon, and L. E. Deaver (2000), Halogen Occultation Experiment confirmation of stratospheric chlorine decreases in accordance with the Montreal Protocol, *J. Geophys. Res.*, *105*, 4483–4490.

Anderson, J. L., et al. (2004), The new GFDL global atmosphere and land model AM2/LM2: Evaluation with prescribed SST simulations, *J. Clim.*, *17*, 4641–4673.

Andrews, A. E., et al. (2001), Mean ages of stratospheric air derived from in situ observations of CO_2, CH_4, and N_2O, *J. Geophys. Res.*, *106*, 32,295–32,314.

Aoki, S., T. Nakazawa, T. Machida, S. Sugawara, S. Morimoto, G. Hashida, T. Yamanouchi, K. Kawamura, and H. Honda (2003), Carbon dioxide variations in the stratosphere over Japan, Scandinavia and Antarctica, *Tellus, Ser. B*, *55*, 178–186.

Austin, J. (2002), A three-dimensional coupled chemistry-climate model simulation of past stratospheric trends, *J. Atmos. Sci.*, *59*, 218–232.

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

Austin, J., R. J. Wilson, F. Li, and H. Vömel (2006), Evolution of water vapor concentrations and stratospheric age of air in coupled chemistry-climate model simulations, *J. Atmos. Sci.*, in press.

Beagley, S. R., J. de Grandpré, J. N. Koshyk, N. A. McFarlane, and T. G. Shepherd (1997), Radiative-dynamical climatology of the first-generation Canadian Middle Atmosphere Model, *Atmos. Ocean*, *35*, 293–331.

Bloom, S., et al. (2005), Documentation and validation of the Goddard Earth Observing System (GEOS) data assimilation system–Version 4, in *Global Modeling Data Assimilation 104606*, *Tech. Rep. Ser. 26*, NASA Goddard Space Flight Cent., Md.

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

Brühl, C., et al. (1996), HALOE ozone channel validation, *J. Geophys. Res.*, *101*, 10,217–10,240.

Collins, W. D., et al. (2004), Description of the NCAR Community Atmosphere Model (CAM3), *Tech. Note NCAR-TN-464+STR*, Natl. Cent. for Atmos. Res., Boulder, Colo.

Cordero, E. C., and P. M. de F. Forster (2006), Stratospheric variability and trends in IPCC model simulations, *Atmos. Chem. Phys. Disc.*, *6*, 7657–7695.

Dameris, M., et al. (2005), Long-term changes and variability in a transient simulation with a chemistry-climate model employing realistic forcings, *Atmos. Chem. Phys.*, *5*, 2121–2145.

Dameris, M., S. Matthes, R. Deckert, V. Grewe, and M. Ponater (2006), Solar cycle effect delays onset of ozone recovery, *Geophys. Res. Lett.*, *33*, L03806, doi:10.1029/2005GL024741.

de Grandpré, J., S. R. Beagley, V. I. Fomichev, E. Griffioen, J. C. McConnell, A. S. Medvedev, and T. G. Shepherd (2000), Ozone climatology using interactive chemistry: Results from the Canadian Middle Atmosphere Model, *J. Geophys. Res.*, *105*, 26,475–26,492.

Douglass, A. R., M. R. Schoeberl, R. S. Stolarski, J. W. Waters, J. M. Russell III, and A. E. Roche (1995), Interhemispheric differences in springtime production of HCl and $ClONO_2$ in the polar vortices, *J. Geophys. Res.*, *100*, 13,967–13,978.

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*, 27,545–27,564.

Egorova, T., E. Rozanov, V. Zubov, E. Manzini, W. Schmutz, and T. Peter (2005), Chemistry-climate model SOCOL: A validation of the present-day climatology, *Atmos. Chem. Phys.*, *5*, 1557–1576.

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

Eyring, V., D. E. Kinnison, and T. G. Shepherd (2005b), Overview of planned coupled chemistry-climate simulations to support upcoming ozone and climate assessments, *SPARC Newsl.*, *25*, 11–17.

Fioletov, V. E., G. E. Bodeker, A. J. Miller, R. D. McPeters, and R. Stolarski (2002), Global ozone and zonal total ozone variations estimated from ground-based and satellite measurements: 1964–2000, *J. Geophys. Res.*, *107*(D22), 4647, doi:10.1029/2001JD001350.

Fomichev, V. I., C. Fu, J. de Grandpré, S. R. Beagley, V. P. Ogibalov, and J. C. McConnell (2004), Model thermal response to minor radiative energy sources and sinks in the middle atmosphere, *J. Geophys. Res.*, *109*, D19107, doi:10.1029/2004JD004892.

Free, M., D. J. Seidel, J. K. Angell, J. Lanzante, I. Durre, and T. C. Peterson (2005), Radiosonde Atmospheric Temperature Products for Assessing Climate (RATPAC): A new data set of large-area anomaly time series, *J. Geophys. Res.*, *110*, D22101, doi:10.1029/2005JD006169.

Froidevaux, L., N. J. Livesey, W. G. Read, R. J. Salawitch, J. W. Waters, B. Drouin, I. A. MacKenzie, H. C. Pumphrey, P. Bernath, C. Boone, R. Nassar, S. Montzka, J. Elkins, D. Cunnold, and D. W. Waugh (2006), Temporal decrease in upper atmospheric chlorine, *Geophys. Res. Lett.*, doi:10.1029/2006GL027600, in press.

Garcia, R. R., and B. A. Boville (1994), "Downward control" of the mean meridional circulation and temperature distribution of the polar winter stratosphere, *J. Atmos. Sci.*, *51*, 2238–2245.

Garcia, R. R., and S. Solomon (1985), The effect of breaking gravity waves on the dynamics and chemical composition of the mesosphere and lower thermosphere, *J. Geophys. Res.*, *90*, 3850–3868.

Gates, W. L., et al. (1999), An overview of the results of the Atmospheric Model Intercomparison Projects (AMIP), *Bull. Am. Meteorol. Soc.*, *80*, 29–56.

Gelman, M. E., A. J. Miller, K. W. Johnson, and R. Nagatani (1996), Detection of long-term trends in global stratospheric temperature from NMC analyses derived from NOAA satellite data, *Adv. Space Res.*, *6*(10), 17–26.

Giorgetta, M. A., and L. Bengtsson (1999), The potential role of the quasi-biennial oscillation in the stratosphere-troposphere exchange as found in water vapor in general circulation model experiments, *J. Geophys. Res.*, *104*, 6003–6020.

Gleason, J. F., et al. (1993), Record low global ozone in 1992, *Science*, *260*, 523–526.

Gregory, A. R., and V. West (2002), The sensitivity of a model's stratospheric tape recorder to the choice of advection schemes, *Q. J. R. Meteorol. Soc.*, *128*, 1827–1846.

Gregory, D. G., G. J. Shutts, and J. R. Mitchell (1998), A new gravity wave drag scheme incorporating anisotropic orography and low-level wave breaking: Impact upon the climate of the UK Meteorological Office unified model, *Q. J. R. Meteorol. Soc.*, *124*, 463–494.

Grewe, V. (2006), The origin of ozone, *Atmos. Chem. Phys.*, *6*, 1495–1511.

Grooß, J.-U., and J. M. Russell III (2005), Technical note: A stratospheric climatology for O_3, H_2O, CH_4, NO_x, HCl and HF derived from HALOE measurements, *Atmos. Chem. Phys.*, *5*, 2797–2807.

Hall, T. M., and D. W. Waugh (1997), Tracer transport in the tropical stratosphere due to vertical diffusion and horizontal mixing, *Geophys. Res. Lett.*, *24*, 1383–1386.

Hall, T. M., D. W. Waugh, K. A. Boering, and R. A. Plumb (1999), Evaluation of transport in stratospheric models, *J. Geophys. Res.*, *104*, 18,815–18,840.

Hansen, J., et al. (2002), Climate forcings in Goddard Institute for Space Studies SI2000 simulations, *J. Geophys. Res.*, *107*(D18), 4347, doi:10.1029/2001JD001143.

Harries, J. E., et al. (1996), Validation of measurements of water vapor from the Halogen Occultation Experiment, HALOE, *J. Geophys. Res.*, *101*, 10,205–10,216.

Hines, C. O. (1997), Doppler-spread parameterization of gravity-wave momentum deposition in the middle atmosphere. Part 2: Broad and quasi monochromatic spectra, and implementation, *J. Atmos. Sol. Terr. Phys.*, *59*, 387–400.

Holton, J. R. (1982), The role of gravity wave induced drag and diffusion in the momentum budget of the mesosphere, *J. Atmos. Sci.*, *39*, 791–799.

Hourdin, F., and A. Armengaud (1999), The use of finite-volume methods for atmospheric advection trace species: 1. Tests of various formulations in a general circulation model, *Mon. Weather Rev.*, *127*, 822–837.

Intergovernmental Panel on Climate Change (2001), *Climate Change 2001: The scientific basis. Contribution of Working Group 1 to the Third Assessment Report*, edited by J. T. Houghton et al., Cambridge Univ. Press, New York.

Iwasaki, T., S. Yamada, and K. Tada (1989), A parameterization scheme of orographic gravity wave drag with the different vertical partitioning, part 1: Impact on medium range forecasts, *J. Meteorol. Soc. Jpn.*, *67*, 11–41.

Jackman, C. H., E. L. Fleming, S. Chandra, D. B. Considine, and J. E. Rosenfield (1996), Past, present, and future modeled ozone trends with comparisons to observed trends, *J. Geophys. Res.*, *101*, 28,753–28,768.

Kalnay, E., et al. (1996), The NCEP/NCAR 40-year reanalysis project, *Bull. Am. Meteorol. Soc.*, *77*, 437–471.

Kiehl, J. T., J. J. Hack, G. B. Bonan, B. A. Boville, D. L. Williamson, and P. J. Rasch (1998), The National Center for Atmospheric Research Community Climate Model: CCM3, *J. Clim.*, *11*, 1131–1150.

Kirchner, I., G. L. Stenchikov, H.-F. Graf, A. Robock, and J. C. Antuna (1999), Climate model simulation of winter warming and summer cooling following the 1991 Mount Pinatubo volcanic eruption, *J. Geophys. Res.*, *104*, 19,039–19,055.

Kley, D., E. J. Stone, W. R. Henderson, J. W. Drummond, W. J. Harrop, A. L. Schmeltekopf, T. L. Thompson, and R. H. Winkler (1979), In situ measurements of the mixing ratio of water vapor in the stratosphere, *J. Atmos. Sci.*, *36*, 2513–2524.

Kley, D., , J. M. Russell III, and C. Phillips (Eds.) (2000), SPARC assessment of upper tropospheric and stratospheric water vapour, *SPARC Rep. 2*, *WMO/TD 1043, WCRP Rep. 113*, World Clim. Res. Program, Geneva, Switzerland. (Available at http://www.aero.jussieu.fr/~sparc/WAVASF-INAL_000206/WWW_wavas/Cover.html).

Kurokawa, J., H. Akiyoshi, T. Nagashima, H. Masunaga, T. Nakajima, M. Takahashi, and H. Nakane (2005), Effects of atmospheric sphericity on stratospheric chemistry and dynamics over Antarctica, *J. Geophys. Res.*, *110*, D21305, doi:10.1029/2005JD005798.

Lamago, D., M. Dameris, C. Schnadt, V. Eyring, and C. Brühl (2003), Impact of large solar zenith angles on lower stratospheric dynamical and chemical processes in a coupled chemistry-climate model, *Atmos. Chem. Phys.*, *3*, 1981–1990.

Lean, J. L. (2000), Evolution of the Sun's spectral irradiance since the maunder minimum, *Geophys. Res. Lett.*, *27*, 2425–2428.

Lean, J. L., G. J. Rottman, H. L. Kyle, T. N. Woods, J. R. Hickey, and L. C. Puga (1997), Detection and parameterization of variations in solar mid- and near-ultraviolet radiation (200 to 400 nm), *J. Geophys. Res.*, *102*, 29,939–29,956.

Lefevre, F., G. P. Brasseur, I. Folkins, A. K. Smith, and P. Simon (1994), Chemistry of the 1991–1992 stratospheric winter: Three dimensional model simulations, *J. Geophys. Res.*, *99*, 8183–8195.

Lin, S.-J. (2004), A 'vertically Lagrangian' finite volume dynamical core for global models, *Mon. Weather Rev.*, *132*, 2293–2307.

Lindzen, R. S. (1981), Turbulence and stress owing to gravity wave and tidal breakdown, *J. Geophys. Res.*, *86*, 9707–9714.

Lott, F., and M. Miller (1997), A new subgrid scale orographic drag parameterization; its testing in the ECMWF model, *Q. J. R. Meteorol. Soc.*, *123*, 101–127.

Lott, F., L. Fairhead, F. Hourdin, and P. Levan (2005), The stratospheric version of LMDz: Dynamical climatologies, arctic oscillation, and impact on the surface climate, *Clim. Dyn.*, *25*, doi:10.1007/s00382-005-0064.

Manzini, E., N. A. McFarlane, and C. McLandress (1997), Impact of the Doppler spread parameterization on the simulation of the middle atmosphere circulation using the MA/ECHAM4 general circulation model, *J. Geophys. Res.*, *102*, 25,751–25,762.

Manzini, E., B. Steil, C. Brühl, M. A. Giorgetta, and K. Krüger (2003), A new interactive chemistry-climate model: 2. Sensitivity of the middle atmosphere to ozone depletion and increase in greenhouse gases and implications for recent stratospheric cooling, *J. Geophys. Res.*, *108*(D14), 4429, doi:10.1029/2002JD002977.

McFarlane, N. A. (1987), The effect of orographically excited gravity wave drag on the general circulation of the lower stratosphere and troposphere, *J. Atmos. Sci.*, *44*, 1775–1800.

Miller, A. J., et al. (2002), A cohesive total ozone data set from the SBUV(/2) satellite system, *J. Geophys. Res.*, *107*(D23), 4701, doi:10.1029/2001JD000853.

Miller, M. J., T. N. Palmer, and R. Swinbank (1989), Parameterization and influence subgrid-scale orography in general circulation and numerical weather prediction models, *Meteorol. Atmos. Phys.*, *40*, 84–109.

Mote, P. W., et al. (1996), An atmospheric tape recorder: The imprint of tropical tropopause temperatures on stratospheric water vapor, *J. Geophys. Res.*, *101*(D2), 3989–4006.

Mote, P. W., T. J. Dunkerton, M. E. McIntyre, E. A. Ray, P. H. Haynes, and J. M. Russell III (1998), Vertical velocity, vertical diffusion, and dilution by midlatitude air in the tropical lower stratosphere, *J. Geophys. Res.*, *103*, 8651–8666.

NASA Panel for Data Evaluation (2006), Chemical kinetics and photochemical data for use in atmospheric studies, Evaluation 15, *JPL Publ. 06–2*, NASA Jet Propul. Lab., Pasadena, Calif.

Newman, P. A., E. R. Nash, and J. E. Rosenfield (2001), What controls the temperature of the Arctic stratosphere during spring?, *J. Geophys. Res.*, *106*, 19,999–20,010.

Numaguti, A. (1993), Dynamics and energy balance of the Hadley circulation and the tropical precipitation zones: Significance of the distribution of evaporation, *J. Atmos. Sci.*, *50*, 1874–1887.

Park, J. H., et al. (1996), Validation of Halogen Occultation Experiment CH_4 measurements from the UARS, *J. Geophys. Res.*, *101*(D6), 10,183–10,204.

Pawson, S., et al. (2000), The GCM-Reality Intercomparison Project for SPARC: Scientific issues and initial results, *Bull. Am. Meteorol. Soc.*, *81*, 781–796.

Pitari, G. (1993), A numerical study of the possible perturbation of stratospheric dynamics due to Pinatubo aerosols: Implications for tracer transport, *J. Atmos. Sci.*, *50*, 2443–2461.

Pitari, G., S. Palermi, G. Visconti, and R. G. Prinn (1992), Ozone response to a CO_2 doubling: Results from a stratospheric circulation model with heterogeneous chemistry, *J. Geophys. Res.*, *97*, 5953–5962.

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

Press, W. H., B. R. Flannery, S. A. Teukolsky, and W. T. Vettering (1989), *Numerical Recipes in Pascal*, 759 pp., Cambridge Univ. Press, New York.

Pope, V., M. Gallani, P. Rowntree, and R. Stratton (2000), The impact of new physical parametrizations in the Hadley Centre climate model–HadCM3, *Clim. Dyn.*, *16*, 123–146.

Ramaswamy, V., et al. (2001), Stratospheric temperature trends: Observations and model simulations, *Rev. Geophys.*, *39*, 71–122.

Randel, W. J., and F. Wu (2006), Biases in stratospheric and tropospheric temperature trends derived from historical radiosonde data, *J. Clim.*, *19*, 2094–2104.

Randel, W. J., F. Wu, S. J. Oltmans, K. Rosenlof, and G. E. Nedoluha (2004), Interannual changes of stratospheric water vapor and correlations with tropical tropopause temperatures, *J. Atmos. Sci.*, *61*, 2133–2148.

Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan (2003), Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, *J. Geophys. Res.*, *108*(D14), 4407, doi:10.1029/2002JD002670.

Roeckner, E., et al. (1996), The atmospheric general circulation model ECHAM-4: Model description and simulation of present-day climate, *Rep. 218*, Max-Planck-Inst. für Meteorol., Hamburg, Germany.

Rozanov, E., M. Schraner, C. Schnadt, T. Egorova, M. Wild, A. Ohmura, V. Zubov, W. Schmutz, and T. Peter (2005), Assessment of the ozone and temperature variability during 1979–1993 with the chemistry-climate model SOCOL, *Adv. Space. Res.*, *35*(8), 1375–1384.

Russell, J. M., III, et al. (1993), The Halogen Occultation Experiment, *J. Geophys. Res.*, *98*(D6), 10,777–10,797.

Russell, J. M., III, et al. (1996), Validation of hydrogen chloride measurements made by HALOE from the UARS platform, *J. Geophys. Res.*, *101*, 10,151–10,162.

Sage, K. H., J. H. Park, M. K. W. Ko, C. H. Jackman, R. A. Plumb, and J. A. Kaye (1999), Models and measurements intercomparison II, *NASA/TM-1999–209554*, NASA Langley Res. Cent., Hampton, Va. (Available at http://hdl.handle.net/2002/12300).

Sankey, D., and T. G. Shepherd (2003), Correlations of long-lived chemical species in a middle atmosphere general circulation model, *J. Geophys. Res.*, *108*(D16), 4494, doi:10.1029/2002JD002799.

Santee, M. L., et al. (1996), Chlorine deactivation in the lower stratospheric polar regions during late winter: Results from UARS, *J. Geophys. Res.*, *101*, 18,835–18,860.

Santer, B. D., J. E. Penner, and P. W. Thorne (2006), How well can the observed vertical temperature changes be reconciled with our understanding of the causes of these changes?, in *Temperature Trends in the Lower Atmosphere: Steps for Understanding and Reconciling Differences. A Report by the U. S. Climate Change Science Program and the Subcommittee on Global Change Research*, edited by T. R. Karl et al., pp. 89–118, Natl. Oceanic and Atmos. Admin., Natl. Climatic Data Cent., Asheville, N. C.

Sato, M., J. E. Hansen, M. P. McCormick, and J. B. Pollack (1993), Stratospheric aerosol optical depths, 1850–1990, *J. Geophys. Res.*, *98*, 22,987–22,994.

Scaife, A. A., N. Butchart, C. D. Warner, and R. Swinbank (2002), Impact of a spectral gravity wave parameterization on the stratosphere in the Met Office Unified Model, *J. Atmos. Sci.*, *59*, 1473–1489.

Schmidt, G. A., et al. (2006), Present-day atmospheric simulations using GISS Model E: Comparison to in situ, satellite, and reanalysis data, *J. Atmos. Sci.*, *19*, 153–192.

Scinocca, J. F. (2003), An accurate spectral non-orographic gravity wave parameterization for general circulation models, *J. Atmos. Sci.*, *60*, 667–682.

Scinocca, J. F., and N. A. McFarlane (2000), The parameterization of drag induced by stratified flow over anisotropic orography, *Q. J. R. Meteorol. Soc.*, *126*, 2353–2393.

Scinocca, J. F., and N. A. McFarlane (2004), The variability of modeled tropical precipitation, *J. Atmos. Sci.*, *61*, 1993–2015.
Sessler, J., P. Good, A. R. MacKenzie, and J. A. Pyle (1996), What role do Type I polar stratospheric cloud and aerosol parameterizations play in modeled lower stratospheric chlorine activation and ozone loss?, *J. Geophys. Res.*, *101*, 28,817–28,835.
Shibata, K., and M. Deushi (2005), Partitioning between resolved wave forcing and unresolved gravity wave forcing to the quasi-biennial oscillation as revealed with a coupled chemistry-climate model, *Geophys. Res. Lett.*, *32*, L12820, doi:10.1029/2005GL022885.
Shibata, K., H. Yoshimura, M. Ohizumi, M. Hosaka, and M. Sugi (1999), A simulation of troposphere, stratosphere and mesosphere with an MRI/JMA98 GCM, *Pap. Meteorol. Geophys.*, *50*, 15–53.
Shibata, K., M. Deushi, T. T. Sekiyama, and H. Yoshimura (2005), Development of an MRI chemical transport model for the study of stratospheric chemistry, *Pap. Meteorol. Geophys.*, *55*, 75–119.
Shine, K. P., et al. (2003), A comparison of model-simulated trends in stratospheric temperatures, *Q. J. R. Meteorol. Soc.*, *129*, 1565–1588.
Smith, A. K., and L. V. Lyjak (1985), An observational estimate of gravity wave drag from the momentum balance in the middle atmosphere, *J. Geophys. Res.*, *90*, 2233–2241.
Solomon, S., R. R. Garcia, and F. Stordal (1985), Transport processes and ozone perturbations, *J. Geophys. Res.*, *90*, 12,981–12,989.
Steil, B., M. Dameris, C. Brühl, P. J. Crutzen, V. Grewe, M. Ponater, and R. Sausen (1998), Development of a chemistry module for GCMs: First results of a multi-annual integration, *Ann. Geophys.*, *16*, 205–228.
Steil, B., C. Brühl, E. Manzini, P. J. Crutzen, J. Lelieveld, P. J. Rasch, E. Roeckner, and K. Krüger (2003), A new interactive chemistry-climate model: 1. Present-day climatology and interannual variability of the middle atmosphere using the model and 9 years of HALOE/UARS data, *J. Geophys. Res.*, *108*(D9), 4290, doi:10.1029/2002JD002971.
Steinbrecht, W., et al. (2006a), Interannual variation patterns of total ozone and temperature in observations and model simulations, *Atmos. Chem. Phys.*, *6*, 349–374.
Steinbrecht, W., et al. (2006b), Long-term evolution of upper stratospheric ozone at selected stations of the Network for the Detection of Stratospheric Change (NDSC), *J. Geophys. Res.*, *111*, D10308, doi:10.1029/2005JD006454.
Stern, W. F., and R. T. Pierrehumbert (1988), The impact of an orographic gravity wave drag parameterization on extended range predictions with a GCM, paper presented at Eighth Conference on Numerical Weather Prediction, Am. Meteorol. Soc., Baltimore, Md.
Stolarski, R. S., and S. 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.
Stolarski, R. S., A. R. Douglass, S. Steenrod, and S. Pawson (2006), Trends in stratospheric ozone: Lessons learned from a 3d chemical transport model, *J. Atmos. Sci.*, *63*, 1028–1041.
Struthers, H., K. Kreher, J. Austin, R. Schofield, G. E. Bodeker, P. V. Johnston, H. Shiona, and A. Thomas (2004), Past and future simulations of NO_2 from a coupled chemistry-climate model in comparison with observations, *Atmos. Chem. Phys.*, *4*, 2227–2239.
Swinbank, R., and A. O' Neill (1994), A stratosphere-troposphere data assimilation system, *Mon. Weather Rev.*, *122*, 686–702.
Taylor, K. E., D. Williamson, and F. Zwiers (2000), The sea surface temperature and sea-ice concentration boundary concentrations for AMIP II simulations, *PCMDI Rep. 60*, Prog. for Clim. Model Diagn. and Intercomparison, Lawrence Livermore Natl. Lab., Livermore, Calif.
Tegtmeier, S., and T. G. Shepherd (2006), Persistence and photochemical decay of springtime total ozone anomalies in the Canadian Middle Atmosphere Model, *Atmos. Chem. Phys. Disc.*, *6*, 3403–3417.
Thomason, L. W., and T. Peter (Eds.) (2006), Assessment of Stratospheric Aerosol Properties (ASAP), *SPARC Rep. 4, WMO/TD 1295, Tech Rep. WCRP-124*, World Clim. Res. Program, Geneva, Switzerland. (Available at http://www.atmosp.physics.utoronto.ca/SPARC/index.html#)
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*, 281–304.
Uppala, S. M., et al. (2005), The ERA-40 re-analysis, *Q. J. R. Meteorol. Soc.*, *131*, 2961–3012, doi:10.1256/qj.04.176.
Warner, C. D., and M. E. McIntyre (1996), On the propagation and dissipation of gravity wave spectra through a realistic middle atmosphere, *J. Atmos. Sci.*, *53*, 3213–3235.
Waugh, D., and T. Hall (2002), Age of stratospheric air: Theory, observations, and models, *Rev. Geophys.*, *40*(4), 1010, doi:10.1029/2000RG000101.
Williamson, D. L., and P. J. Rasch (1994), Water vapor transport in the NCAR CCM2, *Tellus, Ser. A*, *46*, 34–51.
World Meteorological Organization/United Nations Environment Programme (2003), Scientific assessment of ozone depletion: 2002, *Rep. 47*, Global Ozone Res. and Monit. Proj., World Meteorol. Org., Geneva, Switzerland.
Zubov, V., E. Rozanov, and M. Schlesinger (1999), Hybrid scheme for three-dimensional advective transport, *Mon. Weather Rev.*, *127*, 1335–1346.

H. Akiyoshi, T. Nagashima, and M. Yoshiki, National Institute for Environmental Studies, Tsukuba 305-8506, Japan.
J. Austin, Geophysical Fluid Dynamics Laboratory, NOAA, Princeton, NJ 08542-0308, USA.
S. Bekki, L. Jourdain, and M. Marchand, Service d'Aéronomie du Centre National de la Recherche Scientifique, F-75252 Paris, France.
G. E. Bodeker and H. Struthers, National Institute of Water and Atmospheric Research, Lauder, New Zealand.
N. Butchart, Climate Research Division, Met Office, Exeter EX1 3PB, UK.
B. A. Boville, R. R. Garcia, A. Gettelman, D. E. Kinnison, and D. R. Marsh, National Center for Atmospheric Research, Boulder, CO 80307, USA.
C. Brühl, Max Planck Institut für Chemie, D-55128 Mainz, Germany.
M. P. Chipperfield and W. Tian, Institute for Atmospheric Science, University of Leeds, Leeds LS2 9JT, UK.
E. Cordero, Department of Meteorology, San Jose State University, San Jose, CA 95192, USA.
M. Dameris, V. Eyring, and V. Grewe, Institut für Physik der Atmosphäre, Deutsches Zentrum für Luft- und Raumfahrt, Oberpfaffenhofen, D-82234 Wessling, Germany. (veronika.eyring@dlr.de)
M. Deushi and K. Shibata, Meteorological Research Institute, Tsukuba 305-0052, Japan.
V. E. Fioletov, Environment Canada, 4905 Dufferin Street, Toronto, ON, Canada M3H 5T4.
S. M. Frith and J. E. Nielsen, Science Systems and Applications, Inc., Lanham, MD 20706, USA.
M. A. Giorgetta, Max Planck Institut für Meteorologie, D-20146 Hamburg, Germany.
E. Mancini and G. Pitari, Dipartimento di Fisica, Università L'Aquila, Via Vetoio, I-67010 L'Aquila, Italy.
E. Manzini, Istituto Nazionale di Geofisica e Vulcanologia, I-40128 Bologna, Italy.
P. A. Newman, S. Pawson, and R. S. Stolarski, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA.
D. A. Plummer, Ouranos Consortium, 550 rue Sherbrooke Ouest, 19e étage, Montréal, QC, Canada H3A 1B9.
E. Rozanov, Physical-Meteorological Observatory/World Radiation Center, CH-7260 Davos, Switzerland.
M. Schraner, Institute for Atmospheric and Climate Science, Eidgenössische Technische Hochschule, CH-8092 Zurich, Switzerland.
T. G. Shepherd, Department of Physics, University of Toronto, Toronto, ON, Canada M5S 1A7.
D. W. Waugh, Department of Earth and Planetary Sciences, Johns Hopkins University, Baltimore, MD 21218, USA.