



Interactive ozone and methane chemistry in GISS-E2 historical and future climate simulations

D. T. Shindell¹, O. Pechony¹, A. Voulgarakis^{1,*}, G. Faluvegi¹, L. Nazarenko¹, J.-F. Lamarque², K. Bowman³, G. Milly¹, B. Kovari¹, R. Ruedy¹, and G. A. Schmidt¹

¹NASA Goddard Institute for Space Studies and Columbia Earth Institute, New York, NY USA

²National Center for Atmospheric Research, Boulder, CO USA

³Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA USA

*now at: Department of Physics, Imperial College London, London, UK

Correspondence to: D. T. Shindell (drew.t.shindell@nasa.gov)

Received: 27 July 2012 – Published in Atmos. Chem. Phys. Discuss.: 11 September 2012

Revised: 12 February 2013 – Accepted: 13 February 2013 – Published: 5 March 2013

Abstract. The new generation GISS climate model includes fully interactive chemistry related to ozone in historical and future simulations, and interactive methane in future simulations. Evaluation of ozone, its tropospheric precursors, and methane shows that the model captures much of the large-scale spatial structure seen in recent observations. While the model is much improved compared with the previous chemistry-climate model, especially for ozone seasonality in the stratosphere, there is still slightly too rapid stratospheric circulation, too little stratosphere-to-troposphere ozone flux in the Southern Hemisphere and an Antarctic ozone hole that is too large and persists too long. Quantitative metrics of spatial and temporal correlations with satellite datasets as well as spatial autocorrelation to examine transport and mixing are presented to document improvements in model skill and provide a benchmark for future evaluations. The difference in radiative forcing (RF) calculated using modeled tropospheric ozone versus tropospheric ozone observed by TES is only 0.016 W m^{-2} . Historical 20th Century simulations show a steady increase in whole atmosphere ozone RF through 1970 after which there is a decrease through 2000 due to stratospheric ozone depletion. Ozone forcing increases throughout the 21st century under RCP8.5 owing to a projected recovery of stratospheric ozone depletion and increases in methane, but decreases under RCP4.5 and 2.6 due to reductions in emissions of other ozone precursors. RF from methane is 0.05 to 0.18 W m^{-2} higher in our model calculations than in the RCP RF estimates. The surface temperature response to ozone through 1970 follows the increase in forcing due to

tropospheric ozone. After that time, surface temperatures decrease as ozone RF declines due to stratospheric depletion. The stratospheric ozone depletion also induces substantial changes in surface winds and the Southern Ocean circulation, which may play a role in a slightly stronger response per unit forcing during later decades. Tropical precipitation shifts south during boreal summer from 1850 to 1970, but then shifts northward from 1970 to 2000, following upper tropospheric temperature gradients more strongly than those at the surface.

1 Introduction

The new E2 version of the GISS climate model was run with fully interactive reactive chemical constituents in simulations for the Coupled Model Intercomparison Project 5th phase (CMIP5; Taylor et al., 2012) and the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP; Lamarque et al., 2013). Here we describe the model and then evaluate both present-day simulations and ozone trends over recent decades against observations. We then describe the radiative forcing due to ozone for historical simulations and due to ozone and methane for future simulations. Finally, we evaluate the climate response to transient historical ozone changes in an ensemble of coupled ocean-atmosphere climate model simulations.

2 Model description

Here we document the behavior of chemically reactive compounds in the version of the model used in the historical and future simulations in support of CMIP5/ACCMIP. A slightly earlier version of the model, with minor differences in both the climate model and the chemistry model but driven by the same emissions, was used to generate ozone fields to provide to the CMIP5 activity for use by models without chemistry (Lamarque et al., 2010). The average of the historical tropospheric ozone simulated in that version of the model and similar simulations performed by the National Center for Atmospheric Research was described in Cionni et al. (2011).

2.1 Climate model

The new version of the GISS climate model used for CMIP5/ACCMIP simulations is called ModelE2. It is similar to modelE (Schmidt et al., 2006), but has numerous improvements to the physics, including the ability to represent multiple downdrafts and updrafts in convective systems. As transient climate simulations with online, interactive chemistry were performed for both the historical 1850–2005 period and for future scenarios, these simulations were used for both CMIP5 and ACCMIP. All runs analyzed here were performed at 2×2.5 degrees horizontal resolution, with 40 vertical layers extending to 0.1 hPa. The atmosphere was coupled to a full dynamic ocean, either the Russell ocean (Russell et al., 1995) (GISS-E2-R; hereafter E2-R) or the Hybrid Coordinate Ocean Model (HYCOM) ocean (Sun and Bleck, 2006) (GISS-E2-H; hereafter E2-H). Results from the E2-H model are generally very similar to the E2-R model, with differences associated with variations in the base sea-surface temperatures and sea-ice. We arbitrarily focus on the latter in most cases for simplicity, highlighting E2-H only when results differ. To better understand the role of biases in the climate model's internally generated meteorology, we also performed a 2000–2010 simulation in which the meteorology was nudged towards the MERRA reanalysis (Rienecker et al., 2011) and ocean sea-surface temperatures and sea ice cover were prescribed according to observations (Rayner et al., 2003). We refer to that simulation as E2-MERRA. Two versions of the interactive code are archived at CMIP5, denoted by the physics-version = 2 or 3. Physics-version = 2 had a tuned indirect aerosol effect following Hansen et al. (2005), while physics-version = 3 has a parameterized aerosol cloud-albedo effect following Menon et al. (2010). The simulations discussed here are from version 3 unless stated otherwise.

2.2 Chemistry

The chemical mechanism is fully embedded within modelE, so that chemical constituents are treated consistently with the physics of other parts of the model such as surface fluxes of

fundamental physical quantities (e.g. heat) and with transport of momentum and other constituents such as water vapor. Tropospheric chemistry includes basic NO_x - HO_x - O_x - CO - CH_4 chemistry as well as PANs and the hydrocarbons isoprene, alkyl nitrates, aldehydes, alkenes, and paraffins. The lumped hydrocarbon family scheme was derived from the Carbon Bond Mechanism-4 (CBM-4) (Gery et al., 1989) and from the more extensive Regional Atmospheric Chemistry Model (RACM) (Stockwell et al., 1997), following (Houweling et al., 1998). To represent stratospheric chemistry, the model includes chlorine- and bromine-containing compounds, and CFC and N_2O source gases (as well as an “age-of-air” passive tracer). As we use only a single CFC tracer, the ratio of anthropogenic bromine to chlorine is held fixed at year 2000 values, with both released from CFC photolysis in an amount proportional to the total equivalent effective stratospheric chlorine loading in a given year. Aerosol species and chemistry are also included, and chemistry and aerosols are fully integrated, so that these components interact with each other and with the physics of the climate model. The model contains 51 species for gas-phase chemistry interacting via 156 reactions. We use a chemical time step of 1/2 h, including the calculation of photolysis rates using the Fast-J2 scheme (Bian and Prather, 2002). We include transport and phase transformations of soluble species within convective plumes, scavenging within and below updrafts, rain-out within both convective and large-scale clouds, washout below precipitating regions, evaporation of falling precipitation, and both detrainment and evaporation from convective plumes. A full description is given in Shindell et al. (2006b) and references therein.

The only substantial chemistry changes since (Shindell et al., 2006b) are as follows: acetone chemistry has been added to the model (Houweling et al., 1998), we have incorporated a newly identified reaction channel of $\text{OH} + \text{NO}_2$ to form HNO_3 whose branching ratio is pressure and temperature dependent (Butkovskaya et al., 2007), polar stratospheric cloud (PSC) formation in the stratosphere is now dependent upon water vapor, temperature and HNO_3 (Hanson and Mauersberger, 1988), the heterogeneous hydrolysis of N_2O_5 on sulfate now follows Kane et al. (2001) and Hallquist et al. (2003), and the model now includes terpene emissions and oxidation by OH , O_3 and NO_3 (Tsigaridis et al., 2005).

Another development is that now light attenuation by modeled aerosol tracers affects the photolysis rates of gases, following Bian et al. (2003). Modeled aerosol optical depths are passed to the photolysis code at every timestep, while the tabulated optical properties required for different aerosol types (extinction efficiency, single scattering albedo, scattering phase function expansion terms) are prescribed according to prior calculations to be consistent with what is used in the model's radiation code.

2.3 Anthropogenic emissions

Historical emissions are the harmonized set of anthropogenic and biomass burning emissions (NO_x , CO, volatile organic compounds (VOCs) and aerosols) prepared in support of CMIP5 (Lamarque et al., 2010). Future emission scenarios were created by four independent integrated assessment modeling (IAM) groups where the goal was to provide emissions that would lead to radiative forcing in 2100 relative to 1850 of a prescribed amount. The scenarios are denoted ‘representative concentration pathways’ (RCPs) and are named by their nominal 2100 forcing, namely RCP2.6, 4.5, 6.0 and 8.5 (van Vuuren et al., 2011). The four pathways were then harmonized to match historical emission estimates and each other at the year 2000, with harmonization across the 4 RCPs maintained through 2005. All emissions are input at decadal resolution, with intermediate values linearly interpolated.

The long-lived gases methane (CH_4), nitrous oxide (N_2O) and CFCs are not emitted directly in our model, but instead their concentrations are prescribed at the surface according to observations for the past and RCP projections for the future. The exception to this is that methane concentrations after 2005 are calculated internally driven directly by emissions and chemical and soil losses. All three of these long-lived gases are chemically active in the model, so that their vertical structure is determined by the model’s transport and chemistry, and they have important effects on ozone in both the troposphere and stratosphere. CFCs are a lumped family designed to represent the total anthropogenic halogen loading of the atmosphere.

The RCP projections for N_2O and CFC concentrations are based on calculations with the reduced-complexity coupled carbon cycle climate model MAGICC 6.3 to estimate mixing ratios that would result from the IAM RCP emissions (Meinshausen et al., 2011). For methane, we calculate concentrations based on the RCP methane emissions in order to assess the potential for climate driven changes in atmospheric sinks and natural emissions, and include those in radiative forcing, while all other groups’ currently available CMIP5 simulations used the MAGICC-derived concentrations.

2.4 Natural emissions

Natural emissions are generated internally within the GISS model for several sources, and hence vary with climate. For gaseous species in E2-R simulations for the present-day (2000s), these are NO_x from lightning (7.3 Tg N yr^{-1}) and biogenic isoprene (602 Tg yr^{-1}) (Shindell et al., 2006b). Natural emissions of NO_x from soils are prescribed at fixed values (2.7 Tg N yr^{-1}), as are biogenic emissions of alkenes (16 Tg C yr^{-1}), paraffins (14 Tg C yr^{-1}), and terpenes (192 Tg yr^{-1}).

Emissions from wetlands are a large source of methane, but are poorly constrained. Current methane concentrations and growth rates are well known, however. Hence we cali-

brated the year 2005 (the end of our historical runs) methane emissions from wetlands so that methane’s concentration and growth rate at that year would match the observations. Testing this, a nine year simulation with prescribed 2005 emissions and boundary conditions showed a growth rate of -8 to 8 Tg yr^{-1} , in reasonable agreement with recent observations showing a growth rate of ~ -20 to 40 Tg yr^{-1} during 2000–2009 (Dlugokencky et al., 2009) (corresponding to ~ -5 to 10 ppb yr^{-1} increase). We then used this calibrated methane wetland source in all future projections. Note that while this calibrated methane wetland source balances our model, it only matches the true emissions if other emissions and loss rates are correct.

Present day model wetland emissions are $197 \pm 4 \text{ Tg yr}^{-1}$ in the first 5 yr of the 4 transient simulations. During that time period, the net growth rate of methane (averaged over 2006–2010) ranges from 2 to 35 Tg yr^{-1} depending upon the RCP scenario. These correspond to growth rates of about 0.5 to 8 ppb yr^{-1} , similar to those of the last decade. The differences across the scenarios result from the differing chemical lifetimes obtained when the four different sets of precursor emissions are used (anthropogenic emissions of ozone precursors were linearly interpolated between 2000 and 2010 during 2006–2010; see Sect. 5.3 for further discussion). Methane is primarily oxidized by OH, which is quite sensitive to the ratio of NO_x to $\text{CO} + \text{VOC}$ emissions. As these differ in the four scenarios, the methane growth rate is different in the four simulations. While it would have been possible to separately adjust the methane emissions to create a near-zero growth rate in all four cases, it would be physically inconsistent to have different base case (emissions with 2005 climate) methane wetland emissions in each simulation. As our model results are qualitatively consistent across scenarios, we do not believe any small bias in initial growth rates would have a large impact on the simulations. Other natural methane sources are fixed at values of 20 Tg yr^{-1} for termites and 27 Tg yr^{-1} for combined geological + ocean + lake sources. There are no changes in methane emissions due to thawing permafrost or increased release from hydrates over time. The methane soil sink is also constant at 30 Tg yr^{-1} .

3 Present-day composition

We evaluate several aspects of present-day composition in the model, focusing on quantities that are important for radiative forcing and for which observations are available for evaluation. We first discuss ozone, followed by factors that affect the ozone distribution including stratospheric circulation and tropospheric ozone precursors.

3.1 Ozone

3.1.1 Evaluation of total column amounts

A wide variety of observations are available for evaluation of present-day ozone. We first compare the seasonality of zonal mean total column ozone with observations from the Total Ozone Mapping Spectrometer (TOMS) and the Ozone Monitoring Instrument (OMI) (McPeters et al., 2008). The E2-R model captures the zonal mean observed total column ozone in the tropics very well, with biases less than 5 % throughout the year (Fig. 1). At middle to high latitudes, the E2-R model shows cold-season positive biases of up to 17 %. As chemistry is slow at high latitudes during the dark winter months, this suggests that the equator-to-pole circulation is too fast in this model. In the Northern Hemisphere (NH), biases in the warm season are always less than 10 %. In the Southern Hemisphere (SH), negative biases are seen in the austral spring at middle to high latitudes, indicating that the Antarctic ozone hole extends slightly too far equatorward in September–October (as the model has a sizeable cold bias from ~ 50 – 65° S in the upper troposphere and lower stratosphere), and persists about 1 month too long in the polar region. The latter is consistent with a cold bias in the Antarctic lower stratospheric springtime temperatures in our model, as found in most other comparable models (Butchart et al., 2011). Other than those limitations, the onset and recovery of Antarctic ozone depletion is generally captured well in terms of magnitude, location and timing (maximum biases are less than 20 %). The total column ozone results are substantially improved relative to the earlier version of the composition-climate model, which had little seasonality at high latitudes and showed biases of up to 40 % for the Antarctic ozone hole and biases of 20–25 % for other middle to high latitudes (Shindell et al., 2006b). The E2-MERRA run has biases similar in magnitude to those in E2-R, but there is too little seasonality at high Northern latitudes in that run, while the Antarctic shows substantial positive biases in austral summer but the ozone hole duration is now in good agreement with observations.

While this analysis reveals a great deal about the seasonality of total column ozone, it tells us nothing about the longitudinal or vertical structure of ozone. We next examine the vertical structure, subsequently using that information to attempt to separate tropospheric and stratospheric ozone in order to better assess longitudinal variations.

3.1.2 Evaluation of vertical structure

As in prior evaluations of the GISS chemistry model, we have compared tropospheric ozone against a climatology of balloon sonde data where reliable, long-term climatologically representative measurements are available. While their geographical coverage is of course limited, the sondes provide extremely high quality measurements. This provides further

insight into the ozone distribution alongside the broad geographical coverage of the satellites. In addition to a 16 site climatology (Logan, 1999) used in prior evaluations, we have added 10 stations from the Southern Hemisphere Additional Ozone Sondes (SHADOZ) network to the analysis (Thompson et al., 2007). The former dataset is based on data taken primarily from the mid-1980s to the mid-1990s, while the SHADOZ data covers 1998–2009. We have therefore compared with E2-R simulations during both time periods. We also compare with measurements from the Tropospheric Emission Spectrometer (TES) instrument (Osterman et al., 2008). A monthly mean TES ozone climatology from 2005–2009 is sampled at the same locations and altitudes as the sonde data. These results can be compared with both the 2000s E2-R simulation and the E2-MERRA run, which was only performed for the 2000s as it was mainly intended for comparison with satellite observations.

Comparison of the model runs with sonde observations at several selected sites spanning the globe in latitude provides insight into the biases in the model (Fig. 2). At the Arctic Resolute station, the E2-R model has an underestimate during winter at 125 hPa, but otherwise captures measured ozone values quite well in the upper levels. The E2-MERRA run strongly underestimates ozone, especially during boreal winter. The underestimates are much larger than the E2-R model's difference between the 80s/90s and 2000s, indicating that the E2-MERRA biases are not a result of the time difference relative to the sondes. This bias is consistent with the attenuated seasonality in Arctic total column ozone in the E2-MERRA run in comparison with satellite measurements (Fig. 1). At 900 hPa, however, both models show strong seasonality that is not apparent in the sonde observations, causing large positive biases outside of boreal summer. The TES observations also show large seasonality at this level. As TES has fairly low sensitivity in the lower troposphere at high latitudes, retrievals there can be highly influenced by the a priori climatology. As this is smoothed data representative of large spatial scales, this seasonality could be a phenomenon that is present in the Arctic but not seen locally at Resolute, or a bias in the a priori as in the model/sonde comparison shown here. TES shows higher ozone than the sondes at 125 hPa, but this appears to be consistent with the modeled increase between the 80s/90s and 2000s at this location. It is also noteworthy that such differences between the 80s/90s and the 2000s exist only for this level (125 hPa).

The E2-MERRA run shows similar underestimates in the lower stratosphere at the mid-latitude Hohenpeissenberg station, while the E2-R model performs well. Both model versions show high biases in the lower troposphere, particularly at 900 hPa. These may be in part due to comparison of comparatively large area grid boxes with a mountain station. Similar to the models, TES shows much larger values at 900 hPa, supporting the areal-averaging hypothesis. Excessive transport of ozone from the stratosphere is likely not the cause, since if this were true, biases would be expected to increase

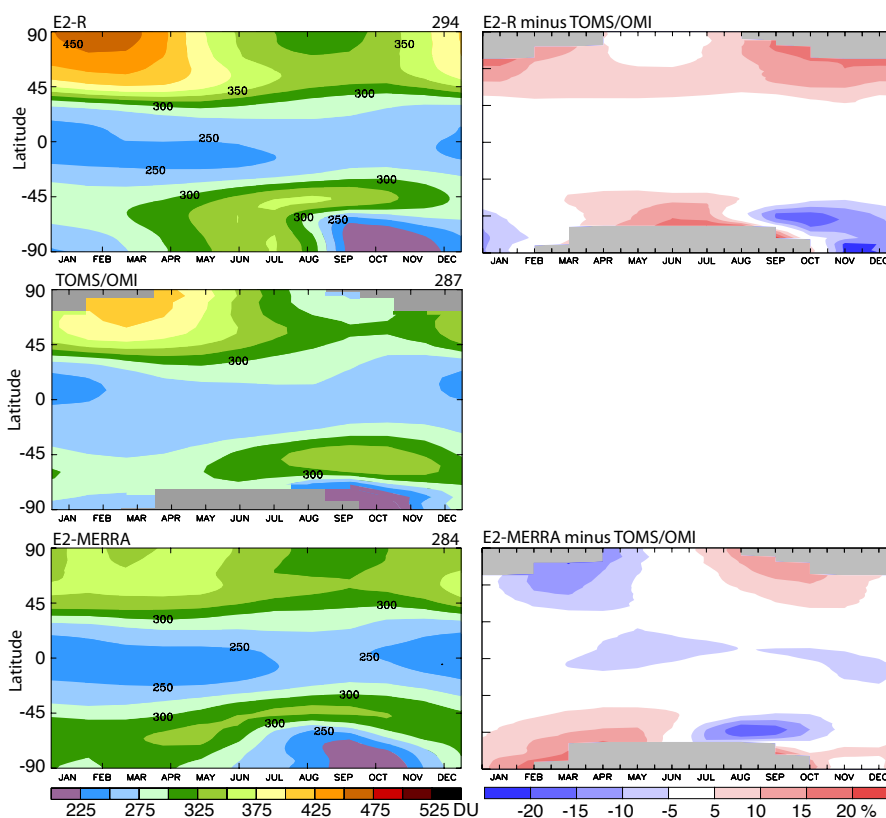


Fig. 1. Total ozone column (DU; left column) in the model simulations and in observations during 2000–2010 (TOMS from 2000–2004, OMI from 2004–2010), and model-observation differences (%; right column). Global mean column values are given in the upper right of each panel.

Table 1. Ozone differences and biases (ppbv) between models and sondes.

Pressure level (hPa)	Avg. difference modelE	Avg. difference E2-R 1985–1996	Avg. difference E2-R 2000s	Avg. difference E2-MERRA 2000s	Avg. bias modelE	Avg. bias E2-R	Standard deviation of observations
125	59.6 (15 %)	51.9 (13 %)	58.3 (15 %)	66.3 (17 %)	−16.0 (4 %)	30.3 (8 %)	92.9
200	37.2 (18 %)	37.3 (18 %)	38.7 (19 %)	26.1 (13 %)	−8.0 (4 %)	12.9 (6 %)	52.2
300	27.7 (37 %)	15.6 (21 %)	15.6 (21 %)	12.3 (16 %)	8.1 (11 %)	4.6 (6 %)	25.6
500	10.3 (22 %)	9.1 (19 %)	10.3 (22 %)	9.3 (20 %)	−4.0 (9 %)	4.7 (10 %)	11.7
900	8.9 (30 %)	9.2 (31 %)	9.6 (32 %)	8.9 (30 %)	−5.6 (19 %)	6.6 (22 %)	8.9

Comparisons are between the models and 26 recommended sites from Logan (1999) (16 sites) and SHADOZ data (Thompson et al., 2007) (10 sites). Average differences are from the month-by-month absolute value differences between the model and the sondes. Values in parentheses are percent difference with respect to observed values at these levels. ModelE is the 4×5 , 23 layer version, while E2-R is run at 2×2.5 degrees with 40 vertical layers. Sonde data from the Logan climatology is primarily from 1985–1996. Though the tropical SHADOZ sondes are more recent, there has been little trend in the tropics. As there have been large trends in the extratropics, the 1985–1996 model is most suitable for comparison with the combined sonde network. We show the E2-R model's 2000s output only to provide a comparison for the E2-MERRA run which was only performed for the later period. Model biases are from the 1985–1996 analysis.

with altitude. Simulations at the remote Hawaiian site of Hilo and at the tropical Natal station are generally fairly good throughout the tropospheric column in both model versions, consistent with the very small biases seen at low latitudes in total column ozone (Fig. 1). TES and sonde data generally also match well at these locations, consistent with the

80s/90s and 2000s output being nearly identical in the E2-R simulations.

In the mid-latitude Southern Hemisphere at Lauder and at the Antarctic Syowa station, the E2-R model overestimates 125 hPa ozone during austral winter, while the E2-MERRA version underestimates. Again, these results are consistent with the biases in total column ozone for these

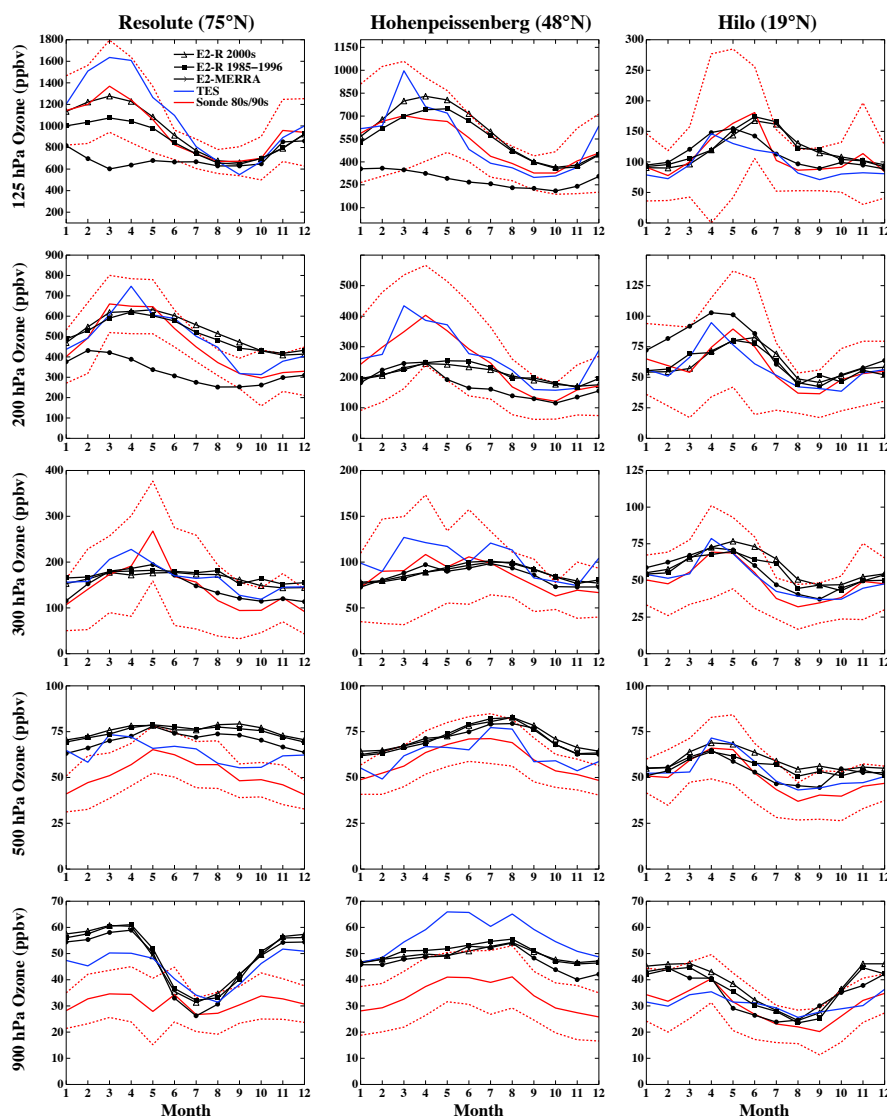


Fig. 2. Comparison of E2-R and E2-MERRA simulations with ozone sonde climatological station data (red, solid is mean and dashed is standard deviation) and with observations from the TES satellite instrument at the same locations. Sonde data from the first six stations shown is from \sim 1985–1996, while from the last three stations is from the 2000s. TES data is from 2005–2009.

regions (Fig. 1). Surprisingly, the situation is reversed at 200 hPa at Lauder. This suggests that the E2-R version has too little downward flux of ozone at this location. However, the E2-R model is fairly realistic at these locations for the mid-troposphere down to the surface.

Comparisons with the tropical SHADOZ stations are typically quite good. The models capture the boreal fall biomass burning season peak at La Reunion and the double peak seasonality at Ha Noi at low levels, for example (Fig. 2). They generally capture the distribution throughout the vertical column quite well, with the exception of an underestimate in the E2-R simulation at Fiji from 200–300 hPa (as at Natal), where the E2-MERRA run performs well. There is little dif-

ference at these locations between the 80s/90s and 2000s model output.

We performed statistical comparisons with the complete 26 station network. We focus on results using the E2-R model's 80s/90s output, as that is most appropriate for the non-SHADOZ sites and has little effect at the SHADOZ locations. We find that modelE2 is fairly similar to the prior modelE (Table 1). The average absolute value difference with respect to sondes has considerably improved at 300 hPa, where the differences were particularly large in modelE, but it is basically the same at other levels. The average bias in the model is also greatly reduced at 300 hPa, but is increased at 125 hPa. Notably, the overall bias is now positive at all

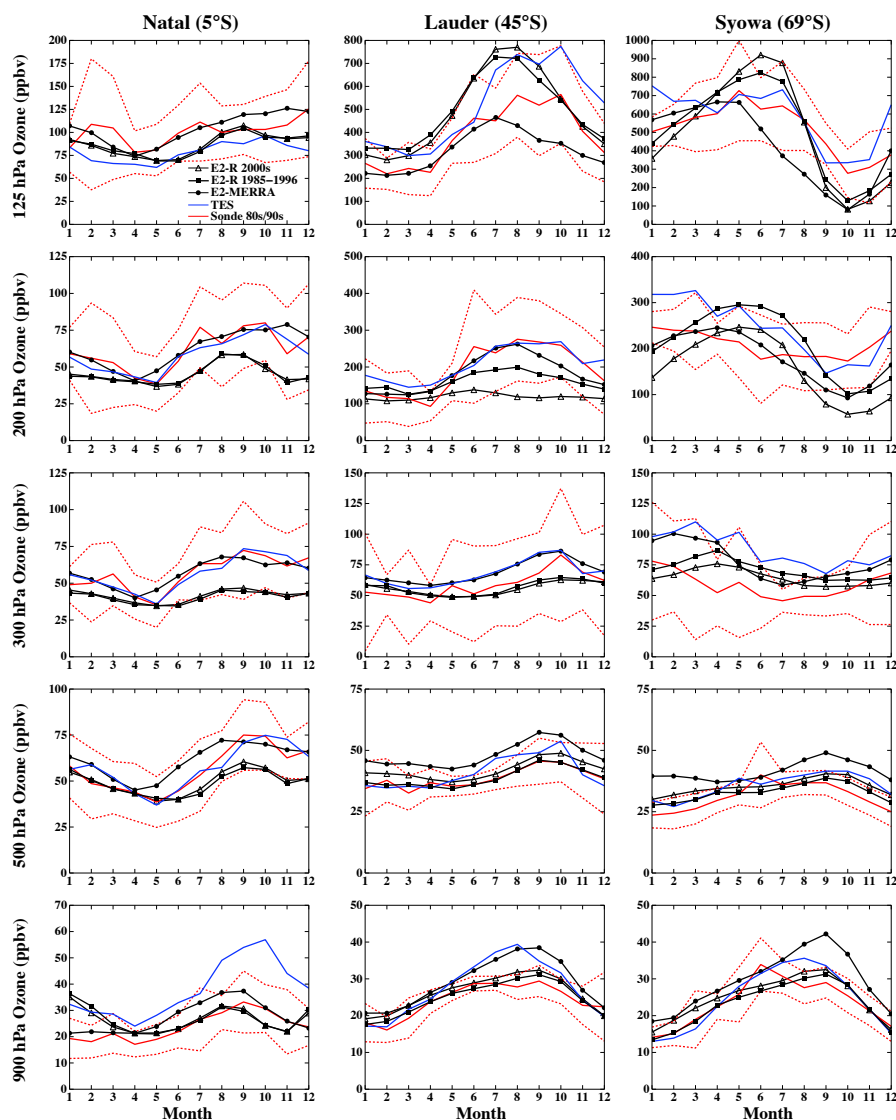


Fig. 2. Continued.

levels. Results from E2-H are very similar to those from E2-R (not shown). The E2-MERRA simulation shows moderately reduced differences in comparison with sondes at most levels, but is not dramatically different when compared with the equivalent E2-R 2000s simulation (E2-MERRA biases are also very similar to those in E2-R).

The TES instrument provides vertical information throughout the atmosphere. TES data generally compare reasonably well with sondes at most locations (Fig. 2), but previous, more extensive comparisons show that in general, TES has positive biases throughout the troposphere typically in the range of about 3–10 ppb (Nassar et al., 2008). We compare zonal means from TES with both model versions. We use monthly mean output from the model, since test comparisons using high temporal resolution data show little difference, consistent with earlier analyses (Aghedo et al., 2011a).

As pointed out previously (Nassar et al., 2008), TES has low sensitivity in the lower troposphere at high latitudes, as seen in our zonal mean analysis in which not enough data was present for comparison at those locations.

In comparison with annual mean values from TES, both models tend to have too much ozone in the upper troposphere everywhere north of about 50° S (Fig. 3). TES typically shows fairly good agreement with sondes at these levels, so the results suggest that there may be too much photochemical production in the model at higher tropospheric levels. This could result at least partially from excessive lightning NO_x at upper levels and/or excessive vertical transport of pollutants from the surface. Our lightning NO_x production of 7.3 Tg N yr⁻¹ is at the higher end of the 2–8 Tg N yr⁻¹ range found in most models, but well within published estimates spanning 1–20 Tg N yr⁻¹ (Schumann and Huntrieser,

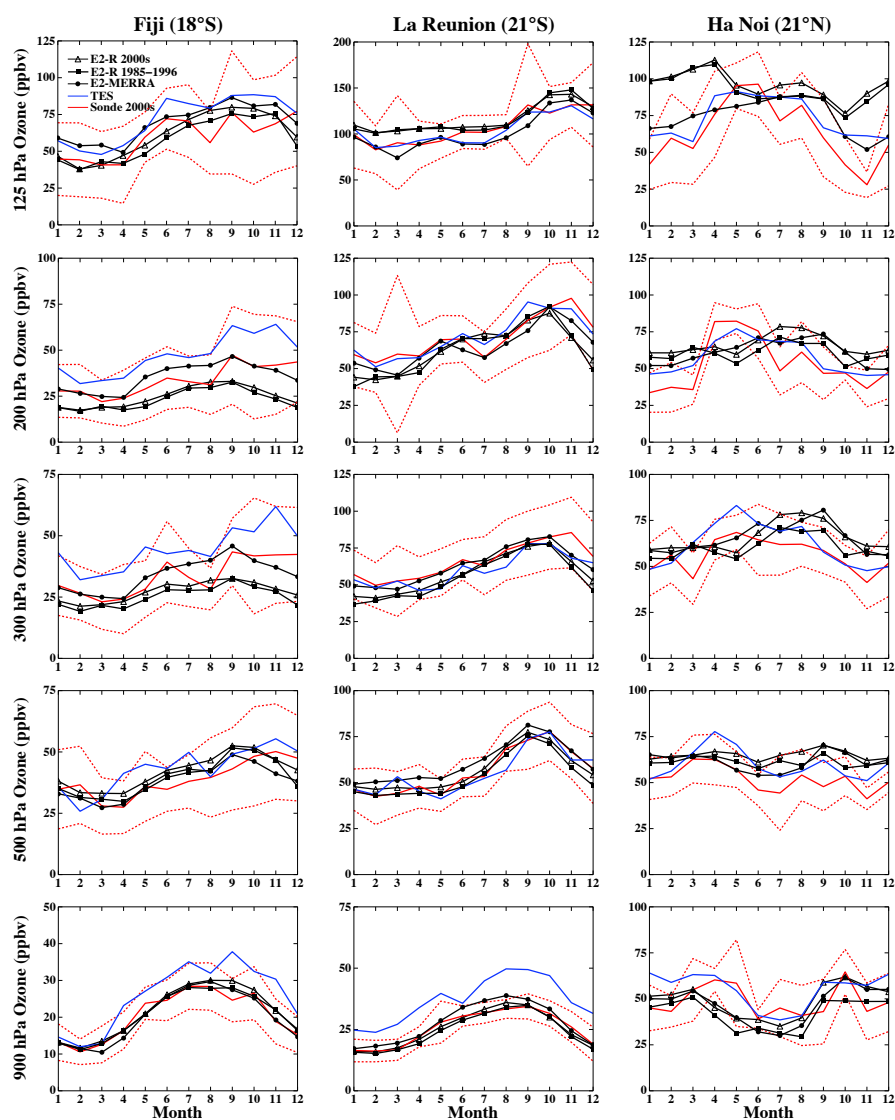


Fig. 2. Continued.

2007). Recent analysis suggests that lightning NO_x production estimates may indeed be biased high (Beirle et al., 2010). However, we note that the model shows no obvious bias in comparison with most sonde data in the upper troposphere. The models also show too little ozone below about 300 hPa in the tropics. However, TES appears to have a positive bias at low levels in the tropics in comparison with sondes (Fig. 2), while the model performs well in this region with respect to sondes. The model also has less stratospheric ozone than TES in most areas, though the TES instruments sensitivity is lower at those altitudes. The comparison between the two model runs shows that when MERRA winds are imposed and observed SSTs used, biases are reduced at high Southern latitudes, but are increased at high Northern latitudes, while remaining similar at lower latitudes (Fig. 3). These results are consistent with those seen in the sonde comparisons.

3.1.3 Evaluation of tropospheric and stratospheric columns

With some knowledge of the vertical structure of ozone, we now examine the column amounts in the troposphere and stratosphere separately. Given the much larger amount of ozone present in the stratosphere and the high gradient near the tropopause, the tropospheric ozone column estimates are quite sensitive to the location of the tropopause, with variations in the definition of the tropopause affecting the tropospheric column amount by $\sim 10\%$ (Wild, 2007). We first compare the annual average E2-R and E2-MERRA tropospheric ozone columns using the model's internally calculated tropopause based on the World Meteorological Organization definition with the TES column calculated using the same definition applied to NCEP reanalyses (Fig. 4). The

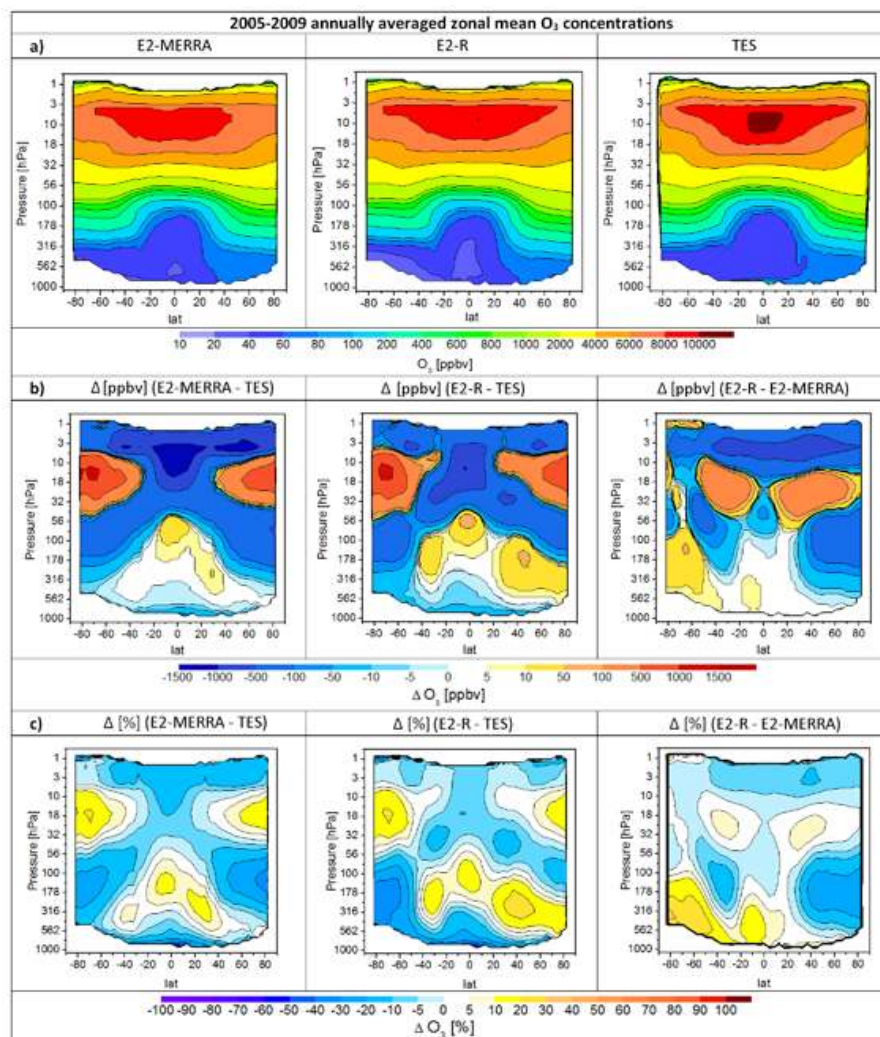


Fig. 3. Zonal mean ozone (ppbv) in the MERRA-nudged run (top left), the ACCMIP/CMIP5 E2-R run (top center) and in TES observations (top right). Differences between the model runs and TES and between the two model runs are shown in the second row in ppbv and in the third row in percent.

E2-MERRA run shows a broadly similar pattern to TES. In the E2-R run, however, the tropospheric column amounts are substantially larger in much of the Northern extratropics, especially at high latitudes, while they are substantially smaller over Southern high latitudes. Biases in the tropospheric column require careful interpretation. The large positive biases poleward of 60° N, for example, result from zonal mean ozone biases of < 15 % (Fig. 3). The positive ozone bias, however, leads to additional heating of the upper troposphere, raising the tropopause and thus adding more ozone into the larger troposphere. Similar, but opposite sign, responses are at work at Southern high latitudes. As the models have identical chemistry, these biases must be related to circulation. Hence while it is certainly useful to evaluate the tropospheric column using a consistently defined tropopause

based on the meteorology present in each individual case, it is then difficult to disentangle differences in the volume contained within the troposphere from differences in ozone itself. We thus utilize a fixed location of the troposphere from the NCEP analyses, in the remainder of our analyses in this section rather than the model's internally calculated location.

In addition to the TES measurements, simultaneous observations of total column ozone from OMI and stratospheric ozone from the Microwave Limb Sounder (MLS) (Froidevaux et al., 2008) allow the tropospheric ozone column to be determined based on the residual of OMI minus MLS ozone (Ziemke et al., 2011). The tropopause in the OMI/MLS analysis is defined using the NCEP reanalysis. OMI minus MLS shows substantially less tropospheric ozone than that seen in the models or TES using the same tropopause definition.

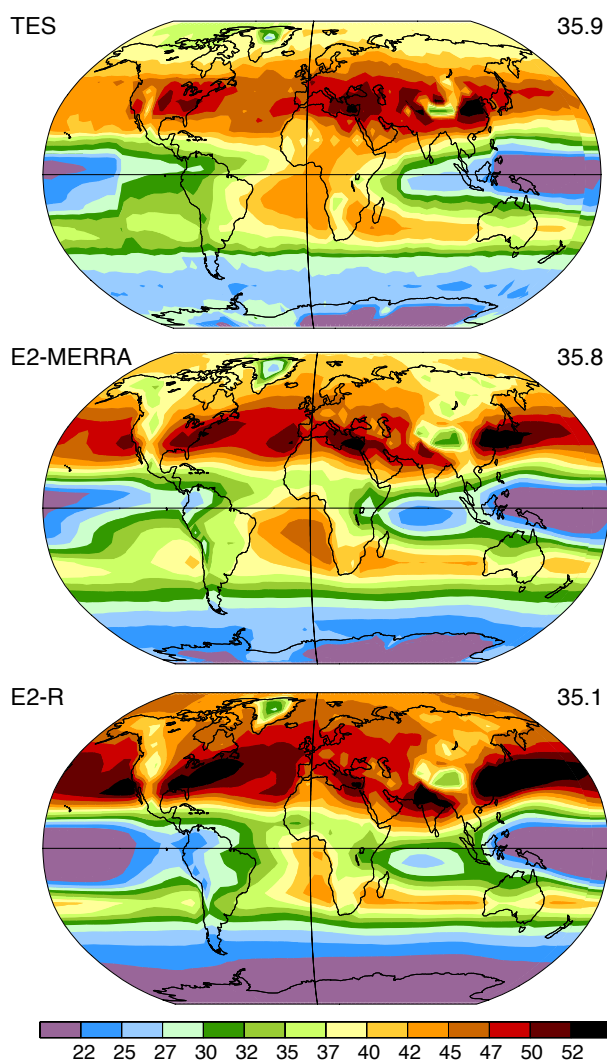


Fig. 4. Annual average 2005–2009 tropospheric ozone column (DU) in the indicated observations and models. The tropopause is defined using the NCEP 2005–2009 monthly values for TES and the model's internally calculated values. Global mean values are in the upper right.

Hence for ease of visual comparison we show the geographical distribution of tropospheric ozone columns from the models and TES using a tropopause one level below the NCEP tropopause along with OMI/MLS observations (which lack vertical information for sampling with other tropopause definitions).

The tropospheric column values in the model agree reasonably well with TES, but the TES and modeled values are substantially larger than the OMI/MLS column values over most of the planet despite the higher tropopause in the latter (Fig. 5). There is good agreement between the model and TES, even at high latitudes, when using a common tropopause location. This indicates that the model's internal tropopause is too high at Northern high latitudes and too low

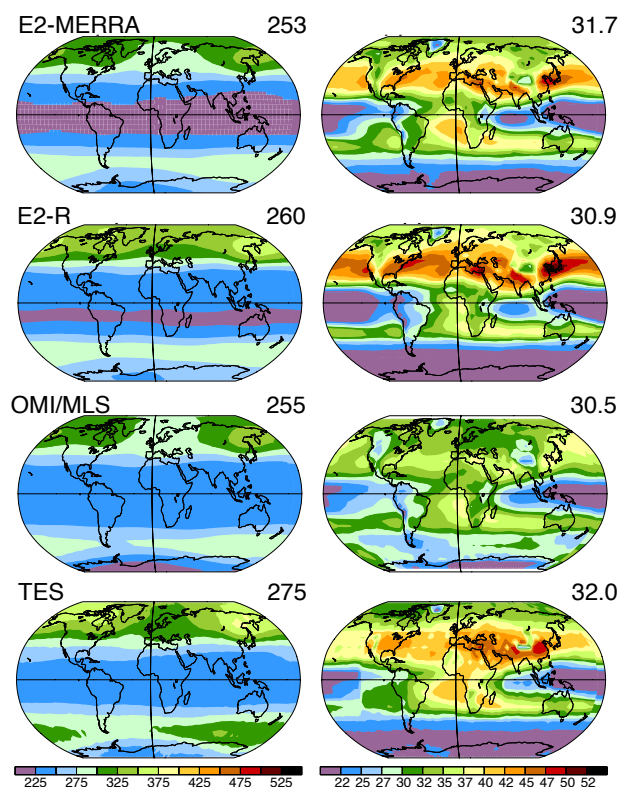


Fig. 5. Annual average stratospheric (left panels) and tropospheric (right panels) ozone column (DU) in the indicated models and observations. OMI/MLS observations are from October 2004 to December 2010, model values are 2005–2009, and TES observations are from 2005–2009. Global mean values are given in the upper right of each panel. The tropopause is the NCEP value for OMI/MLS, and one model level below NCEP for the models and TES.

at Southern high latitudes compared with NCEP. The models (and TES) agree much more closely with OMI/MLS at high Southern latitudes when the NCEP tropopause, rather than one level below the NCEP tropopause, is used to determine the upper boundary of the troposphere, but then the agreement is worse elsewhere. This suggests a consistent difference between biases in the height of the model's tropopause (and thus ozone) relative to NCEP between Southern high latitudes and elsewhere.

Both model versions capture many of the features seen in both satellite tropospheric column datasets, including the minima over high topography regions in the Himalayas, Rockies, Andes, Greenland and Antarctica, the maximum over the Atlantic off the west coast of tropical Africa, and the local minimum over tropical east Africa. Both models also show pollution outflow from Africa out across the Indian Ocean at around 30° S, from North America across the Atlantic and from East Asia across the Pacific that matches the location and extent of that in the observations reasonably well. The Atlantic outflow appears to extend too far

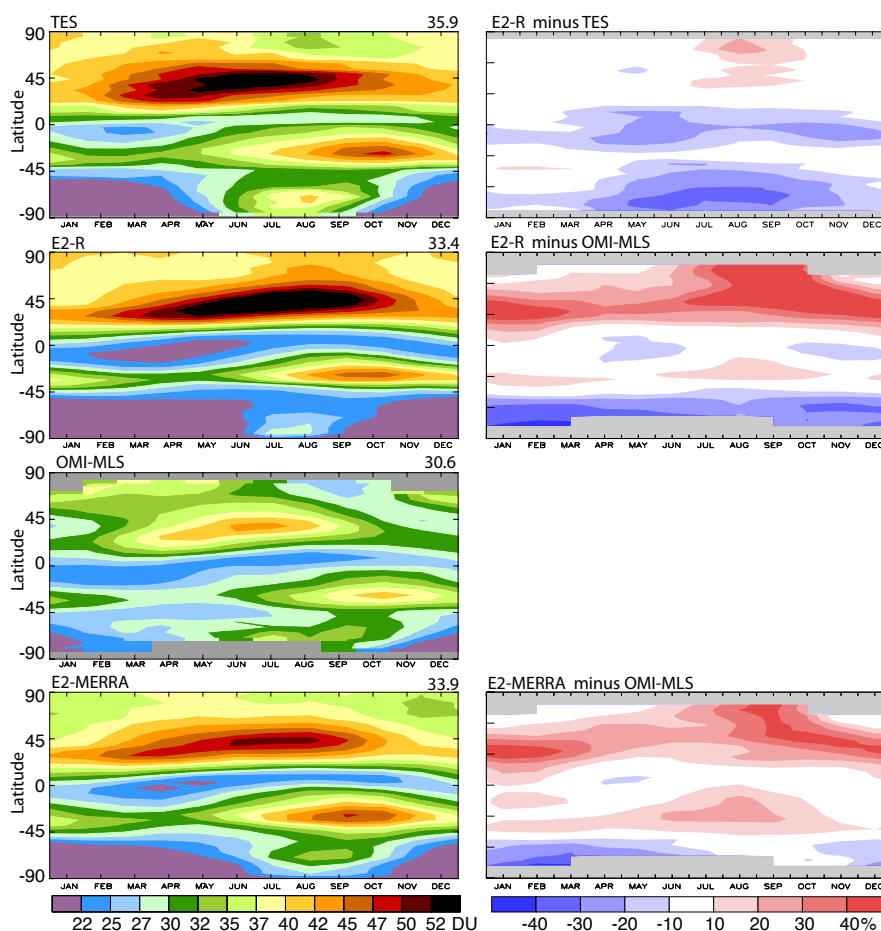


Fig. 6. Seasonality of zonal mean tropospheric ozone column (DU; left column) in the given models and observations and model-observation differences (%; right column). Global mean column values are given in the upper right of each panel. All analyses use the NCEP tropopause definition. OMI-MLS observations are from October 2004–December 2010, TES observations are from 2005–2009. Model differences are based on analysis of the same years as in the observations (2005–2010 or 2005–2009).

north, however, while the outflow from East Asia appears to be too strong, especially relative to TES. Local maxima over regions with both substantial precursor emissions and sunlight, including the eastern Mediterranean, Middle East and the northern part of South Asia, are also captured, as are high values over East Asia and the Eastern, and especially Southeastern, US. Ozone off the coasts of Southern Hemisphere Africa is larger in the E2-MERRA run than the E2-R run, with the former agreeing especially well with TES while both have larger values but similar patterns to OMI/MLS. At high Southern latitudes, both models are much closer to TES than to OMI/MLS, with the E2-MERRA run again showing especially good agreement.

Low latitude tropospheric column values are generally well-simulated, consistent with the comparisons against sondes (Fig. 2). The good agreement between the model and sondes at Hilo and the fact that the E2-MERRA run has too little ozone in comparison with both Arctic and Northern mid-latitude sondes suggests that the apparent Northern middle

and high latitude positive bias in tropospheric column ozone with respect to OMI/MLS data seen even in the E2-MERRA run may in fact be an artifact of vertical sampling (as discussed above).

The stratospheric ozone column's geographic distribution is generally well simulated in both the E2-R and E2-MERRA models. Both show a Northern extratropical wave 1 pattern (a planetary scale wave with 1 node in the zonal direction) with a local maximum over eastern Siberia, a wave 1 pattern in the Southern extratropics with a maximum at a similar longitude, and little longitudinal variation in the tropics (Fig. 5). The amplitude of the Northern extratropical wave 1 is captured quite well in the E2-MERRA run, but is underestimated in the E2-R simulation. Analysis of the TES stratospheric column using the NCEP tropopause, rather than one level below as in Fig. 5, shows a very similar global mean column of 271 DU. Thus using either definition, TES has nearly 10% more stratospheric ozone than MLS, which may account for

the model having generally less ozone than TES in the stratosphere (Fig. 3).

We also analyze the seasonality of the tropospheric column. Using the NCEP tropopause in all analyses, the model captures the general seasonal pattern of mid-latitude maxima in the spring/summer in each hemisphere, a year-round local minimum in the tropics, and a winter maximum in the Antarctic (Fig. 6). The E2-R simulation tends to agree well with TES over most of the year, though there is a persistent negative bias in the Southern tropics and at Southern high latitudes. Note that TES appears somewhat low compared with sondes at Natal (5° S) at 125 hPa (Fig. 2), so that small biases may not be significant. The high latitude biases show marked seasonality, suggesting transport from the stratosphere may be slightly too strong in the Arctic during boreal fall (though generally good in the Northern Hemisphere) and may be too weak during austral winter in the Antarctic. In comparison with OMI/MLS, the modeled mid-latitude maxima are too large and values are again too large during July–September at high Northern latitudes. The E2-MERRA run shows roughly the same global mean tropospheric ozone column as the E2-R run, with both models falling between the TES and OMI/MLS values.

The apparent high biases in E2-R at Northern mid-latitudes with respect to OMI/MLS are somewhat reduced in E2-MERRA, while those at Southern mid-latitudes are increased (though Southern high latitudes seem slightly more realistic). The comparison suggests that downward flux of ozone from the stratosphere is larger in the MERRA run than in the free-running model in the Southern Hemisphere, but smaller in the Northern Hemisphere. The comparison against sondes at Hohenpeissenberg, however, suggests that the slower transport in E2-MERRA may in fact be too slow (Fig. 2), and in fact the E2-R results agree much more closely with TES.

The model-sonde comparisons together with the tropospheric ozone column analysis thus indicates that the E2-R run has reasonable downward fluxes of stratospheric air in the Northern Hemisphere while the E2-MERRA run has too little. The situation is reversed in the Southern Hemisphere where the E2-MERRA run performs well while the E2-R model has too little downward flux.

3.1.4 Ozone metrics and budget

It is difficult to find simple quantitative metrics to summarize comparisons between models and observations since the fields exhibit spatial variations in the vertical and horizontal and temporal variability at various scales such as seasonal and interannual, all of which could be examined. In addition to the comparison with sondes shown previously (Table 1), we have examined the absolute value of the percent difference in the zonal mean versus latitude analysis (Fig. 1) for the different model configurations, averaging over various latitude bands. The global mean area-weighted differences

are just over 4 % in all the E2 configurations (Table 2). The E2-R model outperforms the older E-R model (Shindell et al., 2006b) globally and in each latitude band. As noted previously, the older version had poor seasonality at high latitudes, and indeed improvements are larger in absolute terms in the extratropics, especially the Northern extratropics, for E2-R relative to E-R, than they are in the tropics. While the E2-MERRA simulation is similar to E2-R globally and in the Southern extratropics, it is worse in the tropics and better in the Northern extratropics. The E2-H model is quite similar to E2-R.

We have also explored the representation of the observed annual average horizontal spatial distribution in the model by calculating the linear Pearson correlation coefficient R^2 on the fields shown in Fig. 5 (Table 2). Correlations are extremely high for the stratospheric columns, and also generally quite good for the troposphere, especially in comparison with TES. It appears that the use of observed SST/sea-ice and nudging towards reanalysis winds improves the distribution of ozone in the troposphere (much more so comparing with OMI/MLS). However, these correlations must be interpreted with some caution, as the R^2 between TES and OMI/MLS is only 0.61 for the troposphere (though it is 0.97 for the stratosphere). Our correlation with OMI/MLS in the troposphere ($R = 0.85$) compares well with other recent models that found correlations of 0.82 ± 0.07 , while our bias of 2.6 DU is well within the reported range of -4.9 to 4.0 DU (Young et al., 2013).

We also examine the degree of inhomogeneity in ozone, which provides a different measure of how well the models capture the observed spatial distribution. In the stratosphere, this is primarily a function of large scale dynamics. In the troposphere it is primarily indicative of how efficiently the models transport pollutants from source regions, though chemical processing timescales will also play a role (Shindell et al., 2010). We use the spatial autocorrelation, in which each grid point value is related to the mean value on circles around this grid point. The circles for a given characteristic radius distance are established such that they are equal area circles on the sphere around all grid points. Autocorrelations are calculated after removing the mean and normalizing the field by its standard deviation to better discern spatial variations. Here we analyze the mean autocorrelation over a distance of 1–3000 km for each point, a length scale chosen to emphasize spatial correlations over relatively short distances that are especially characteristic of tropospheric distributions. Autocorrelation values in both the E2-R and E2-MERRA versions of the model agree well with both observational datasets for the stratosphere. In the troposphere, they agree well with TES, but are too high compared with OMI/MLS. Thus there is a possibility that horizontal mixing may be too rapid in the model, but it is difficult to know given the sensitivity of the comparison to the satellite dataset used for this evaluation.

Another common way to characterize tropospheric ozone is to examine its budget terms. The ozone budget for

Table 2. Evaluation of model fields against satellite observations.

Absolute value % difference vs. TOMS/OMI (2000–2010 average, zonal mean vs. month)	E-R	E2-H	E2-R	E2-MERRA	
Global	5.4	4.0	4.2	4.1	
Tropics	2.7	1.9	2.0	3.2	
Southern Extratropics	6.5	5.0	5.3	5.3	
Northern Extratropics	10.2	7.7	7.5	5.0	
Evaluation against OMI/MLS and TES tropospheric and stratospheric columns			E2-R	E2-MERRA	
R^2 correlation with OMI/MLS: Troposphere			0.71	0.78	
R^2 correlation with TES: Troposphere			0.85	0.87	
R^2 correlation with OMI/MLS: Stratosphere			0.89	0.89	
R^2 correlation with TES: Stratosphere			0.95	0.95	
Evaluation against MOPITT mid-tropospheric CO (500 hPa retrieval level)			E2-R	E2-MERRA	
R^2 correlation; April			0.81	0.81	
R^2 correlation; October			0.56	0.64	
Autocorrelations: average 1–3000 km			OMI/MLS	TES	
Tropospheric ozone	0.68	0.84	0.85	0.82	
Stratospheric ozone	0.85	0.86	0.90	0.88	
			OMI	E2-R	E2-MERRA
Tropospheric NO ₂	0.44		0.56	0.52	
			MOPITT	E2-R	E2-MERRA
Tropospheric CO; April	0.94		0.95	0.95	
Tropospheric CO; October	0.69		0.81	0.82	

the troposphere is sensitive to the definition of the tropopause, however. Using the 150 ppbv ozone contour to define a “chemical” tropopause, the net chemical production is 1085 Tg yr⁻¹, dry deposition is 1378 Tg yr⁻¹, and stratosphere-troposphere exchange (STE) is 292 Tg yr⁻¹. Using the model’s “nominal” tropopause of 150 hPa, STE is 462 Tg yr⁻¹. The latter value of STE compares well with constraints from observations, which give a best estimate of 450 Tg yr⁻¹ (range 200 to 870 Tg yr⁻¹) (Murphy and Fahey, 1994), a range of 450–590 Tg yr⁻¹ at 100 hPa (Gettelman et al., 1997a), and a value for the downward extrapolar flux (~80–100 % of total downward flux) of 470 Tg yr⁻¹ for the year 2000 (Olsen et al., 2003). Using the 150 ppb contour gives a value that is clearly on the low end of the observationally-constrained range. STE in E2-MERRA is 502 Tg yr⁻¹, again using the 150 hPa tropopause, slightly larger but also quite consistent with observation-based constraints. Downward flux in the NH is 53 % of the total in the E2-R model, slightly less than the 55–57 % inferred from satellite data (Gettelman et al., 1997b; Olsen et al., 2003). The tropospheric ozone burden is 439 Tg using the nominal 150 hPa level, and 352 Tg using the 150 ppb chemical tropopause. Our burden is similar to that found in previous multi-model evaluations which used the chemical tropopause and found an ozone burden of 344 ± 39 Tg (Stevenson et

al., 2006) and 337 ± 23 Tg (Young et al., 2013) (versus our 352 Tg), and similar to the 335 ± 10 Tg estimated from observations (Wild, 2007). There are differences in the budget terms leading to that burden, however, which are 442 ± 309, 1003 ± 200 and 552 ± 168 Tg yr⁻¹ for net chemical production, dry deposition and STE, respectively in Stevenson et al. (2006) (versus our 1085, 1378 and 292 Tg yr⁻¹). As discussed above, however, observational constraints are only available for STE, and these are quite sensitive to the tropopause definition.

3.2 Tracers of stratospheric circulation

To better understand the biases in stratospheric ozone, which the above analysis suggests are largely transport related, we have evaluated circulation using an age-of-air tracer and nitrous oxide (N₂O). A compact correlation exists between these two in the atmosphere, and hence analysis of this correlation provides a useful diagnostic of circulation in models. The N₂O data is a monthly climatology derived from observations from the Atmospheric Chemistry Experiment (ACE) satellite instrument onboard SCISAT-1 during 2004–2009, while the age data comes from CO₂ and SF₆ data from balloon flights, all as described in Strahan et al. (2011) and Andrews et al. (2001). The observed correlation between N₂O and age is captured reasonably well in the model, but

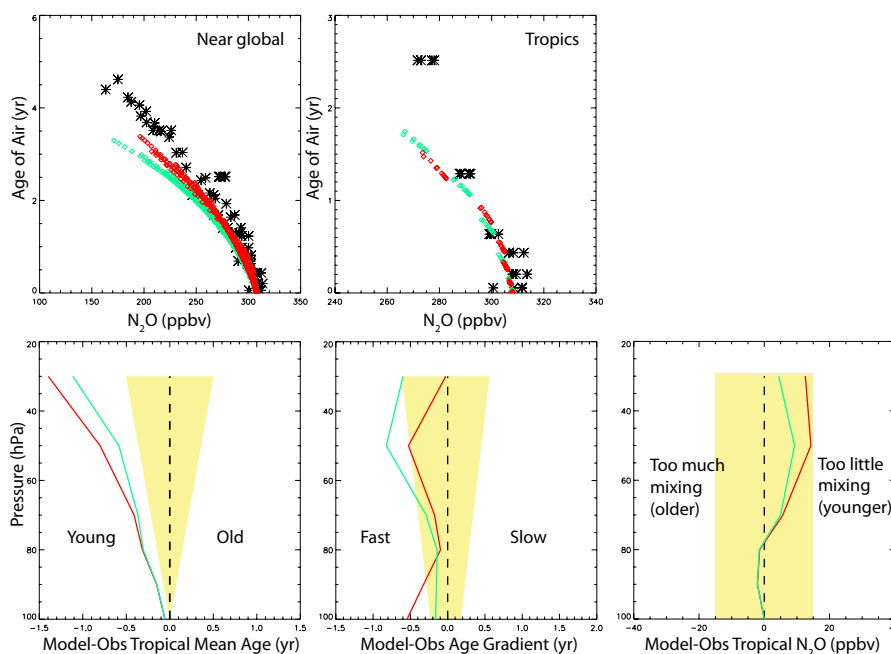


Fig. 7. Diagnostics of stratospheric circulation from the E2-R simulations (green), the E2-MERRA run (red) and observations (black). N_2O and age of air points are from 150 to 50 hPa, 65°S to 65°N plus 50–30 hPa for 40°S to 40°N (upper left) or from 100–30 hPa from 10°S to 10°N for the tropical analysis (upper right), in both cases with model N_2O normalized to observations at 100 hPa, 10°S – 10°N . The age gradient is the difference between the mean ages averaged over 35 – 50°N and 10°S – 10°N . Yellow areas are approximate ± 1 sigma uncertainties from the observations.

the maximum age values in E2-R are less than observations while the N_2O range is in relatively good agreement (Fig. 7). This suggests that the circulation is somewhat too fast, leading to too much photolytic destruction of N_2O for a parcel of a given age. As described in Strahan et al. (2011), analysis of tropical N_2O diagnoses the effect of horizontal mixing, analysis of the horizontal mean age gradient evaluates the tropical ascent rate, and analysis of tropical mean age examines the influence of both ascent and mixing together. Thus for the E2-R simulation, the horizontal mixing appears to be reasonably good, while the tropical ascent is indeed too fast and hence tropical mean ages are too short (with values comparable to other models which showed too rapid circulation in Strahan et al., 2011). Similar biases are apparent, although generally to a lesser extent, in the E2-MERRA simulation.

The stratospheric circulation being too fast in the E2-R model is consistent with the overestimates of total column ozone during the polar winter when transport dominates over chemistry (Fig. 1), and in the annual average positive biases in zonal mean ozone at high latitudes from about 30–5 hPa in both hemispheres (Fig. 3). Biases in total column ozone in the E2-MERRA run are of comparable magnitude to those in the E2-R run, but different seasonality (Fig. 1). Comparison with ozone sondes in the lower stratosphere showed substantial improvement in Southern Hemisphere mid-latitudes using MERRA and observed SSTs, but Northern middle and high latitudes were substantially worse. Thus although the

E2-MERRA run's circulation seems to be quite similar (and marginally better) in the annual mean in comparison with the observed N_2O -age correlation, it appears that the seasonality and differing hemispheric biases lead to comparable magnitudes of high latitude ozone biases. Note that this analysis is not purely an evaluation of the MERRA product, as while winds are relaxed towards MERRA they still reflect the influence of the model's internally derived circulation.

3.3 Ozone precursors

Ozone is not directly emitted, but is a secondary pollutant in the troposphere that is formed photochemically in the presence of nitrogen dioxide (NO_2) and carbon monoxide (CO) or hydrocarbons. Realistic reproduction of tropospheric ozone therefore requires realistic simulations of these precursors. Many hydrocarbons play a role, but global observations are not available for most of these other than methane. As surface methane is prescribed at observed amounts, comparison of methane with observations is not informative of model processes, and thus we concentrate our evaluation of precursors on NO_2 and CO.

3.3.1 Nitrogen dioxide

In addition to observing ozone, OMI also measures NO_2 (Celarier et al., 2008). To facilitate comparisons with OMI

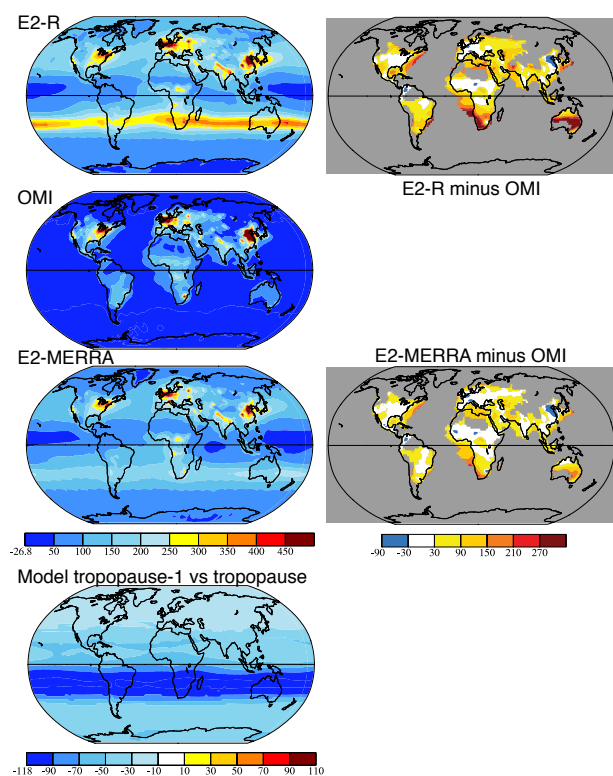


Fig. 8. Annual average 2005–2010 tropospheric NO_2 column (10^{13} molecules cm^{-2}) at 13:30 local time in the E2-R run (top left), OMI observations (2nd row left), and the E2-MERRA run (3rd row left). Percentage differences between the models and OMI are shown in the right column for areas with at least 50×10^{13} molecules cm^{-2} . Also shown is the column difference in the E2-MERRA run when sampled using the model's tropopause and one level below the tropopause (lower left).

tropospheric NO_2 , we saved the tropospheric NO_2 column at $\sim 13:30$ local time during all our model runs. Those column values include all NO_2 from the surface up through the highest level with a mid-layer pressure greater than (or equal to) the pressure of the model's tropopause (meteorological). Hence the tropopause may be either in the layer above or in the upper half of the top layer included in our column in this case. The definition of the tropopause is important, as the OMI data processing to yield tropospheric columns has removed stratospheric NO_2 based on stratospheric values and the tropopause location in the TM4 model. This removal of stratospheric NO_2 is clearly imperfect as it leads to large negative values in the OMI data in many areas (hence the removal of the estimated stratospheric NO_2 has taken out more NO_2 than is present in the entire column measurement) (Fig. 8). The model simulations show large biases in areas where there is substantial NO_2 in the vicinity of the tropopause, namely in mid-latitude bands in both hemispheres, but especially the Southern. That these are related to stratosphere-troposphere exchange is implied by the fact that

the biases vary greatly between the E2-R and E2-MERRA simulations, which differ primarily in their circulation. They also differ in the strength of convection, which transports NO_2 from surface sources to the upper troposphere and increases production of lightning NO_x , which is 9.6 Tg N yr^{-1} in E2-MERRA and 7.3 in E2-R. However, both these differences would tend to increase upper tropospheric NO_x , while E2-MERRA has less NO_x than E2-R, indicating that differences in STE are responsible (along with possible differences in the location of the tropopause in the two simulations). The importance of NO_2 near the tropopause is also highlighted by looking at the impact of sampling the model up to the tropopause versus up to one level below the tropopause (Fig. 8). Removing the top level of the tropospheric column lowers the global mean by 23 % in the E2-MERRA run, a large portion of the 61 % overestimate relative to OMI seen in that simulation. Using an ozone concentration of 150 ppbv to define a ‘chemical tropopause’ also reduces the global mean column amount by 23 %. In either case, the remaining bias is on the order of 38 %. Much of this bias may be due to an underestimate of NO_2 by OMI, which finds tropospheric column values roughly 15–30 % below validation measurements (Celarier et al., 2008). However, at least part of the positive NO_2 bias in the model seems real, and may account for some of the positive ozone bias seen in the upper troposphere in comparisons with TES.

While the model appears to overestimate NO_2 in remote areas, annual mean values reported by OMI in most remote locations are less than 50×10^{13} molecules cm^{-2} . Values are as low as -25×10^{13} molecules cm^{-2} , suggesting that in remote regions, where values are comparable in magnitude to the apparent biases, values are not terribly reliable. Hence we show differences relative to OMI only in areas where values are greater than 50×10^{13} molecules cm^{-2} (Fig. 8). Over source regions, the model typically underestimates the NO_2 column, especially over north-west Europe and East Asia. The underestimate over NE China is particularly large, more than a factor of two in some locations, suggesting that surface emissions there are likely too low in the Lamarque et al. (2010) inventory. There is also an associated high bias over the Sea of Japan and the Pacific just east of Japan, suggesting that export of NO_2 from source regions may be too high in East Asia, as was also seen for ozone. More generally, positive NO_2 biases are typically correlated with positive ozone biases in outflow regions, presumably due to overestimated transport of both species, while biases in these two compounds are not well correlated in other areas. Evaluation of the autocorrelation of the tropospheric NO_2 field shows that in both models the values are too high, though they are somewhat closer to the observations in the E2-MERRA run (Table 2). This indeed indicates that horizontal transport from source regions is too rapid in the model.

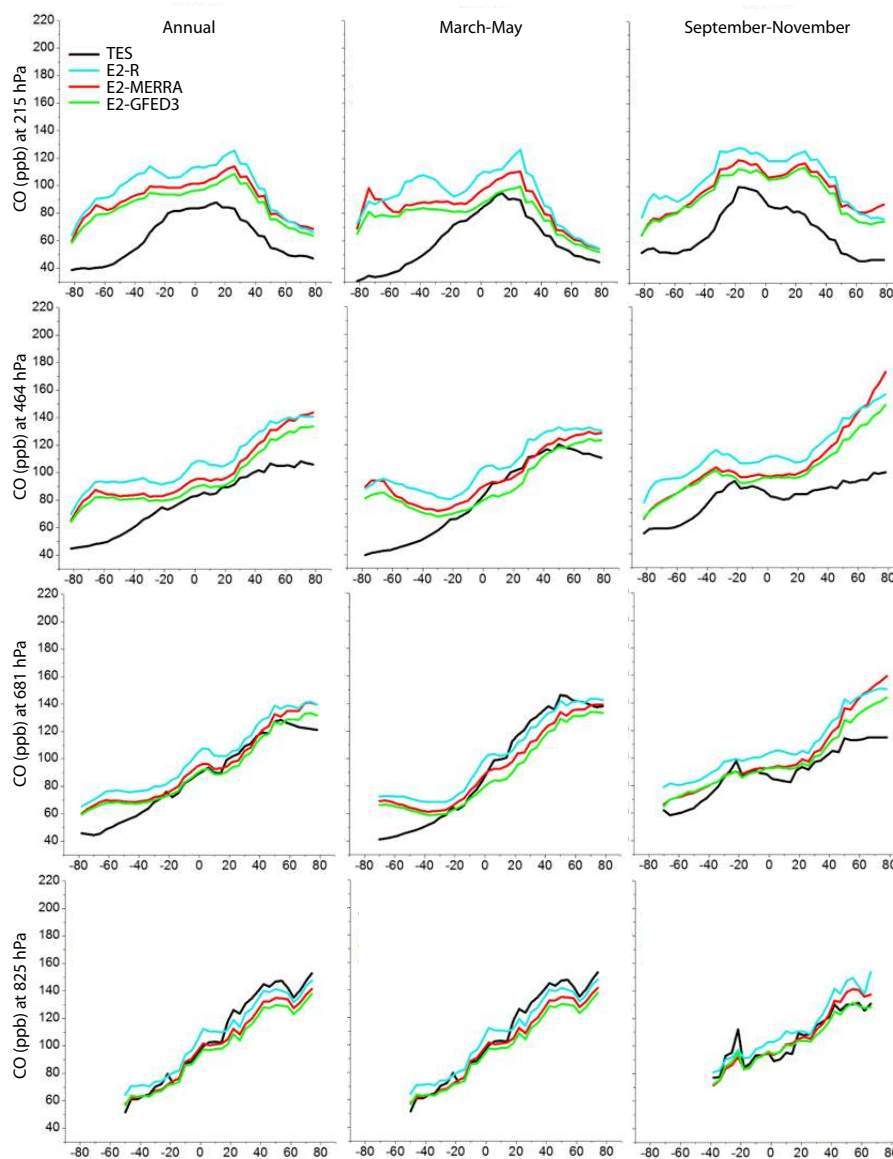


Fig. 9. Zonal mean CO at the indicated pressure levels in 2005–2009 TES observations and in three model simulations for the annual average and boreal spring and fall.

3.3.2 Carbon monoxide

In addition to the ozone observations shown previously, the TES instrument also provides vertical information on CO in the troposphere. We show comparisons of zonal mean TES CO at several pressure levels (nominal levels of the broader TES averaging kernels) in Fig. 9. For the annual average, the model does very well in capturing the near-surface (825 hPa) level zonal mean, though the E2-R model tends to slightly overpredict CO in the Southern Hemisphere while the E2-MERRA run underpredicts in the Northern Hemisphere. The model shows positive biases at higher levels, with high latitude biases appearing first (e.g. at 681 hPa) and biases at all

latitudes in the middle and upper troposphere. Interestingly, biases are approximately the same magnitude at all latitudes in the upper troposphere, for example at 215 hPa (Fig. 9). This suggests that the progressively larger biases with altitude are likely the result of the ratio of production of CO from hydrocarbon oxidation relative to CO's own oxidation being too large, rather than from excessive vertical transport of surface sources. In the upper troposphere, production of CO is almost exactly split between production from methane and that from non-methane VOCs (the latter primarily from paraffins). Upper tropospheric biases are typically greatest in fall and least in spring in the Northern extratropics, while biases are largest during winter and fall and least in spring

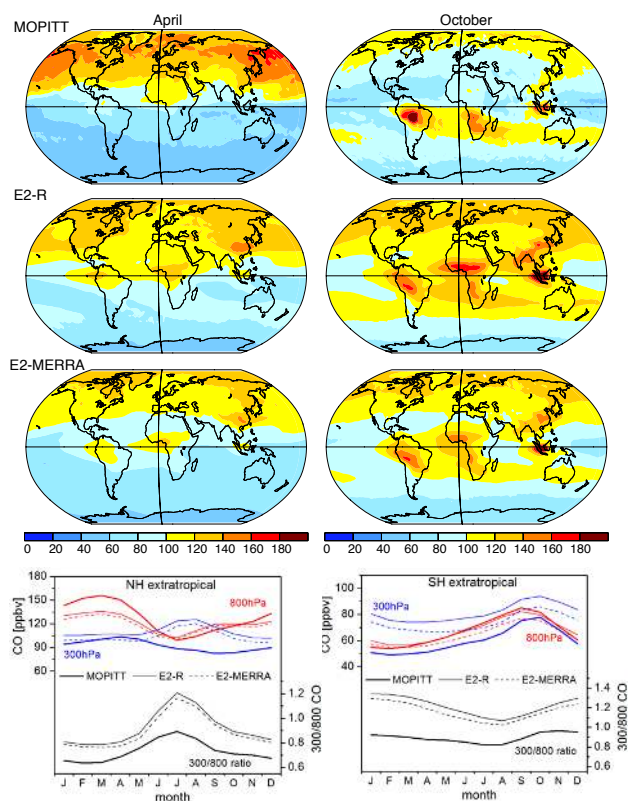


Fig. 10. Mid-tropospheric CO (ppb) in MOPITT observations and in the E2-R and E2-MERRA simulations for 2005–2009 (top three rows). Values are given for the broad retrieval level centered near 500 hPa. Seasonal cycle of CO (ppb) in MOPITT observations and the models for the retrieval levels centered at 300 and 800 hPa and their ratio in the extratropics (bottom row).

in the Southern Hemisphere. Upper tropospheric biases in the tropics are similar throughout the year. The positive biases in the upper troposphere, along with the positive biases in NO_2 , may contribute to the positive biases in upper tropospheric ozone seen in the model in comparison with TES (Fig. 3). Additional evaluation of a slightly earlier, but very similar, version of the GISS-E2 model was presented in Voulgarakis et al. (2011) which compared ozone/CO correlations against those seen by TES. That analysis demonstrated that the GISS-E2 model performed quite well in capturing the observed correlations, even in relatively remote areas where they were highly influenced by transport in addition to emissions.

An additional simulation that is otherwise identical to E2-MERRA except that it uses the GFED-3 biomass burning emission inventory (van der Werf et al., 2010) shows fairly similar results (Fig. 9). Near-surface (825 hPa) level CO is in better agreement with TES during boreal fall, but the Northern Hemisphere underestimates in the model are slightly larger in that run during other seasons. This suggests that biases in biomass burning emissions in the Lamarque et

al. (2010) inventory are not likely to be a large contributor to the model biases with respect to TES observations.

While TES provides valuable information on the vertical structure of CO in the troposphere, observations from the Measurements of Pollution in the Troposphere (MOPITT) instrument provide better spatial coverage (Deeter, 2003). We therefore evaluate the CO distribution by comparison with MOPITT data averaged over 2005–2009, focusing on April when Northern Hemisphere CO concentrations peak and October when biomass burning emissions of CO in the tropics are large. We focus on the MOPITT retrieval level centered around 500 hPa, representative of a broad area of the mid-troposphere. The model is sampled using the MOPITT averaging kernel and a priori profiles. As in the comparisons with TES, the model shows too little seasonality in the Northern extratropics (Fig. 10). In the Southern extratropics, the E2-R version again shows too little seasonality, but the E2-MERRA version looks quite realistic. The tropics show a transition from relatively little CO during April to large CO concentrations located over land areas in October. While South America stands out in the MOPITT observations, the model shows highest CO over the maritime continent. CO emissions during the 2000s are interpolated between the 2000 and 2010 datasets. As for the former is based on 1996–2004 emissions, it includes the effects of large fires over Indonesia during the extremely large 1997–8 El Niño event. The model also shows more CO over Africa north of the equator, while MOPITT shows larger abundances to the south, again suggesting regional differences in biomass burning emissions. Looking at the seasonality in the lower and upper troposphere shows that the models perform reasonably well in the lower troposphere, but substantially overestimate CO in the upper levels especially in local summer (Fig. 10). These results are consistent with the increasing high bias with altitude in the models seen in the comparison with TES (Fig. 9). Biases in the 300/800 hPa CO ratio are fairly similar throughout the year and in both hemispheres, supporting the hypothesis that these result from large-scale production of CO from hydrocarbon oxidation being too large relative to CO oxidation.

Statistical comparison of the model's mid-tropospheric CO with MOPITT shows that the models perform better in April than October, and that the two versions have nearly equal R^2 spatial correlations (Table 2). The R^2 correlations reported here of 0.81 and 0.56 for April and October are substantially better than those reported in an earlier multi-model evaluation (Shindell et al., 2006a) that found values from 0.19 to 0.72 for April and from 0.05 to 0.31 for October using the 350 hPa level (which has similar, but slightly higher correlations to the 850 hPa level so is comparable to the 500 hPa level used here). This may reflect a more realistic set of emissions. Consistent with the seasonal differences in spatial correlations, analysis of the autocorrelation shows that while the models have distributions that are quite realistic during April, both models have CO distributions that are

substantially too smooth during October. Autocorrelations in the E2-GFED3 simulations are nearly identical, suggesting that overly smooth October field is not due to the use of climatological biomass burning emissions. Hence horizontal mixing may be too rapid in the tropical mid-troposphere.

3.4 Hydroxyl (OH)

The abundance of hydroxyl radicals is the key indicator of the oxidizing capacity of the troposphere. As very few direct measurements of OH have been made, and global observations do not exist, rather than examining OH directly we evaluate it via the lifetime of methane. As methane concentrations are prescribed at observed values, and oxidation by OH is the only chemical loss for methane, its tropospheric chemical lifetime is a nearly direct proxy for OH abundance.

The present day GISS model value for the full residence time of methane is 9.0–9.4 yr (range of 2006 values in the four RCPs), while it is 9.6–10.0 yr for the tropospheric chemical lifetime (Voulgarakis et al. (2012) report a slightly higher value, 10.6 yr, for the period centered around 2000 from the portion of the simulations with fixed methane concentration and using a different tropopause definition). The atmospheric lifetime is in good agreement with constraints derived from observations, which yield a value of 9.1 ± 0.9 yr (Prather et al., 2012). Thus at the global scale, oxidation capacity seems realistic in the model. This suggests that the upper tropospheric CO biases reported above more likely reflect a problem in the lumped hydrocarbon chemistry than with modeled OH, though biases in the vertical profile of OH could also play a role. That is more difficult to evaluate, however.

4 Historical trends and variability

We performed multiple realizations of historical simulations starting in 1850 and going to 2005 driven by varying observed forcings including natural effects (volcanic aerosols, solar variability, orbital variation) and anthropogenic changes in emissions, land use and concentrations of well-mixed greenhouse gases. Each simulation was started from a different point in a long control run.

4.1 Long-term ozone changes

We compare long-term trends in column and stratospheric ozone with available data extending back several decades. TOMS instruments have been flown since 1979, while some limited earlier satellite data has been used to estimate total ozone columns even further back in the ERA reanalysis (a combination of model and observations). We compare the ozone column trends simulated in E2-R with both the longer-term estimate (using ERA-interim (Dee et al., 2011) for the recent period and ERA-40 (Uppala et al., 2005) for the earlier time) and the trends seen by TOMS (Fig. 11). Ozone losses at Northern middle to high latitudes are quite similar in mag-

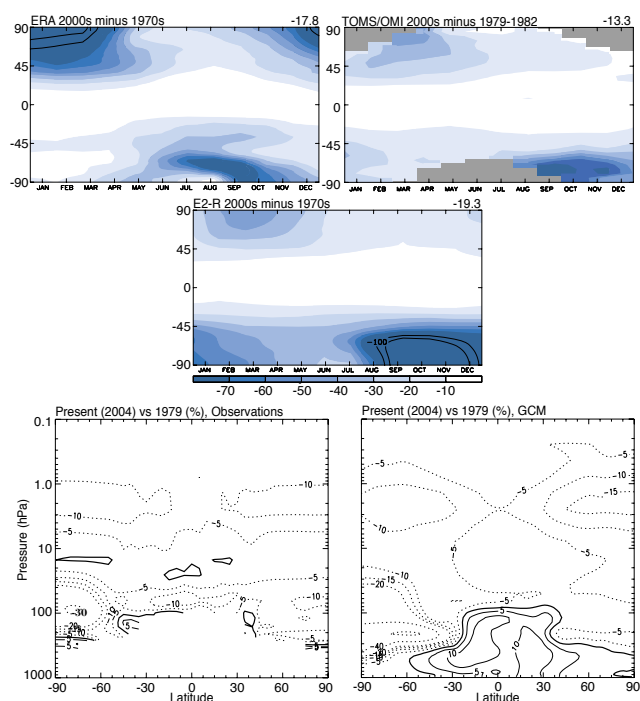


Fig. 11. Ozone trends in total column (top three panels; DU) and zonal mean (lower two panels; %).

nitude and seasonality to those seen in the TOMS record, though they are substantially less than those in the ERA estimate. At Southern middle to high latitudes, modeled depletion is again similar in magnitude to TOMS, but it extends too far equatorward and lasts too long into the austral summer. These biases in the simulated Southern Hemisphere trends are similar to those in the present-day climatology (Fig. 1). An analysis of 1980–2000 ozone trends in the CMIP5 models shows that the GISS-E2 models are among the best at matching the observed global trends, which are underpredicted in most models, but that the GISS models overestimate ozone decreases in some regions, particularly 35–60° S (Eyring et al., 2012).

We also compare modeled trends with observations of zonal mean stratosphere ozone changes (Randel and Wu, 1999). Losses throughout most of the stratosphere are captured well in spatial pattern and magnitude (Fig. 11), although Antarctic depletion is overestimated and extends too far from the pole (as in the column analysis, and consistent with the jet being located too far equatorward) and too far towards the surface. The overestimate of ozone losses around 200 hPa despite evidence of too little stratosphere-troposphere ozone flux in the SH suggests that local chemistry is too vigorous, consistent with a cold bias in the E2-R model relative to ERA-interim that maximizes at 60° S, 180 hPa. The secondary loss maxima in the upper stratosphere at high latitudes in both hemispheres are well captured.

Table 3. Surface ozone trends (ppbv yr⁻¹).

Station name	Latitude	Longitude	Altitude (m)	Time Period	Observed trend	E2-R trend*
Barrow	71.32° N	156.6° W	11	Mar 1973–Dec 2005	0.04	0.05 to 0.07
Arkona/Zingst	54.4° N	13.26° E	42	Sep 1956–Jun 2004	0.36	0.22 to 0.26
Mace Head	53.33° N	9.9° W	25	Nov 1987–Sep 2006	0.18	0.04 to 0.17
Hohenpeissenberg	47.89° N	11.02° E	985	Jan 1995–Dec 2007	0.32	0.01 to 0.10
Zugspitze	47.42° N	10.98° E	2960	Jan 1995–Dec 2002	0.40	0.26 to 0.46
Mauna Loa	19.54° N	155.58° W	3397	Sep 1973–Dec 2004	0.15	0.16 to 0.20
Samoa	14.24° S	170.57° W	42	Sep 1975–Dec 2004	-0.03	0.08 to 0.11
South Pole	89.90° S	24.8° W	2810	Jan 1975–Dec 2004	-0.05	-0.03 to -0.07

* Model trends are given as the range over five ensemble members.

Though long-term measurements of tropospheric ozone are quite limited, several locations have surface observations extending back to the 1970s. We compare trends in the model with the small number of stations reporting surface ozone for at least 20 yr (Derwent et al., 2007; Oltmans et al., 2006) as in the analysis of Lamarque et al. (2010) (Table 3). For the six sites in the NH, the model either captures the observed increase fairly well or underestimates. Underestimates are large at Arkona and Hohenpeissenberg. Note that Hohenpeissenberg and Zugspitze are within the same model grid box (at different altitudes), and the much lower trend at low elevation may reflect the stronger influence of surrounding areas at that level, or perhaps a trend in STE that does not extend far enough down in the model. At Samoa, the model shows an increasing trend in surface ozone while the observations show a decrease. At the South Pole, the model captures the observed decreasing surface ozone trend. These results are similar to the GISS results shown in Lamarque et al. (2010), though trends are more positive at the two SH sites.

For comparison with other studies, we also analyzed the 1850 ozone budget. Again using the present-day 150 ppbv ozone contour to define the tropopause, the net chemical production is 428 Tg yr⁻¹, dry deposition is 788 Tg yr⁻¹, and STE is 360 Tg yr⁻¹. The chemistry and deposition terms are much lower than their present-day counterparts due to reduced precursor emissions, but the STE value is 68 Tg yr⁻¹ larger owing to the greater amount of stratospheric ozone prior to depletion by halocarbons. The 1850 tropospheric ozone burden using the chemical tropopause is 251 Tg, 71 % of the present-day value, while using the nominal 150 hPa tropopause it is 357 Tg, 81 % of the present-day value. Preindustrial lightning NO_x production is 7.1 Tg N yr⁻¹, only 0.2 Tg N yr⁻¹ less than the present-day amount. The tropospheric burden below the chemical tropopause was found to be 239 ± 22 Tg in a recent multi-model analysis (Young et al., 2013), similar to our 251 Tg. In that study, net chemical production and dry deposition were much lower in 1850, as in our results, while the change in STE ranged from minimal to ~ 50 % greater during 1850, qualitatively consistent with our findings.

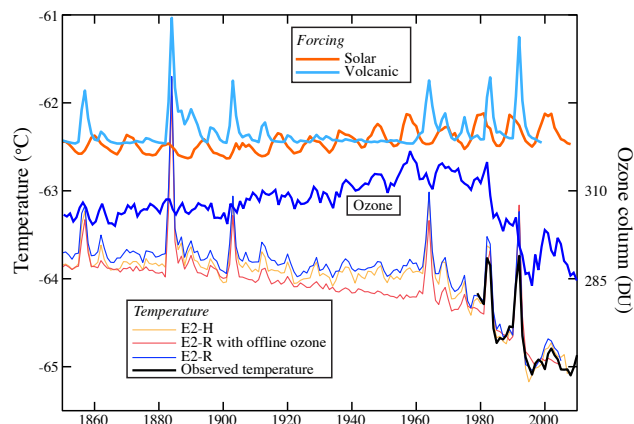


Fig. 12. Comparison of MSU lower stratospheric temperatures in model simulations and observations. Historical total column ozone (Dobson units; DU) and solar and volcanic forcing timeseries are shown for comparison (the vertical scale is arbitrary for forcings).

4.2 Ozone variability

In addition to responding to composition changes, variability in ozone on decadal to interannual timescales has both an unforced component and a component driven by variations in solar and volcanic forcing. The response to volcanism is seen clearly in the ~10–15 DU decrease in total column ozone following both the eruptions of El Chichon (1982) and Mt. Pinatubo (1991) (Fig. 12). The response to solar forcing is most clearly isolated during the 1936–1960 period when there was little volcanism and only weak long-term trends. During this time, ozone variations are ~3–5 DU. The magnitude of both the ozone response to solar and volcanic forcings are in reasonable agreement with observations (Chandra et al., 1999; Randel et al., 1995).

The combined influence of volcanic aerosols, solar flux changes and ozone changes causes stratospheric temperature variations. The long, observed Microwave Sounding Unit (MSU) temperature timeseries in the lower stratosphere is captured very well in the GISS-E2 interactive chemistry

simulations, both the multi-decadal trend due to ozone-depleting substances and shorter term variations (Fig. 12). Note the clear differences in temperature variability during the 1936–1960 period when only solar forcing was substantial in the interactive chemistry runs as compared with simulations using offline decadal mean ozone (physics-version = 1 in the CMIP5 archive). This shows that use of coupled ozone may be quite important in capturing realistic stratospheric temperature variability.

We examined variability by calculating standard deviations of monthly means during a decade using all five ensemble members (i.e. 50 yr per month) and using 1997–2005 observations (i.e. 9 years per month). Variability of polar stratospheric ozone depletion is generally captured reasonably well in the model. In the Arctic, variability is confined to high latitudes during winter through April, extending to mid-latitudes in May and June as polar loss is mixed to lower latitudes. The amount of depletion over the whole polar area varies in the springtime Arctic, with significant wave structure and much stronger amplitude in the E2-R model than in E2-H (Fig. 13), with the former model showing better agreement with observations. In the Southern Hemisphere, model variability is largest at the edge of the vortex during September and October (Fig. 13), extending over the whole polar cap by November when the vortex breakup typically occurs. As discussed previously, this breakup of the vortex takes place too late in our model, so that the observed variability extends over the whole polar cap by October. This timing is captured well in the E2-MERRA simulation.

The large interannual variation in the extent of the Antarctic ozone hole is well captured, consistent with year-to-year variations in meteorology influencing whether temperatures are below PSC formation thresholds at the edges of the Antarctic polar vortex (as interior values are typically well below). The broader extent of variability in the equivalent season in the Arctic occurs because temperatures are typically near the PSC threshold values. It is interesting that atmosphere-ocean interactions appear to play such a large role in Arctic interannual variability in the stratosphere, as evidenced by the difference between the E2-R and E2-H simulations (Fig. 13). This suggests that ocean conditions are affecting planetary waves, which play a larger role in Arctic than Antarctic dynamical variability.

5 Radiative forcing

We performed several types of radiative forcing calculations. We calculated instantaneous forcing (iRF) at the tropopause in which no part of the climate system is allowed to respond to the forcing. These values are averaged over time during a simulation in which radiative transfer is calculated twice at each point in time with the only difference in the two calculations being the ozone field. Calculations of iRF were performed online during the transient climate change simula-

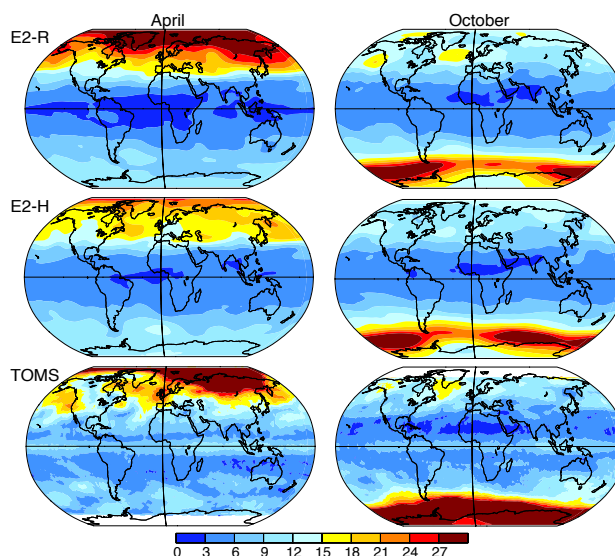


Fig. 13. Standard deviation of total column ozone (DU) in April (left) and October (right) in the E2-R (top row), E2-H (2nd row) and TOMS observations (bottom row). Models are 2000–2009, while TOMS data is from 1997–2005.

tions, and were also performed in separate simulations using prescribed ozone changes and constant background atmospheric conditions.

We also calculated forcing allowing stratospheric temperatures, but nothing else, to adjust to the ozone changes, which we call RF following the standard definition (Forster et al., 2007). These RF values were calculated in offline computations using the GISS radiative transfer code, prescribed tropospheric conditions (temperature, water vapor, surface conditions, long-lived greenhouse gases and aerosols), and various input ozone fields, with an assumption of fixed dynamical-heating in the stratosphere. Ozone forcing is the difference between the two radiative transfer calculations with only the ozone altered. Finally, we performed calculations of the effective radiative forcing (ERF) due to ozone which is defined here as the radiative flux perturbation when SSTs and sea ice are held fixed but all other adjustments are allowed in response to the imposed ozone changes. This forcing thus includes changes such as those in water vapor, lapse rate and cloud cover induced by the ozone changes. There is thus substantially more internal noise in these simulations, so that longer runs are required to isolate forcings. Hence these fixed-SST runs were performed for 40 yr (achieving a global mean 2-sigma uncertainty of 0.07 W m^{-2} , with a longer 70 yr run to diagnose the stratospheric forcing, for which the uncertainty was reduced to 0.06 W m^{-2}). In contrast, uncertainties in iRF and RF calculations due to interannual meteorological variability are much less than 0.01 W m^{-2} .

Calculations of RF due to GISS model ozone changes were also performed using the NCAR radiative transfer model (RTM). Not only does this help test the sensitivity to

the particular RTM, but those calculations used the preindustrial (1850) ozone 150 ppb contour to define tropospheric ozone, and hence they also test the sensitivity to the choice of troposphere/stratosphere ozone boundary. Calculations were performed with the CCSM4 RTM (see description and references in <http://www.cesm.ucar.edu/models/ccsm4.0/cam/>) in an offline setting (Conley et al., 2012) analogous to the GISS RTM calculations, except that the location of the tropopause (where flux changes are reported) is based on the NCAR/NCEP reanalysis tropopause pressure while the GISS RTM used that model's meteorological tropopause.

5.1 Present-day influence of ozone: model versus TES

To characterize potential biases in our RF calculations, we compare the present-day tropospheric ozone radiative forcing with that calculated using the observed ozone climatology from TES. We calculate the RF using either TES or modeled ozone below the model's tropopause and evaluate the difference between these. Hence these calculations do not diagnose forcing relative to preindustrial or zero ozone, but rather the forcing bias due to model/TES ozone differences. We also calculate iRF using modeled and TES ozone below the tropopause to provide a basis for comparison with independent TES-based RF calculations, which are instantaneous. The global mean tropospheric ozone iRF using the model's ozone is extremely close to that using TES', with the model value lower by 0.022 W m^{-2} at the tropopause (0.024 W m^{-2} longwave (LW) and -0.003 W m^{-2} shortwave (SW)), with a similar value at the top-of-the-atmosphere (TOA) where the model value is lower by 0.028 W m^{-2} (0.019 W m^{-2} LW, 0.009 W m^{-2} SW). A similar calculation of model-tes differences using the TES instantaneous radiative kernels (Worden et al., 2011) finds an even smaller value of 0.006 W m^{-2} for LW at the TOA (Bowman et al., 2012). For comparison, the tropopause iRF due to all present-day tropospheric ozone (relative to zero) is 1.826 W m^{-2} , hence the global mean bias is $\sim 1\%$.

For RF, model-tes differences are even smaller, with a difference of -0.016 W m^{-2} at either the tropopause or TOA (Fig. 14). This good global mean agreement clearly results from a cancellation of positive biases in the model in the NH (0.10 W m^{-2} RF) and negative biases in the SH (-0.13 W m^{-2} RF). These large-scale biases are consistent with the zonal mean tropospheric ozone differences shown in Fig. 3, and they follow differences in the tropospheric column fairly closely (though with latitudinal dependence due to available sunlight). Modeled long-term surface ozone trends in the NH are generally about right or underestimated, as discussed in Sect. 4.1, suggesting that the positive bias there has a substantial systematic component rather than resulting purely from overestimated tropospheric ozone increases. Negative biases in the SH are likely related to both overestimated downward flux of stratospheric air there

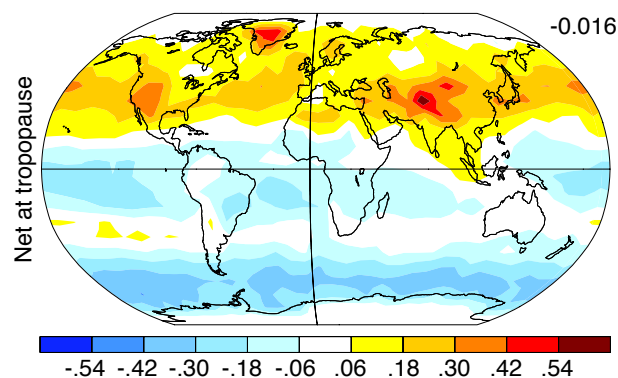


Fig. 14. Difference in tropopause RF (W m^{-2}) calculated using modeled (E2-R) and TES tropospheric ozone. The global mean value is in the upper right.

and potentially overestimated Antarctic ozone depletion (e.g. Fig. 11). Note that the total TOA LW iRF due to tropospheric ozone (relative to zero ozone) is 0.71 W m^{-2} , while the more climate-relevant net tropopause iRF in this case is 1.83 W m^{-2} . Such large discrepancies between TOA and tropopause fluxes are not seen when RF is used, however. This highlights how iRF can be a poor indicator of RF, and suggests that to obtain the net tropopause RF, TES estimates of TOA LW ozone iRF (e.g. Aghedo et al., 2011b; Worden et al., 2008) need to be adjusted down by $\sim 16\%$ ((tropopause net RF bias)/(TOA LW iRF bias) = $0.016/0.019$).

As a percentage of the iRF due to all present-day tropospheric ozone (relative to zero), the NH biases are typically $+5\text{--}15\%$, tropical biases are less than 10% everywhere except for the pronounced low bias over the central Pacific, and the biases over the Southern Ocean are typically $30\text{--}50\%$. Though comparison with TES ozone indicates biases, the model does not always show such biases in comparison with sonde data. For example, the NH mid-latitude positive bias in the model is particularly large over the Western US. Yet in comparison with the sonde climatology from Boulder, CO, the model shows very little bias at most levels. In particular, biases in the middle and upper troposphere, where forcing efficiency is greatest, are 0 ppb at 200 hPa , $+16 \text{ ppb}$ at 300 hPa and $+11 \text{ ppb}$ at 500 hPa , suggesting that the model's ozone forcing at this location is unlikely to be greatly overestimated. Similarly at the Syowa station on the Antarctic coast, the model is quite low compared with TES at 300 hPa but is actually a bit too large compared with the sonde measurements. At 200 hPa the model appears to have a genuine low bias during austral spring and summer.

5.2 Historical ozone forcing

We have evaluated the radiative forcing due to preindustrial (1850) to present-day ozone changes in the E2-R model using both iRF and RF. Calculations of iRF were performed

Table 4. Global mean annual average 1850 to 2000 ozone radiative forcings (W m^{-2}).

	iRF: transient* (± 0.03)	iRF: prescribed ozone* (± 0.02)	RF (± 0.07)	ERF (± 0.07)	AR4 iRF	AR4 RF	AR4 ERF
Whole atmosphere ozone	0.28	0.22	0.17 (0.17)	0.22	0.44	0.28	0.26
Tropospheric ozone	0.22	0.17	0.26 (0.30)	0.24	0.41	0.34	0.34
Stratospheric ozone	0.06	0.05	-0.09 (-0.13)	0.00	0.03	-0.06	-0.08

iRF and RF values are given at the tropopause and ERF values are given at the top-of-the-atmosphere. RF values are given using the offline GISS RTM and in parentheses using the offline NCAR RTM with their respective tropopause definitions (see text). Uncertainties for iRF and ERF calculations are from the year-to-year variability in meteorology (runs are 10 yr or more for iRF, 40 yr or more for ERF) and are given in parentheses below the forcing name (95 % confidence). There is no variability in offline RF calculations.

* Transient iRF values are based on reruns of 1850s and 1995–2004 transients, while prescribed ozone iRF values are from a present-day simulation using offline ozone fields from the transient runs.

online during the E2-R transient simulations, providing a time-varying estimate of iRF due to ozone changes throughout the atmosphere. These calculations also include any changes in ozone forcing that occur due to differences in background conditions, such as temperature, cloud cover and water vapor concentration, however. A second set of iRF calculations was performed rerunning the transient simulations during the 1850s and from 1995–2004 and using either the whole atmosphere ozone change or only the ozone changes below the tropopause to diagnose the iRF due to tropospheric ozone, with the iRF due to stratospheric ozone calculated as the difference between the whole atmosphere and tropospheric ozone iRFs. A third set of iRF calculations used a more common setup in which ozone fields from the transient runs averaged over the 1850s and 1995–2004 were used as input to an atmosphere-only simulation with non-interactive chemistry in which the double call to the radiation code used the offline preindustrial and present-day ozone fields. In these prescribed ozone simulations, ozone either throughout the atmosphere or only below the tropopause was set to preindustrial values in an initial radiation call, while ozone and all other conditions were set to present-day for all other purposes (i.e. this removes any influence of changes in the background state of the atmosphere). All calculations use the meteorological tropopause.

The transient and prescribed ozone iRF calculations thus differ in two main ways: the background state of the atmosphere varies with time in the former (e.g. temperature, pressure, other trace gases) while constant present-day conditions for all background variables are used in the latter (note that the tropopause itself moves in the transient case as the temperature profile changes, but it does not in the prescribed case) and the atmosphere is coupled to a dynamic ocean in the transient while prescribed ocean conditions are used in the offline calculations, which alters the base distribution of relevant fields such as clouds. To separate these, another calculation using offline archived ozone fields and also including background state changes (but still with prescribed ocean conditions) was performed which finds a whole atmosphere iRF of 0.29 W m^{-2} , consistent with the online transient results (Table 4). This indicates that changes in background

state account for the differences in the iRF values, adding $\sim 0.05 \text{ W m}^{-2}$ (23 %) relative to constant background conditions. While much of the change in the background state of the troposphere is not due to ozone, changes in ozone concentration in the vicinity of the tropopause can have a large impact on temperatures in the upper troposphere/lower stratosphere, so variations in the height of the tropopause may in fact be largely due to ozone trends. Thus it is unclear which type of ozone iRF calculation provides a more reasonable assessment of ozone's influence on Earth's energy balance.

The offline calculation of stratospherically-adjusted forcing using GISS ozone and the GISS RTM shows a tropospheric ozone RF of 0.26 W m^{-2} (Table 4), a value slightly lower than our prior estimates and most of those in the literature (Forster et al., 2007). Stratospheric ozone RF is -0.09 W m^{-2} , well within the range of previous estimates. Calculations separating the tropospheric and stratospheric contributions using the NCAR RTM show stronger forcings in both regions, but the whole atmosphere ozone RF calculated by the two methods gives identical results of 0.17 W m^{-2} . Separation of ozone forcing by altitude region is thus less robust than the whole atmosphere forcing, and does not provide a good indication of the role of different drivers (Shindell et al., 2013). Note that for consistency, the same tropopause is used at 1850 and 2000 in both methods so that the RF calculations are not biased by changes in the volume of the troposphere itself. For 1970 versus 1850, before substantial stratospheric ozone depletion, the tropospheric ozone RF calculated using the NCAR RTM is 0.23 W m^{-2} , while the GISS RTM whole atmosphere value is slightly lower at 0.20 W m^{-2} . Hence there is some sensitivity to the RTM and/or the definition of the tropopause, with differences on the order of 0.04 W m^{-2} or less.

While the RF calculation requires offline RTM simulations, temporal variations can be examined using the iRF, keeping in mind that the magnitude is not as good a predictor of climate response. The whole atmosphere ozone iRF shows a gradual increase from about 1860 to the early 20th century, with the most rapid growth from about 1950 to 1980 (Fig. 15). Net ozone iRF through 1970 was almost

entirely due to change in LW forcing resulting from tropospheric ozone increases. Starting in the 1970s, a clear increase in SW forcing began while longwave forcing began to decline. These changes are coincident with large decreases in total column ozone due to stratospheric depletion. iRF remained fairly constant from 1980–2004, as increases in tropospheric ozone, the dominant contributor to the whole atmosphere ozone forcing, slowed due to the impact of pollution controls while the stratospheric ozone-induced trends in LW and SW radiation largely offset one another. Variability increased during this latter period, however, with strong interannual changes associated with volcanic eruptions that deplete stratospheric ozone in the presence of large chlorine and bromine abundances. While the model's total ozone column abundance appears to be high compared with the early data in the ERA reanalysis (Fig. 15), we note that those times have few observations and that present day ozone columns are lower than the present-day TES data (Fig. 5). We utilize ERA ozone simply to show the long-term trends, and the comparison with the model shows that the model exhibits roughly the correct timing and magnitude of ozone depletion. This is true both with regards to the longer-term gradual reduction due to halocarbons over the past several decades and the short-term response to large volcanic eruptions (El Chichon in 1982 and Mt. Pinatubo in 1991). The model does not show the large excursion around 2005, but this appears to be unforced variability.

For RF, values at the tropopause and at the TOA are generally about the same (within 0.01 W m^{-2}) for cases where ozone changes are in the troposphere, although the SW and LW components can vary markedly. TOA and tropopause RF for whole atmosphere ozone differ, however, as stratospheric ozone changes can push RF at these two levels in opposite directions (implying non-zero heat flux across the tropopause). Similarly, the iRF for prescribed offline ozone is -0.31 W m^{-2} for the whole atmosphere, and 0.08 W m^{-2} for the troposphere, values quite different than the tropopause values shown in Table 4. We utilize tropopause forcing as a better indicator of the forcing applied to the lower atmosphere where climate response takes place.

In 1970, the whole-atmosphere ozone RF is 0.20 W m^{-2} , 94 % of the iRF (both using the GISS RTM). We take this to be a representative RF/iRF ratio for ozone forcing driven primarily by tropospheric ozone changes. In 2005, this ratio decreases to 61 % as the RF drops to 0.17 W m^{-2} , consistent with our understanding that iRF provides a relatively poor indicator of the forcing due to stratospheric ozone changes. We estimate the tropospheric ozone RF using the change in the ozone burden throughout the troposphere (below the 150 ppb ozone contour) times the normalized radiative forcing (NRF) where the latter is defined as the RF per unit burden change (Fig. 15). We determine NRF based on the ozone change during 1901 to 1970, when the signal is large enough to be seen clearly but before stratospheric ozone changes became large, and find that the NRF is $0.0034 \pm 0.0006 \text{ W m}^{-2}$ per Tg (us-

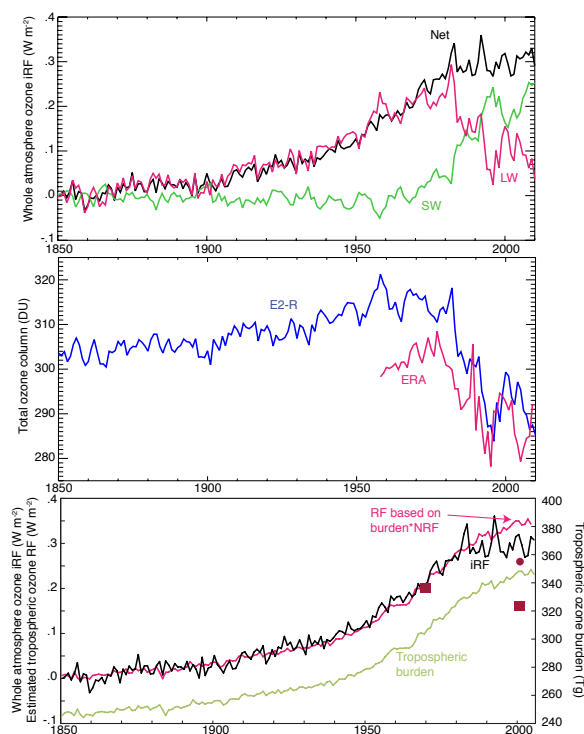


Fig. 15. Historical evolution of ozone burdens and radiative forcing. Time varying instantaneous tropopause radiative forcing due to ozone (W m^{-2}) from the transient simulations (top; as longwave (LW), shortwave (SW) and net), total column ozone (DU) in the model and in ERA reanalysis (middle), and iRF along with the tropospheric ozone burden (DU) and ozone RF estimated using the tropospheric burden times the normalized radiative forcing (NRF) (bottom). The solid red symbols in the lower panel show the whole atmospheric (square) and tropospheric (circle) ozone RF calculated in the offline radiative transfer model.

ing the $\text{RF}_{1970}/\text{iRF}_{1970}$ to convert from iRF to RF). The estimated 2005 RF using this method is 0.34 W m^{-2} . This estimate is somewhat larger than the calculated tropospheric RF for 2000, but overall the estimate based on tropospheric burden*NRF appears to provide a reasonably good indication of tropospheric ozone RF. In contrast, the lack of an increase in iRF from 1970 to 2000 reflects the influence of stratospheric ozone losses, and the difference between the NRF-based estimate and the iRF is similar to the RF from stratospheric ozone.

The change in the ozone forcing trend over the last 30 yr visible in the global mean is caused by a distinct spatial pattern of forcing. The evolution of zonal mean ozone RF over time shows that the rapid increase in ozone forcing after 1950 was mostly in the Northern Hemisphere during 1951–1975 (Fig. 16). During the last 30 yr forcing increased substantially both at Northern middle latitudes and throughout the

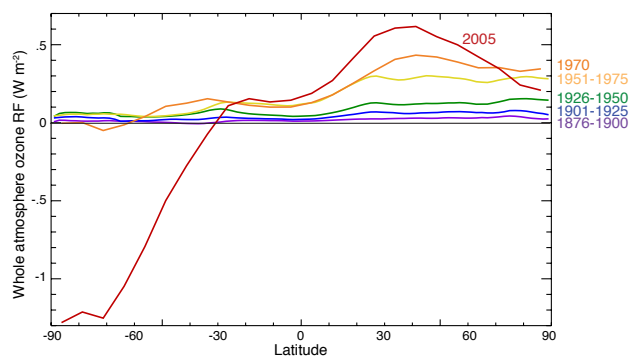


Fig. 16. Zonal mean ozone radiative forcing (W m^{-2}) for the indicated time periods relative to the 1850–1875 mean. Forcing values for 1951–1975 and earlier 25 yr periods are the transient $\text{iRF} \cdot \text{RF}_{1970}/\text{iRF}_{1970}$, while the values at 1970 and 2005 are from the offline RF radiative transfer calculations.

tropics while decreasing very sharply at Southern high latitudes. Northern high latitude RF decreased slightly during the latter quarter of the 20th century as some Arctic stratospheric ozone loss began to occur, while the large decreases in RF at Southern high latitudes are due to Antarctic ozone depletion during this period.

Examining the full spatial structure of the preindustrial to present-day forcing, we see that ozone changes caused substantial positive RF throughout most of the NH, with especially large values in the clear sky subtropical regions and at NH mid-latitudes (Fig. 17). In contrast, stratospheric ozone changes led to reductions in RF over Antarctica. Comparison of ERF and RF shows that the distribution of ERF is much more uneven. The LW ERF shows clear signs of shifts in circulation, with noticeable changes particularly in stormtrack regions related to cloud cover shifts. These are to some extent compensated for by changes in SW. The SW ERF also shows strong changes over the Southern Ocean, which are apparent in the stratospheric ozone ERF as well but not in the tropospheric ozone ERF. Hence these indicate a shift in the polar jet location in response to stratospheric ozone depletion, a well-documented phenomenon (e.g. Polvani et al., 2011). While prior work indicated that ERF is a better predictor of temperature response than iRF or RF (Hansen et al., 2005), clearly even with a fairly long simulation the inclusion of meteorological changes leads to a much noisier forcing field, and we note that the difference between ozone ERF and RF is not statistically significant at the 95 % confidence level.

5.3 Future ozone and methane forcing

5.3.1 Ozone forcing

Stratospheric ozone recovers from the late 20th century depletion in all four RCP scenarios (Fig. 18) as the effect of projected decreases in halogen-containing gases capable of

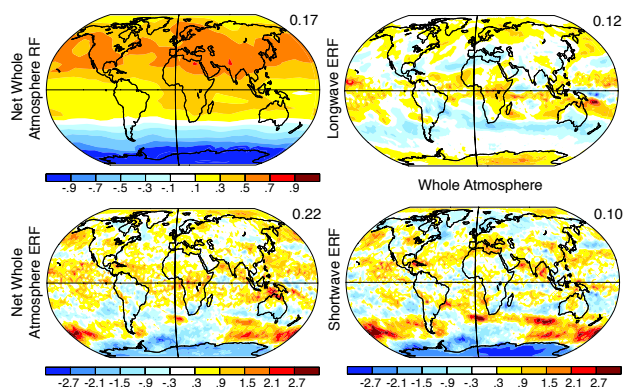


Fig. 17. Preindustrial to present-day ozone RF (top left) and ERF (lower left, right column) (W m^{-2}) due to the ozone changes in the E2-R simulations. Note difference in scale for RF and ERF. The ERF values are from 40-yr runs. Global mean values are given in the upper right of each panel.

reaching the stratosphere outweighs projected increases in N_2O . In RCP2.6 and 4.5, values return to a level only slightly less than their 1950–1970 average by 2100. In the RCP6.0 and especially the RCP8.5 scenario, there is a super-recovery with projected future values larger than those seen in the 1950–1970 period. This can be attributed to the larger greenhouse gas loading in those scenarios, which leads to greater cooling of the stratosphere and hence reduces the rates of chemical reactions that destroy stratospheric ozone outside the polar regions (the extrapolar dominates the total stratospheric burden). Tropospheric ozone remains fairly constant for several decades under RCP4.5 and RCP6.0, decreasing thereafter, while it decreases throughout the 21st century under RCP2.6. In contrast, the tropospheric burden increases throughout the 21st century under RCP8.5. These trends reflect the varying projections in emissions of tropospheric precursors. The large increase in tropospheric ozone in the RCP8.5 scenario results from increases in methane, as all other ozone precursors emissions decrease in all the RCP scenarios. The decreases are largest in the RCP2.6 scenario, leading to the lowest tropospheric ozone burden in that scenario.

Tropospheric ozone RF, again using tropospheric burden \times NRF, under the four RCPs shows a diverse range of trends (Fig. 18). Forcing increase by about 0.15 W m^{-2} under RCP8.5, while decreasing by about 0.08 to 0.18 W m^{-2} in the other RCPs. The difference between the scenarios' ozone forcing trends are qualitatively similar to the results from the CAM model, though we find a larger decrease in tropospheric ozone forcing under RCP4.5 (Cionni et al., 2011). Kawase et al. (2011) also shows qualitatively similar behavior of tropospheric column ozone under the 4 RCPs, though again there are some differences in the intermediate RCPs with weaker changes in their model for RCPs 4.5 and 6.0. Stratospheric ozone forcing would return to near-zero during

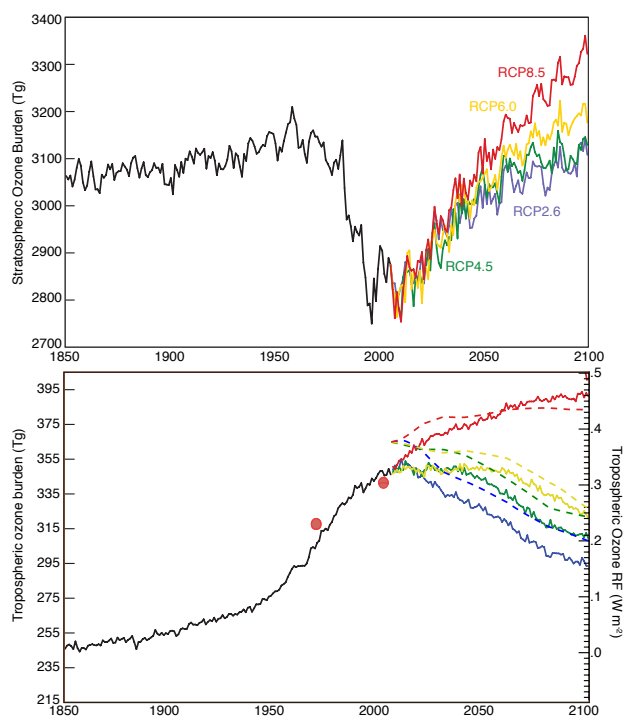


Fig. 18. Stratospheric ozone burden (top) and tropospheric ozone burden based on the 150 ppb ozone chemical tropopause along with the ozone RF estimated using the burden times NRF (bottom, solid lines). Also shown are the tropospheric ozone forcings estimated for the four RCPs using the MAGICC model (dashed lines). The red solid circles show tropospheric ozone RF using the NCAR RTM and 150 ppb ozone chemical tropopause.

the 21st century as ozone levels return to near their pre-industrial values (though there could be a small forcing due to the super-recovery under RCP8.5).

The tropospheric ozone forcing trends found here are also fairly similar to those estimated using the reduced-form model MAGICC (Meinshausen et al., 2011) (Fig. 18). Accounting for the offset in present-day tropospheric ozone forcing, the results are nearly identical for RCPs 2.6, 4.5 and 6.0. The increased forcing under RCP8.5 is substantially less ($\sim 0.08 \text{ W m}^{-2}$) in the MAGICC results, however.

High-latitude ozone recovery is fairly similar under all the scenarios (Fig. 19). The Antarctic recovers to approximately its pre-1950 values late in the 21st century in RCPs 2.6, 4.5 and 6.0, with a super-recovery in RCP 8.5. In the Arctic, a super-recovery occurs in all scenarios to some extent, and recovery begins more rapidly than in the Antarctic, where there is little change prior to 2020. Increases in transport contribute to the super-recoveries in both polar regions. We note that the September–November ozone depletion within the Antarctic 60–90° S band over 1980–2000 is 29 %, in good accord with observations that also show depletion of ~ 30 % (Bodeker et al., 2005).

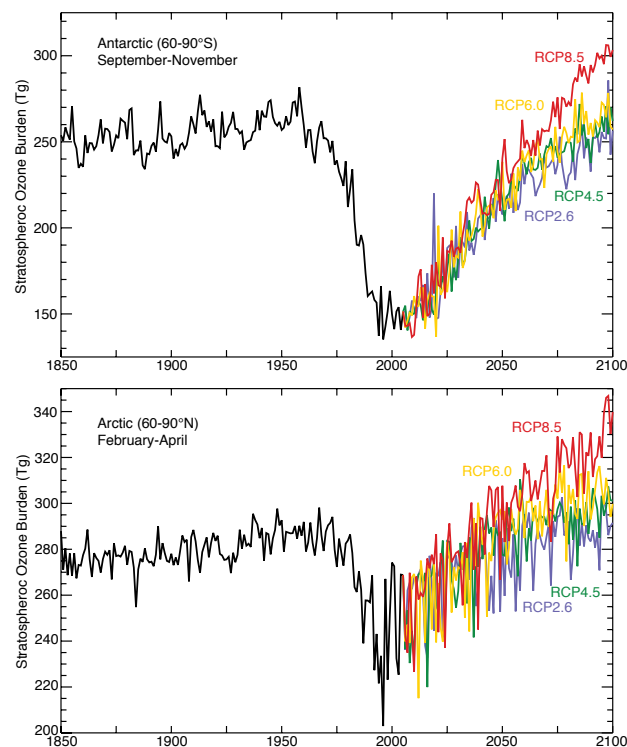


Fig. 19. High latitude stratospheric ozone burdens (Tg) during springtime.

5.3.2 Methane distribution

Since methane abundances in the future are calculated internally in our model, they are a function of changes in both anthropogenic and natural emissions and changes in oxidation rates. While anthropogenic emissions come from the RCPs, wetland emissions and oxidation rates are projected in the model simulations. This leads to a spatially and temporally varying distribution of methane in the model. We have compared the spatial distribution with satellite observations from SCIAMACHY (Schneising et al., 2011). The satellite observations cover 2003–2005. As the model values during this period were prescribed, we performed a separate run with the same setup as used in the RCP simulations but for repeating 2005 anthropogenic emissions. We analyze the last 5 yr of a 20-yr run, allowing the methane distribution to fully adjust to the spatial pattern of sources and sinks. The model is sampled using the same a priori methane and averaging kernel as in the satellite retrieval (Schneising et al., 2009, 2011).

Comparison shows that the model captures many of the features of the observed distribution, but that tropical emissions appear to be too low (Fig. 20). As the absolute amount in the model can be easily adjusted by altering the sources within their fairly large uncertainty ranges (Denman et al., 2007), we offset the model to match the global mean in SCIAMACHY when plotting the difference. This

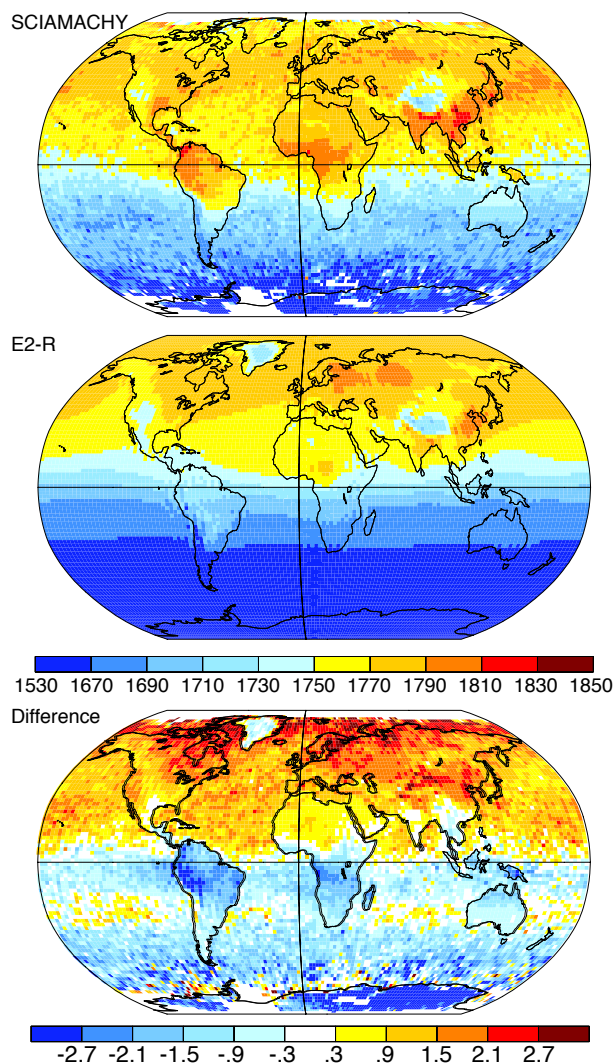


Fig. 20. Observed methane from the SCIAMACHY instrument over 2003–2005 (ppb; top) compared with modeled methane for 2005 conditions (ppb; middle) and difference between the two (percent; bottom). The top two panels show methane concentration averaged over the broad vertical weighting function of SCIAMACHY. In the difference map, the model has been offset by 1.8 % to make a global mean difference of zero in order to emphasize the spatial pattern.

small adjustment, less than 2 %, is smaller than the variation across the RCPs based on their present-day emissions of ozone precursors (see Sect. 5.3.3). Though there is a general tendency for the model to have too much methane at middle to high Northern latitudes and too little in the Southern Hemisphere, most of the biases are fairly localized suggesting that emissions, rather than atmospheric oxidation, are the cause. Emissions seem especially low over the Amazon and central Africa; they also appear low over Southeast Asia, Southern China, India and Texas/Louisiana. The tropical biases likely reflect too little emission from tropical wetlands, while the biases in Asia may be due to underestimated emissions from

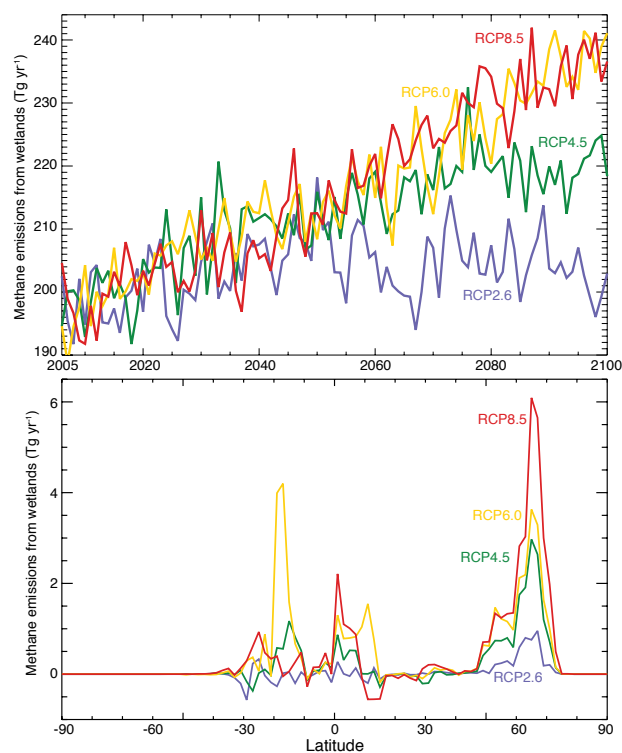


Fig. 21. Methane emissions from wetlands (Tg yr^{-1}) projected under the RCP scenarios (top; global total) and zonal mean change in methane emissions from wetlands between the 2090s and 2006–2014 (Tg yr^{-1} per 2 degree band; bottom).

rice cultivation and those in the US may reflect underestimated emissions from oil and gas-related activities. The low bias in the model over Antarctica is consistent with an overestimate of the downward flux from the stratosphere, which would mix air containing low methane values too low down in the atmosphere.

5.3.3 Methane emissions and oxidation

Wetland emissions projections are based on a model that calculates emission changes relative to present day based on anomalies in temperature and precipitation (Shindell et al., 2004). Changes in land-use do not affect wetland area in these simulations. Wetland methane emissions increase in all the RCPs, but especially in RCP6.0 and RCP8.5 where they are roughly 40 Tg yr^{-1} greater by 2100 relative to 2005 (Fig. 21). For RCP8.5, this is a relatively small contribution to emission trends as anthropogenic methane emissions under this scenario rise by 549 Tg yr^{-1} over this same period. For RCP6.0 and RCP4.5, however, the increased emissions from wetlands of about 40 and 25 Tg yr^{-1} are important as anthropogenic emissions are projected to decrease by 68 and 36 Tg yr^{-1} , respectively. Thus the increased wetland emissions offset roughly 60 and 70 % of the anthropogenic reductions achieved in these scenarios. Regionally, emissions

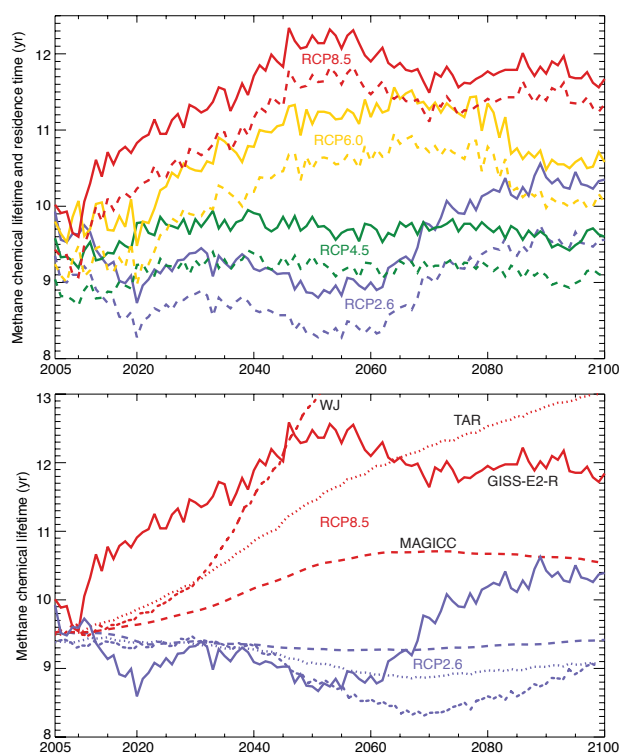


Fig. 22. Tropospheric methane chemical lifetimes (solid lines) and atmospheric residence times (dashed lines) in GISS simulations (upper panel), and tropospheric methane chemical lifetimes (thick solid lines) compared with lifetime estimates calculated using the IPCC TAR formulation (dotted lines), estimates from the MAGICC model used in creating the RCP methane concentrations (dashed lines) and estimates using the method of Wang and Jacob (WJ; long short dashes) (lower panel). Only two scenarios are shown in the lower panel for clarity.

in the Arctic increase by 11, 14 and 23 Tg yr⁻¹ going from the RCP4.5 to 6.0 to 8.5, with increases largely in Eurasia. In contrast, emissions from the tropics increase by 5 Tg yr⁻¹ in the RCP4.5 scenario, by 20 Tg yr⁻¹ in the RCP6.0 scenario, but by only 8 Tg yr⁻¹ in the RCP8.5 scenario. Changes in the tropics from 0–15° N are primarily due to changes in Africa, while the substantial changes between 20 and 30° S are largely driven by wetlands in South America. This region, south of the Amazon, both warms and receives more precipitation in the future projections, leading to enhanced methane emissions, while the Amazon warms but dries and hence shows little change in wetland emissions. Rainfall in the RCP8.5 scenario shifts strongly to the southwest, however, so that it eliminates some of the increase in South American wetland methane emissions seen in the RCP6.0 projection despite the warmer climate. Note that the comparison with SCIAMACHY suggested that our emissions from tropical wetlands are underestimated (Fig. 20), and hence the anomalies induced by climate change are likely to be similarly on the low side.

Methane's main sink, chemical oxidation, also changes with time in the RCP simulations. We examine changes using tropospheric chemical lifetimes and atmospheric residence times. Similar calculations in the literature are typically performed for steady-state conditions, but in our transient runs an adjustment to account for the methane growth rate is required (though this has only a modest effect on the results). Hence tropospheric chemical lifetime is defined as (tropospheric burden)/(tropospheric chemical loss – tropospheric growth rate × f) where f is the fraction of the methane chemical sink in the troposphere: $f = (\text{tropospheric chemical loss})/(\text{flux to stratosphere} + \text{tropospheric chemical loss})$. The atmospheric residence time is defined as (burden)/(chemical loss + soil sink – growth rate). Note that examination of global chemical lifetimes and tropospheric residence times shows very similar trends.

Trends in the methane residence time closely track trends in tropospheric chemical lifetime, indicating that these are not driven by variation in losses to the stratosphere (the loss to soils is fixed) (Fig. 22). The RCPs also have very similar wetland emissions out to about 2050, demonstrating that changes in natural sources are also not driving the trends in methane's lifetime. The trends in chemical lifetime follow OH, for which the mass-weighted tropospheric abundance decreases by 7, 2, 15 and 17% in RCPs 2.6, 4.5, 6.0 and 8.5, respectively, in 2100 relative to 2000 (using 3 yr averages at each time). Physical climate changes affect methane oxidation primarily by altering reaction rates and water vapor abundance. Comparing the OH in a simulation with 2100 climate and present-day emissions against OH in a present-day simulation Voulgarakis et al. (2012) found a change in methane chemical lifetime of –0.22 yr per C warming. Based on this relationship, climate changes should have decreased the methane chemical lifetime by –0.2 to –0.8 yr across the RCPs, while in fact the lifetime increased by 0.7 to 1.7 yr (7–16%) in the RCPs other than RCP6.0, under which it remained at its 2000 value. Hence the increases in methane lifetime must be driven by OH decreases resulting from changes in the emissions of ozone precursors, include methane itself, and overhead stratospheric ozone column (Voulgarakis et al., 2012), which more than offset the effect of a warmer climate (except in the case of RCP6.0, where these effects are approximately in balance).

Various reduced-form relationships have been used for estimating the response of OH to emission changes. We have calculated changes in methane's tropospheric chemical lifetime based on the IPCC Third Assessment Report formulation (Prather et al., 2001) for comparison with the actual time-evolution of changes (Fig. 22). While there is good agreement for the first 40 yr of the RCP2.6 scenario, the parameterization generally provides a relatively poor estimate of the actual methane lifetime change in the model. We have also calculated lifetime trends based on the assumption that OH is proportional to $S_N/(S_C)^{3/2}$ where S_N is the source of NO_x and S_C is the source of CO + VOCs + CH₄, both in

moles yr^{-1} (Wang and Jacob, 1998). These also yield trends that do not resemble those in the full chemistry-climate simulations very closely. Finally, we show lifetime trends generated by the reduced-form MAGICC model based on the TAR emission sensitivities but also including an estimate of the OH response to temperature and water vapor (Meinshausen et al., 2011). These estimates were used in creating the RCP methane concentration projections. The MAGICC lifetime trends generally have a shape closer to those in the full chemistry-climate model than the TAR formula based estimates, suggesting that the addition of responses to temperature and water vapor in MAGICC has improved the results. However, there are substantial differences relative to the full model. In particular, the lifetime in the RCP8.5 and 6.0 scenarios is significantly longer in the full model over nearly the entire 21st century. Trends in the RCP4.5 scenario are smoother in MAGICC, but generally in rough agreement with those in the model, while the GISS model's RCP2.6 simulation shows a sharp upward trend in the latter half of the century not seen in MAGICC.

Though there is some ambiguity in how the published relationships between OH and VOC emissions apply to different mixtures of VOCs, the VOC emissions have a fairly small influence on OH trends. It may be that the various parameterizations are too simplified to accurately capture the response to changes in emissions in the RCPs, which include quite large changes in both the magnitude and location of emissions. Additionally, factors other than emissions, such as changes in photolysis rates due to changing overlying ozone or changes in tropospheric temperature and water vapor, some or all of which are not included in the parameterizations may play important roles. This topic is further analyzed in the ACCMIP project (Voulgarakis et al., 2012).

We note also that the nitrate model used in the E2-R simulations (Bauer et al., 2007) appears to overpredict nitrate aerosol concentrations, and biases in this model could also affect gaseous nitrogen species. We therefore performed a sensitivity study rerunning the RCP8.5 scenario for 40 yr while including a somewhat arbitrary removal of NH_3 transport within convective plumes which appears to remove the nitrate bias (which is primarily in the upper troposphere). The methane chemical lifetime was 5 % less in this version of the model, but the trends were very similar, with an increase of 1.37 yr in the sensitivity study versus 1.39 yr in the original simulation. Hence we do not believe this bias affects the methane projections.

5.3.4 Methane concentration and forcing

With fully interactive chemistry and wetland methane emissions responding to climate, we would expect the methane projected in our model to be different from the MAGICC estimates provided along with the RCPs, and indeed this is the case (Fig. 23). Differences are especially pronounced in the RCP8.5 and 6.0 simulations. The GISS values do not di-

verge substantially from the RCP concentrations for the 2.6, 4.5 and 6.0 cases until ~ 2030 – 2040 , but differences are substantial in the RCP8.5 case even in the 2020s.

The approximate contribution of increasing methane emissions from wetlands can be easily assessed by looking at the percentage increase in the total methane source due to the wetland emission growth, accounting for the feedback of methane on its own lifetime (Prather et al., 2001). The rise in methane from wetlands from 2005 to 2100 increased the 2005 emissions by 1, 5, 8 and 8 percent in the RCP 2.6, 4.5, 6.0 and 8.5 scenarios, respectively. Accounting for the methane/OH feedback in the GISS model at 2100, $\delta\ln[\text{OH}]/\delta\ln[\text{CH}_4] = -0.47$ (Voulgarakis et al., 2012), these would lead to increases in methane concentration of 16, 138, 234 and 225 ppb, respectively in the four scenarios. The actual changes were -373 , 70, 409 and 2689 ppb, indicating that increases in wetland methane emissions were important in the RCP4.5 and 6.0 scenarios, but played only a minor role for RCP2.6 and 8.5. Differences between our calculated concentrations and those estimates in the RCPs were 130, 252, 518 and 691 ppb for the RCP 2.6, 4.5, 6.0 and 8.5 scenarios, respectively, indicating that similarly the growth in wetland methane emissions was not the main driver of differences relative to RCPs (though it was a fairly large component in the RCP4.5 case). Instead, the main driver of those differences was the trends in chemical lifetime in our model relative to those estimated in MAGICC.

The higher methane concentrations in our simulations lead to a correspondingly larger methane RF relative to the RCP concentrations (Fig. 23). Forcing is calculated for both sets of concentrations using the formula provided in the IPCC TAR (Ramaswamy et al., 2001), which accounts for changes in N_2O . Methane RF in our interactive chemistry-climate simulations is 0.05, 0.09, 0.18 and 0.16 W m^{-2} greater than that calculated using the RCP concentrations for the 2.6, 4.5, 6.0 and 8.5 RCPs, respectively. Hence the other CMIP5 models, which all use the RCP methane concentrations, may have a systematic low bias in their methane radiative forcing.

6 Climate response

We performed transient climate simulations for the historical period (1850 to 2005) driven only by ozone changes. We ran the same GISS-E2-R atmosphere-ocean model, but without interactive composition. Instead, aerosols and other conditions (e.g. well-mixed greenhouse gases, solar irradiance, etc.) were fixed at 1850 values and monthly mean ozone fields from the fully interactive model described previously were input to the model as the only time-varying forcing. An ensemble of five simulations was performed. Output from continuations of the control simulation from the same starting points was subtracted to remove any long-term climate drift due to the very slow response of the deep ocean.

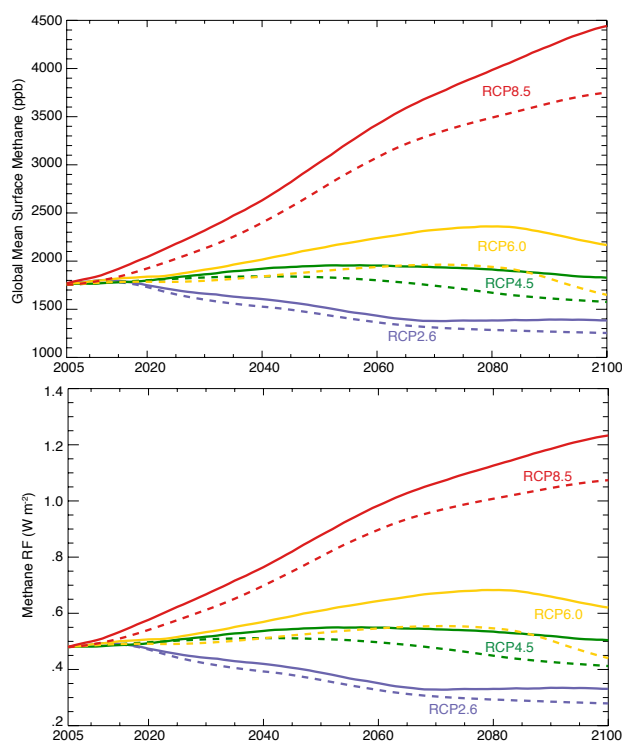


Fig. 23. Methane concentrations and forcing (W m^{-2}) relative to 1850. Concentration values (top panel) projected in the GISS model (solid lines) and the RCP values calculated with MAGICC (dashed lines), and the resulting radiative forcing (bottom panel).

The global mean annual average temperature shows gradual warming through the 1980s, after which temperatures decrease substantially (Fig. 24). This time evolution follows the RF (Fig. 15), while differing markedly from the $i\text{RF}$, reinforcing the notion that RF is the more useful metric for ozone changes (especially those in or near the stratosphere). The response to tropospheric ozone forcing from 1850 to 1970 is roughly proportional to the RF (following the model's transient climate sensitivity). The response to ozone during 1990 to 2010 is somewhat larger than would be expected given the small forcing decrease during that period (Fig. 15), though not markedly so. The model may be more sensitive to ozone forcing at higher altitudes. However, we note that the ozone forcing during this latter period changed more rapidly than during 1850 to 1970, and the model's net radiation at the TOA (or net heat at the surface) becomes more and more out of balance during 1990–2010 (Fig. 24). Hence the transient response is likely greater than the eventual equilibrium response to the higher altitude ozone forcing.

Examining the spatial pattern of forcing and response, we find little close correlation. In the broadest sense, ozone RF from 1850 to 1970 is largely in the Northern Hemisphere middle to high latitudes, and the zonal mean response is greatest there (Fig. 25). However, the response is also large over the Southern Ocean, where feedbacks appear to cause

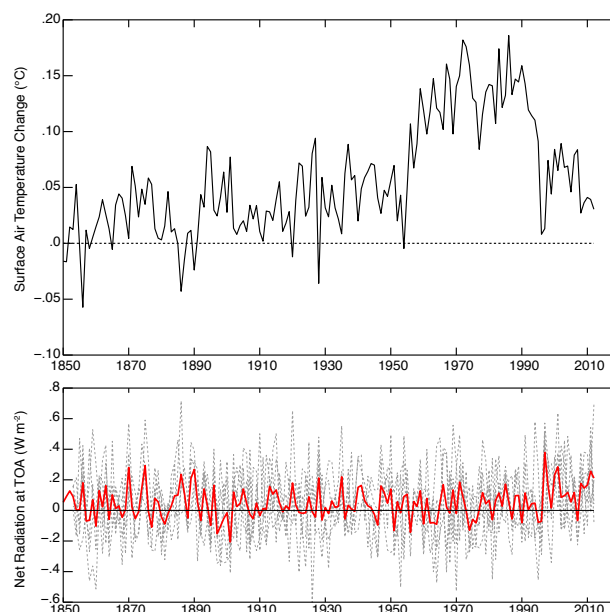


Fig. 24. Global mean annual average temperature change due to ozone forcing (top; ensemble mean) and net radiation at the top-of-the-atmosphere (TOA) in response to ozone forcing (red solid ensemble mean; black dashed individual members).

a local maximum in climate sensitivity. From 1970 to 2000, forcing remained positive over most of the tropics and NH, but large negative forcing occurred over SH middle to high latitudes. This negative forcing appears to be responsible for cooling over most of the Earth during this time in response to ozone forcing. Comparing the total 1850 to 2000 temperature response to ozone with various estimates of the forcing indicates that none of the metrics ($i\text{RF}$, RF or ERF) provides a good indication of the zonal mean responses, though ERF captures more of the meridional variations than the others, especially at SH middle to high and NH high latitudes. This suggests that at small spatial scales, regional patterns in climate sensitivity, e.g. those associated with changes in North Atlantic overturning, can dominate over regional patterns in forcing, especially for forcing by stratospheric ozone changes.

We have also compared the response with estimates of the zonal mean temperature responses averaged over broad latitude bands made using the Absolute Regional Temperature Potentials (ARTP) (Shindell, 2012). We find that the tropical and NH mid-latitude responses are in good agreement with the ARTP-based estimates for 1850–1970. This holds using the mean ARTP response coefficients calculated for CO_2 and sulfate and especially using those calculated specifically for tropospheric ozone (the latter give estimates within 10% of the actual response). The Arctic response is quite sensitive to the choice of coefficients, however, with the ARTP-based estimate being too large using the CO_2 and sulfate mean and too small using the ozone responses (though the latter have

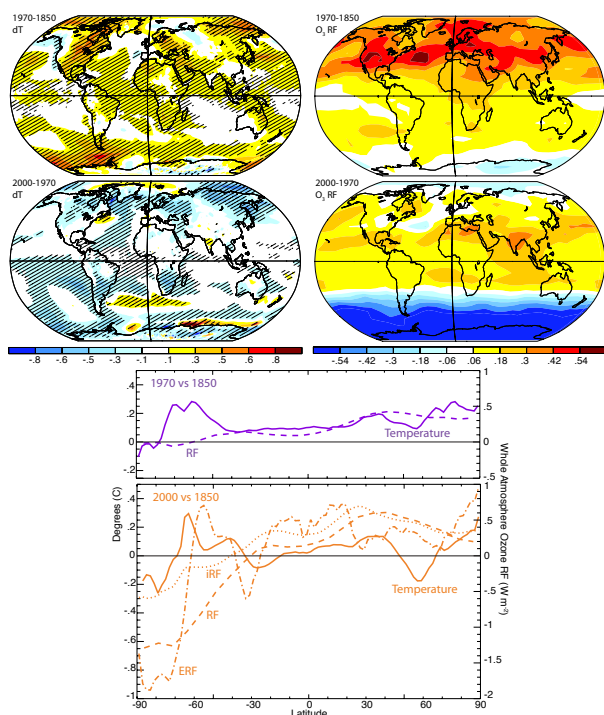


Fig. 25. Annual average surface temperature changes (C; left maps) and whole atmosphere ozone RF (W m^{-2} ; right maps) over the given time periods and zonal mean annual average surface temperature changes and ozone forcing for the indicated definitions of forcing (line plots). Hatching in the left column maps indicates 95 % statistical significance for the ensemble mean response relative to the control variability.

large uncertainties, Shindell and Faluvegi, 2009). The ARTP-based estimate of response in the SH extratropics is much lower than the actual response, however (calculated using the CO_2 and sulfate mean response coefficients as these were not calculated for ozone). The ARTP-based estimates do not correspond well with the actual response in any region for the 1970 to 2000 temperature changes, indicating that they are not suitable for estimating the effect of stratospheric ozone forcing. The results, however, suggest that the ARTP-based estimates for the response to worldwide tropospheric ozone forcing are robust in the tropics and NH mid-latitudes, consistent with the findings for aerosol forcing when evaluating the ARTP against other climate model results (Shindell, 2012).

Turning to precipitation, we see that for 1850 to 1970, the relatively larger warming of the NH suppressed the northward progression of the ITCZ during boreal summer (Fig. 26). Similarly, South Asian monsoon rainfall was pushed south. This behavior is opposite to the general view of thermally-driven changes that draw the ITCZ into a warm-

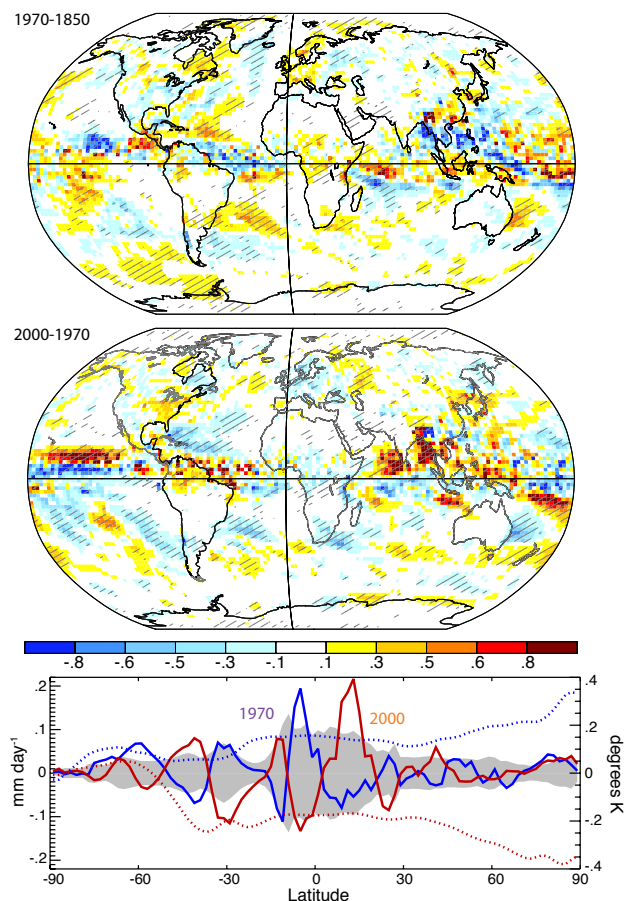


Fig. 26. June–August precipitation changes (mm day^{-1}) in response to ozone forcing comparing 1970 (1966–1975) with 1850 (1850–1859) and 2000 (1996–2005) with 1970. Hatching in the upper panels indicates 95 % statistical significance for the ensemble mean response relative to the control variability. Solid lines in the bottom panel show the zonal mean precipitation while dotted lines show zonal mean temperature changes at 200 hPa. The shaded area indicates the 2-sigma uncertainty range for precipitation changes.

ing hemisphere (Chiang and Bitz, 2005; Kang et al., 2009). While that mechanism has been demonstrated in responses to surface climate changes (e.g. from GHGs or aerosols), it appears that ozone may induce a different type of response by locally heating the atmosphere aloft. Indeed, the zonal mean precipitation response, both during 1850 to 1970 and during 1970 to 2000 when the shift is reversed, appears to follow temperature responses to ozone at around 200 hPa more closely than it does the surface temperature response (Fig. 26). This is especially noticeable during 1970 to 2000, when the 200 hPa temperature change is almost exactly opposite the 1850 to 1970 change but the extrapolar surface temperature response is much weaker than the 1970 vs. 1850 change. Note that changes in zonal mean temperatures are qualitatively similar at 500 hPa, but not as pronounced as those at 200 hPa (especially for 2000 vs. 1970).

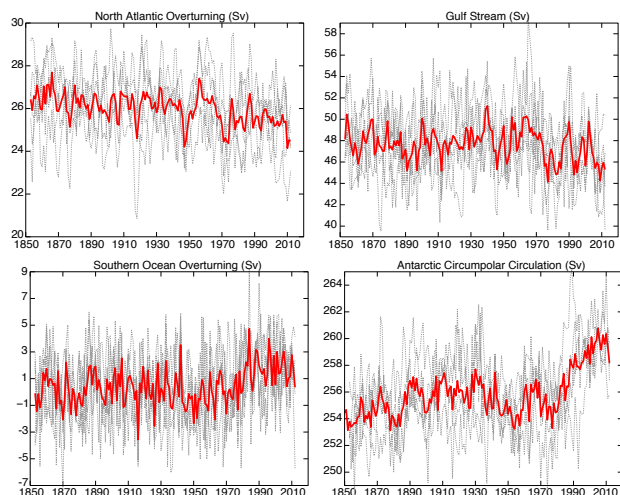


Fig. 27. Ocean and heat content trends due to ozone changes showing ensemble mean (red solid) and results from individual ensemble members (black dashed). The North Atlantic overturning circulation in the ocean is at 48° N, 900 m depth, the Gulf Stream is the maximum at $20\text{--}24$ or $30\text{--}33^{\circ}$ N, the Southern Ocean overturning is at 54° S, 3000 m depth, and the Antarctic Circumpolar Current is through the Drake Passage.

Circulation also changes in response to ozone forcing. While temperature changes in the NH over land were generally as large or larger than those over the oceans, changes in the SH were typically greatest over the oceans. The close agreement between the meridional structure of ERF and surface temperature response in the SH indicates that much of the enhanced sensitivity of temperature over the Southern Ocean relative to areas to the north or south stems from atmospheric adjustments to ozone forcing, such as the response of water vapor, lapse rate or clouds. While those atmospheric effects dominate during 1850 to 1970, substantial shifts in ocean circulation also play a role after that time. Large trends in the ocean are especially apparent after 1970 in the Antarctic circumpolar circulation, which increases markedly, and in Southern Ocean overturning, which also increases (Fig. 27). At the same time, there is a significant slowdown in the Gulf Stream (~ 2 Sv) and in North Atlantic Overturning (~ 1 Sv), which may account for the large cooling seen in the North Atlantic (Fig. 25). Antarctic sea-ice trends are qualitatively similar to those shown for the Antarctic circumpolar current, with fairly constant values from the 1850s through the 1970s, with a decline in area thereafter. The average SH area covered with sea-ice decreases by $\sim 6\%$ during the last 30 years of the simulation (1981–2010) relative to the first 100 yr. Note that as the Antarctic circumpolar circulation is substantially stronger than observed in this model, the increase of about 2% from 1980 to 2010 is more realistic than the absolute change.

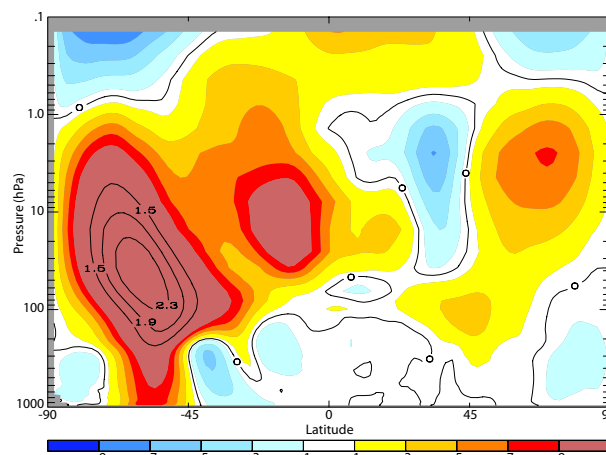


Fig. 28. Annual average zonal wind change (m s^{-1}) between 2000 (1996–2005) and 1970 (1966–1975) in simulations driven by ozone forcing alone.

The strong trends in the Southern Ocean appear to be driven by wind anomalies induced by stratospheric ozone depletion. Ozone losses lead to cooling at high latitudes, increasing the temperature gradient that drives the polar jet (e.g. Polvani et al., 2011; Shindell et al., 2001). While the zonal wind anomalies caused by ozone depletion maximize in the stratosphere, they extend to the surface (Fig. 28). As the near-surface zonal winds maximize around 50° S, this enhancement represents both a strengthening and poleward shift of the winds, consistent with recent observations (Lee et al., 2013). These changes enhance the circumpolar current and in turn accelerate Southern Ocean overturning, a change that is also qualitatively consistent with observed trends in tracers of ocean circulation over the past few decades (Waugh et al., 2013). These strong circulation responses may be the reason the tropospheric ozone ARTP values are not well-suited to stratospheric ozone forcing. We hypothesize that the North Atlantic Overturning changes are a response to the Southern Ocean shifts via the broader thermohaline circulation rather than a response to local NH forcing.

While the enhancement of zonal winds by stratospheric ozone depletion has been well documented, the direct link from ozone to atmosphere to ocean response is more clearly illuminated here than in previous studies which have typically included either multiple forcings simultaneously or explored the response to ozone forcing alone in atmosphere-only models. The results imply that stratospheric ozone depletion could have an impact on long-term carbon storage given the prominent role of Southern Ocean overturning in transporting carbon to the deep ocean. The coupled ocean responses are likely also at least part of the reason for the appearance of a TOA energy imbalance in the model after 1970, as ocean responses can introduce significant time lags into the system, and may also be the cause of the apparent change

in climate sensitivity between the pre-1970 tropospheric ozone forcing and the post-1970 stratospheric/tropospheric ozone forcing.

7 Discussion and conclusions

We find that the model captures many observed features of the distribution and trends for gaseous species important to climate change. The current GISS-E2 model exhibits substantial improvements relative to the GISS-E CMIP3 version. In particular, the seasonality of stratospheric ozone is much more realistic, although the stratospheric circulation is still too rapid. The largest biases in the troposphere in comparison with sondes at around 300 hPa have also been reduced by nearly half, though comparisons at other tropospheric levels show little change in model skill.

The availability of a large amount of new observations, especially near-global satellite datasets, allows evaluation of far more global fields than previously. We performed correlation, bias, and spatial autocorrelation analyses. In comparison with monthly, zonal mean total column ozone observations, the new model shows improvement in all regions. The largest remaining differences relative to observations occur in the NH extratropics, and a significant portion of those is due to circulation as nudging toward the MERRA reanalysis reduces them by about one-third (while having almost no effect on SH extratropical differences). The model captures the spatial distribution of annual average stratospheric ozone very well (R^2 from 0.89 to 0.95), with results in the troposphere showing less skill and a stronger dependence on the observational dataset used ($R^2 = 0.70$ vs. OMI/MLS, $R^2 = 0.85$ vs. TES). Comparisons against MO-PITT CO reveal a strong seasonality in model skill, with reduced agreement with the satellite data during the boreal fall biomass burning season (as in prior studies). Using spatial autocorrelation analysis, we find that the model realistically captures spatial inhomogeneity in the distribution of stratospheric ozone. In the troposphere, results are more mixed, as the modeled NO₂ distribution is too smooth, as is the boreal fall CO distribution, while the boreal spring CO is realistic and tropospheric ozone is realistic in comparison with TES but too well-mixed in comparison with OMI/MLS. We suggest that the spatial autocorrelation can provide a useful metric for evaluating models and especially atmospheric mixing, though it is of course also sensitive to the distribution of emissions and removal processes. Though substantial ozone biases remain in some areas, they tend to be in locations where ozone has less of an impact on RF, and so the difference in the global mean tropospheric ozone RF between calculations using the model's distribution and that observed by TES is only 0.016 W m^{-2} (though regional biases can be substantially larger).

Evaluating historical ozone forcing, we find that whole atmosphere ozone forcing is robust to various subjective

choices, while separation of ozone forcing into tropospheric and stratospheric components is sensitive to the definition of the tropopause. We also find that when ozone changes are primarily in the troposphere, ERF and RF are quite similar. Examining the climate response in simulations driven only by imposed ozone changes, we find that the global mean surface temperature response to radiative forcing due primarily to increased tropospheric ozone from 1850 to ~ 1970 closely tracks that forcing. In addition, regional responses across the tropics and NH mid-latitudes are consistent with estimates based on the pattern of RF using the ARTP metric, though temperature responses at high latitudes and at smaller spatial scales, and precipitation responses, are less tightly coupled to RF. Although the Antarctic ozone hole is too large and persists too long in our model, its depth and the timing of its onset are realistic. Stratospheric ozone depletion causes whole atmosphere ozone RF to decrease over the 1980 to 2000 period. The climate response to this forcing, however, does not follow the RF or ERF as closely at either global or regional scales. Stratospheric ozone depletion appears to cause substantial changes in circulation within the coupled atmosphere-ocean system, including an increase in the Antarctic circumpolar current and in Southern Ocean overturning. Such responses are not captured well by any metric of radiative forcing.

For future scenarios, ozone increases under RCP8.5 owing to growth in methane and stratospheric ozone recovery, while decreasing in the other RCPs due to reductions in ozone precursors (despite stratospheric ozone recovery occurring in those RCPs as well). The interactive composition-climate model provides more physically-based estimates of the time-evolution of ozone and methane forcing than the simple box model used in generating the RCP estimates. Those RCP forcing estimates underestimate the combined methane and ozone forcing by about 0.05, 0.1, 0.2 and 0.25 W m^{-2} under the RCP 2.6, 4.5, 6.0 and 8.5 scenarios, respectively. These values are roughly 4–6 % of the projected 2100 forcing under the RCPs, so the biases are not terribly large at the global scale.

The GISS model now includes fully interactive chemistry and aerosols in our climate simulations, and for the first time we have included interactive methane in our future projections. Additional processes that are not yet included, such as climate-sensitive fire occurrence affecting future emissions or the detrimental effects of ozone on carbon uptake and the effect of CO₂ fertilization on methane emissions, may have substantial impacts on climate-composition projections. For the latter, observations of the response to CO₂ increases from ~ 25 –100 % show that these lead to an increase in methane from wetlands of 13 % and methane from rice paddies of 43 % according to van Groenigen et al. (2011). Hence inclusion of such processes could lead to substantial changes in the prognostic methane found here.

Acknowledgements. We thank Susan Strahan for providing the N₂O and Age-of-air data, and NASA ACMAP and MAP for funding. Resources supporting this work were provided by the NASA High-End Computing (HEC) Program through the NASA Center for Climate Simulation (NCCS) at Goddard Space Flight Center. The SCIAMACHY methane retrievals (WFMDv2.0.2) were obtained from the Institute of Environmental Physics (IUP), University of Bremen, Germany and we thank Oliver Schneising and Michael Buchwitz for their assistance. Model output is available via the WGCM PCDMI ESG gateway. The National Center for Atmospheric Research is operated by the University Corporation for Atmospheric Research under sponsorship of the National Science Foundation.

Edited by: B. N. Duncan

References

- Aghedo, A. M., Bowman, K. W., Shindell, D. T., and Faluvegi, G.: The impact of orbital sampling, monthly averaging and vertical resolution on climate chemistry model evaluation with satellite observations, *Atmos. Chem. Phys.*, 11, 6493–6514, doi:10.5194/acp-11-6493-2011, 2011a.
- Aghedo, A. M., Bowman, K. W., Worden, H. M., Kulawik, S. S., Shindell, D. T., Lamarque, J. F., Faluvegi, G., Parrington, M., Jones, D. B. A., and Rast, S.: The vertical distribution of ozone instantaneous radiative forcing from satellite and chemistry climate models, *J. Geophys. Res.*, 116, D01305, doi:10.1029/2010JD014243, 2011b.
- Andrews, A., Boering, K., Daube, B., Wofsy, S., Loewenstein, M., Jost, H., Podolske, J., Webster, C., Herman, R., Scott, D., Flesch, G., Moyer, E., Elkins, J., Dutton, G., Hurst, D., Moore, F., Ray, E., Romashkin, P., and Strahan, S.: Mean ages of stratospheric air derived from in situ observations of CO₂, CH₄, and N₂O, *J. Geophys. Res.-Atmos.*, 106, 32295–32314, doi:10.1029/2001JD000465, 2001.
- Bauer, S. E., Koch, D., Unger, N., Metzger, S. M., Shindell, D. T., and Streets, D. G.: Nitrate aerosols today and in 2030: a global simulation including aerosols and tropospheric ozone, *Atmos. Chem. Phys.*, 7, 5043–5059, doi:10.5194/acp-7-5043-2007, 2007.
- Beirle, S., Huntrieser, H., and Wagner, T.: Direct satellite observation of lightning-produced NO_x, *Atmos. Chem. Phys.*, 10, 10965–10986, doi:10.5194/acp-10-10965-2010, 2010.
- Bian, H. and Prather, M.: Fast-J2: Accurate simulations of photolysis in global climate models, *J. Atmos. Chem.*, 41, 281–296, 2002.
- Bian, H., Prather, M., and Takemura, T.: Tropospheric aerosol impacts on trace gas budgets through photolysis, *J. Geophys. Res.-Atmos.*, 108, 4242, doi:10.1029/2002JD002743, 2003.
- Bodeker, G. E., Shiona, H., and Eskes, H.: Indicators of Antarctic ozone depletion, *Atmos. Chem. Phys.*, 5, 2603–2615, doi:10.5194/acp-5-2603-2005, 2005.
- Butchart, N., Charlton-Perez, A., Cionni, I., Hardiman, S., Haynes, P., Kruger, K., Kushner, P., Newman, P., Osprey, S., Perlwitz, J., Sigmond, M., Wang, L., Akiyoshi, H., Austin, J., Bekki, S., Baumgaertner, A., Braesicke, P., Bruhl, C., Chipperfield, M., Dameris, M., Dhomse, S., Eyring, V., Garcia, R., Garny, H., Jockel, P., Lamarque, J., Marchand, M., Michou, M., Morgenstern, O., Nakamura, T., Pawson, S., Plummer, D., Pyle, J., Rozanov, E., Scinocca, J., Shepherd, T., Shibata, K., Smale, D., Teysse, H., Tian, W., Waugh, D., and Yamashita, Y.: Multi-model climate and variability of the stratosphere, *J. Geophys. Res.-Atmos.*, 116, D05102, doi:10.1029/2010JD014995, 2011.
- Butkovskaya, N., Kukui, A., and Le Bras, G.: HNO₃ Forming Channel of the HO₂ + NO Reaction as a Function of Pressure and Temperature in the Ranges of 72–600 Torr and 223–323 K, *J. Phys. Chem. A*, 111, 9047–9053, 2007.
- Celari, E., Brinksma, E., Gleason, J., Veeckind, J., Cede, A., Herman, J., Ionov, D., Goutail, F., Pommereau, J., Lambert, J., van Roozendaal, M., Pinardi, G., Wittrock, F., Schonhardt, A., Richter, A., Ibrahim, O., Wagner, T., Bojkov, B., Mount, G., Spinei, E., Chen, C., Pongetti, T., Sander, S., Bucsel, E., Wenig, M., Swart, D., Volten, H., Kroon, M., and Levelt, P.: Validation of ozone monitoring instrument nitrogen dioxide columns, *J. Geophys. Res.-Atmos.*, 113, D15S15, doi:10.1029/2007JD008908, 2008.
- Chandra, S., Ziemke, J. R., and Stewart, R. W.: An 11-year solar cycle in tropospheric ozone from TOMS measurements, *Geophys. Res. Lett.*, 26, 185–188, 1999.
- Chiang, J. and Bitz, C.: Influence of high latitude ice cover on the marine Intertropical Convergence Zone, *Clim. Dynam.*, 25, 477–496, doi:10.1007/s00382-005-0040-5, 2005.
- Cionni, I., Eyring, V., Lamarque, J. F., Randel, W. J., Stevenson, D. S., Wu, F., Bodeker, G. E., Shepherd, T. G., Shindell, D. T., and Waugh, D. W.: Ozone database in support of CMIP5 simulations: results and corresponding radiative forcing, *Atmos. Chem. Phys.*, 11, 11267–11292, doi:10.5194/acp-11-11267-2011, 2011.
- Conley, A. J., Lamarque, J.-F., Vitt, F., Collins, W. D., and Kiehl, J.: PORT, a CESM tool for the diagnosis of radiative forcing, *Geosci. Model Dev. Discuss.*, 5, 2687–2704, doi:10.5194/gmdd-5-2687-2012, 2012.
- Dee, D., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hersbach, H., Holm, E., Isaksen, L., Kallberg, P., Kohler, M., Matricardi, M., McNally, A., Monge-Sanz, B., Morcrette, J., Park, B., Peubey, C., de Rosnay, P., Tavolato, C., Thepaut, J., and Vitart, F.: The ERA-interim reanalysis: configuration and performance of the data assimilation system, *Q. J. Roy. Meteor. Soc.*, 137, 553–597, doi:10.1002/qj.828, 2011.
- Deeter, M. N.: Operational carbon monoxide retrieval algorithm and selected results for the MOPITT instrument, *J. Geophys. Res.*, 108, 4399, doi:10.1029/2002JD003186, 2003.
- Denman, K. L., Brasseur, G., Chidthaisong, A., Ciais, P., Cox, P. M., Dickinson, R. E., Hauglustaine, D., Heinze, C., Holland, E., Jacob, D., Lohmann, U., Ramachandran, S., Dias, P. L. d. S., Wofsy, S. C., and Zhang, X.: Couplings between changes in the climate system and biogeochemistry, in: *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited, Cambridge University Press, Cambridge, UK and New York, NY, USA, 2007.
- Derwent, R., Simmonds, P., Manning, A., and Spain, T.: Trends over a 20-year period from 1987 to 2007 in surface ozone at the atmospheric research station, Mace Head, Ireland, *Atmos. Environ.*, 41, 9091–9098, doi:10.1016/j.atmosenv.2007.08.008, 2007.

- Dlugokencky, E. J., Bruhwiler, L., White, W. C., Emmons, L. K., Novelli, P. C., Montzka, S. A., Masarie, K. A., Lang, P. M., Crotwell, A. M., Miller, J. B., and Gatti, L. V.: Observational constraints on recent increases in the atmospheric CH₄ burden, *Geophys. Res. Lett.*, 36, L18803, doi:10.1029/2009GL039780, 2009.
- Eyring, V., I. Cionni, J. M. Arblaster, J. Sedláček, J. Perlwitz, P. J. Young, S. Bekki, D. Bergmann, P. Cameron-Smith, W. J. Collins, G. Faluvegi, K.-D. Gottschaldt, L. W. Horowitz, D. E. Kinnison, J.-F. Lamarque, D. R. Marsh, D. Saint-Martin, D. T. Shindell, K. Sudo, S. Szopa, and S. Watanabe: Long-term changes in tropospheric and stratospheric ozone and associated climate impacts in CMIP5 simulations, *J. Geophys. Res.*, submitted, 2012.
- Forster, P., Ramaswamy, V., Artaxo, P., Bernsten, T., Betts, R., Fahey, D. W., Haywood, J., Lean, J., Lowe, D. C., Myhre, G., Nganga, J., Prinn, R., Raga, G., Schulz, M., and Van Dorland, R.: Changes in atmospheric constituents and in radiative forcing, in: *Climate Change 2007: The Physical Science Basis, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited, Cambridge University Press, Cambridge, UK and New York, NY, USA, 2007.
- Froidevaux, L., Jiang, Y., Lambert, A., Livesey, N., Read, W., Waters, J., Browell, E., Hair, J., Avery, M., Mcgee, T., Twigg, L., Sumnicht, G., Jucks, K., Margitan, J., Sen, B., Stachnik, R., Toon, G., Bernath, P., Boone, C., Walker, K., Filipiak, M., Harwood, R., Fuller, R., Manney, G., Schwartz, M., Daffer, W., Drouin, B., Cofield, R., Cuddy, D., Jarnot, R., Knosp, B., Perun, V., Snyder, W., Stek, P., Thurstans, R., and Wagner, P.: Validation of aura microwave limb sounder stratospheric ozone measurements, *J. Geophys. Res.-Atmos.*, 113, D15S20, doi:10.1029/2007JD008771, 2008.
- Gery, M. W., Whitten, G. Z., Killus, J. P., and Dodge, M. C.: A photochemical kinetics mechanism for urban and regional scale computer modeling, *J. Geophys. Res.*, 94, 925–956, 1989.
- Gottelman, A., Holton, J., and Rosenlof, K.: Mass fluxes of O₃, CH₄, N₂O and CF₂Cl₂ in the lower stratosphere calculated from observational data, *J. Geophys. Res.-Atmos.*, 102, 19149–19159, 1997a.
- Gottelman, A., Holton, J. R., and Rosenlof, K. H.: Mass fluxes of O₃, CH₄, N₂O, and CF₂Cl₂ in the lower stratosphere calculated from observational data, *J. Geophys. Res.*, 102, 19149–19159, 1997b.
- Hallquist, M., Stewart, D. J., Stephenson, S. K., and Cox, R. A.: Hydrolysis of N₂O₅ on sub-micron sulfate aerosols, *Phys. Chem. Chem. Phys.*, 5, 3453–3463, 2003.
- Hansen, J., Sato, M., Ruedy, R., Nazarenko, L., Lacis, A., Schmidt, G. A., Russell, G., Aleinov, I., Bauer, M., Bauer, S., Bell, N., Cairns, B., Canuto, V., Chandler, M., Cheng, Y., Del Genio, A., Faluvegi, G., Fleming, E., Friend, A., Hall, T., Jackman, C., Kelley, M., Kiang, N., Koch, D., Lean, J., Lerner, J., Lo, K., Menon, S., Miller, R., Minnis, P., Novakov, T., Oinas, V., Perlwitz, J., Rind, D., Romanou, A., Shindell, D., Stone, P., Sun, S., Tausnev, N., Thresher, D., Wielicki, B., Wong, T., Yao, M., and Zhang, S.: Efficacy of climate forcings, *J. Geophys. Res.-Atmos.*, 110, D18104, doi:10.1029/2005jd005776, 2005.
- Hanson, D. and Mauersberger, K.: Laboratory studies of the nitric acid tridhydrate: Implications for the south polar stratosphere, *Geophys. Res. Lett.*, 15, 855–858, 1988.
- Houweling, S., Dentener, F., and Lelieveld, J.: The impact of non-methane hydrocarbon compounds on tropospheric photochemistry, *J. Geophys. Res.*, 103, 10673–10696, 1998.
- Kane, S. M., Caloz, F., and Leu, M.-T.: Heterogeneous Uptake of Gaseous N₂O₅ by (NH₄)₂SO₄, NH₄H₂SO₄, and H₂SO₄ Aerosols, *J. Phys. Chem. A*, 105, 6465–6470, 2001.
- Kang, S., Frierson, D., and Held, I.: The Tropical Response to Extratropical Thermal Forcing in an Idealized GCM: The Importance of Radiative Feedbacks and Convective Parameterization, *J. Atmos. Sci.*, 66, 2812–2827, doi:10.1175/2009JAS2924.1, 2009.
- Kawase, H., Nagashima, T., Sudo, K., and Nozawa, T.: Future changes in tropospheric ozone under Representative Concentration Pathways (RCPs), *Geophys. Res. Lett.*, 38, L05801, doi:10.1029/2010gl046402, 2011.
- Lamarque, J.-F., Bond, T. C., Eyring, V., Granier, C., Heil, A., Klimont, Z., Lee, D., Liousse, C., Mieville, A., Owen, B., Schultz, M. G., Shindell, D., Smith, S. J., Stehfest, E., Van Aardenne, J., Cooper, O. R., Kainuma, M., Mahowald, N., McConnell, J. R., Naik, V., Riahi, K., and van Vuuren, D. P.: Historical (1850–2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: methodology and application, *Atmos. Chem. Phys.*, 10, 7017–7039, doi:10.5194/acp-10-7017-2010, 2010.
- Lamarque, J.-F., Shindell, D. T., Josse, B., Young, P. J., Cionni, I., Eyring, V., Bergmann, D., Cameron-Smith, P., Collins, W. J., Doherty, R., Dalsoren, S., Faluvegi, G., Folberth, G., Ghan, S. J., Horowitz, L. W., Lee, Y. H., MacKenzie, I. A., Nagashima, T., Naik, V., Plummer, D., Righi, M., Rumbold, S. T., Schulz, M., Skeie, R. B., Stevenson, D. S., Strode, S., Sudo, K., Szopa, S., Voulgarakis, A., and Zeng, G.: The Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP): overview and description of models, simulations and climate diagnostics, *Geosci. Model Dev.*, 6, 179–206, doi:10.5194/gmd-6-179-2013, 2013.
- Lee, S., and S. B. Feldstein.: Detecting Ozone- and Greenhouse Gas-Driven Wind Trends with Observational Data, *Science*, 339, 563–567, 2013.
- Logan, J. A.: An analysis of ozonesonde data for the troposphere: Recommendations for testing 3-D models and development of a gridded climatology for tropospheric ozone, *J. Geophys. Res.*, 104, 16115–16149, 1999.
- McPeters, R., Kroon, M., Labow, G., Brinksma, E., Balis, D., Petropavlovskikh, I., Veefkind, J., Bhartia, P., and Levelt, P.: Validation of the aura ozone monitoring instrument total column ozone product, *J. Geophys. Res.-Atmos.*, 113, D15S14, doi:10.1029/2007JD008802, 2008.
- Meinshausen, M., Raper, S. C. B., and Wigley, T. M. L.: Emulating coupled atmosphere-ocean and carbon cycle models with a simpler model, *MAGICC6 – Part 1: Model description and calibration*, *Atmos. Chem. Phys.*, 11, 1417–1456, doi:10.5194/acp-11-1417-2011, 2011.
- Menon, S., Koch, D., Beig, G., Sahu, S., Fasullo, J., and Orlikowski, D.: Black carbon aerosols and the third polar ice cap, *Atmos. Chem. Phys.*, 10, 4559–4571, doi:10.5194/acp-10-4559-2010, 2010.
- Murphy, D. and Fahey, D.: An estimate of the flux of stratospheric reactive nitrogen and ozone into the troposphere, *J. Geophys. Res.-Atmos.*, 99, 5325–5332, 1994.

- Nassar, R., Logan, J., Worden, H., Megretskaia, I., Bowman, K., Osterman, G., Thompson, A., Tarasick, D., Austin, S., Claude, H., Dubey, M., Hocking, W., Johnson, B., Joseph, E., Merrill, J., Morris, G., Newchurch, M., Oltmans, S., Posny, F., Schmidlin, F., Vomel, H., Whiteman, D., and Witte, J.: Validation of Tropospheric Emission Spectrometer (TES) nadir ozone profiles using ozonesonde measurements, *J. Geophys. Res.-Atmos.*, 113, D15S17, doi:10.1029/2007JD008819, 2008.
- Olsen, M. A., Douglass, A. R., and Schoeberl, M. R.: A comparison of Northern and Southern Hemisphere cross-tropopause ozone flux, *Geophys. Res. Lett.*, 30, 1412, doi:10.1029/2002GL016538, 2003.
- Oltmans, S., Lefohn, A., Harris, J., Galbally, I., Scheel, H., Bodeker, G., Brunke, E., Claude, H., Tarasick, D., Johnson, B., Simmonds, P., Shadwick, D., Anlauf, K., Hayden, K., Schmidlin, F., Fujimoto, T., Akagi, K., Meyer, C., Nichol, S., Davies, J., Redondas, A., and Cuevas, E.: Long-term changes in tropospheric ozone, *Atmos. Environ.*, 40, 3156–3173, doi:10.1016/j.atmosenv.2006.01.029, 2006.
- Osterman, G. B., Kulawik, S. S., Worden, H. M., Richards, N. A. D., Fisher, B. M., Eldering, A., Shephard, M. W., Froidevaux, L., Labow, G., Luo, M., Herman, R. L., Bowman, K. W., and Thompson, A. M.: Validation of Tropospheric Emission Spectrometer (TES) measurements of the total, stratospheric, and tropospheric column abundance of ozone, *J. Geophys. Res.-Atmos.*, 113, D15S16, doi:10.1029/2007jd008801, 2008.
- Polvani, L. M., Waugh, D. W., Correa, G. J. P., and Son, S. W.: Stratospheric Ozone Depletion: The Main Driver of Twentieth-Century Atmospheric Circulation Changes in the Southern Hemisphere, *J. Climate*, 24, 795–812, doi:10.1175/2010jcli3772.1, 2011.
- Prather, M. J., Ehhalt, D., Dentener, F., Derwent, R. G., Dlugokencky, E. J., Holland, E. A., Isaksen, I. S., Katima, J., Kirchhoff, V., Matson, P., Midgley, P., and Wang, M.: Atmospheric chemistry and greenhouse gases, in: *Climate Change 2001*, edited by: Houghton, J. T., Cambridge Univ. Press, Cambridge, pp. 239–287, 2001.
- Prather, M. J., Holmes, C. D., and Hsu, J.: Reactive greenhouse gas scenarios: Systematic exploration of uncertainties and the role of atmospheric chemistry, *Geophys. Res. Lett.*, 39, L09803, doi:10.1029/2012gl051440, 2012.
- Ramaswamy, V., Boucher, O., Haigh, J. D., Hauglustaine, D. A., Haywood, J., Myhre, G., Nakajima, T., Shi, G. Y., and Solomon, S.: Radiative forcing of climate change, in: *Climate Change 2001*, edited by: J. T. Houghton, 349–416, Cambridge Univ. Press, Cambridge, 2001.
- Randel, W. J. and Wu, F.: A stratospheric ozone trends data set for global modeling studies, *Geophys. Res. Lett.*, 26, 3089–3092, 1999.
- Randel, W. J., Wu, F., Russell III, J. M., Waters, J. W., and Froidevaux, L.: Ozone and temperature changes in the stratosphere following the eruption of Mt. Pinatubo, *J. Geophys. Res.*, 100, 16753–16764, 1995.
- Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., Kent, E. C., and Kaplan, A.: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, *J. Geophys. Res.*, 108, 4407, doi:10.1029/2002JD002670, 2003.
- Rienecker, M., Suarez, M., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M., Schubert, S., Takacs, L., Kim, G., Bloom, S., Chen, J., Collins, D., Conaty, A., Da Silva, A., Gu, W., Joiner, J., Koster, R., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C., Reichle, R., Robertson, F., Ruddick, A., Sienkiewicz, M., and Woollen, J.: MERRA: NASA's modern-era retrospective analysis for research and applications, *J. Climate*, 24, 3624–3648, doi:10.1175/JCLI-D-11-00015.1, 2011.
- Russell, G. L., Miller, J. R., and Rind, D.: A coupled atmosphere-ocean model for transient climate change, *Atmos. Ocean*, 33, 683–730, 1995.
- Schmidt, G. A., Ruedy, R., Hansen, J. E., Aleinov, I., Bell, N., Bauer, M., Bauer, S., Cairns, B., Canuto, V., Cheng, Y., Del Genio, A., Faluvegi, G., Friend, A. D., Hall, T. M., Hu, Y., Kelley, M., Kiang, N. Y., Koch, D., Lacis, A. A., Lerner, J., Lo, K. K., Miller, R. L., Nazarenko, L., Oinas, V., Perlwitz, J., Perlwitz, J., Rind, D., Romanou, A., Russell, G. L., Sato, M., Shindell, D. T., Stone, P. H., Sun, S., Tausnev, N., Thresher, D., and Yao, M.-S.: Present day atmospheric simulations using GISS ModelE: comparison to in-situ, satellite and reanalysis data, *J. Climate*, 19, 153–192, 2006.
- Schneising, O., Buchwitz, M., Burrows, J. P., Bovensmann, H., Bergamaschi, P., and Peters, W.: Three years of greenhouse gas column-averaged dry air mole fractions retrieved from satellite – Part 2: Methane, *Atmos. Chem. Phys.*, 9, 443–465, doi:10.5194/acp-9-443-2009, 2009.
- Schneising, O., Buchwitz, M., Reuter, M., Heymann, J., Bovensmann, H., and Burrows, J. P.: Long-term analysis of carbon dioxide and methane column-averaged mole fractions retrieved from SCIAMACHY, *Atmos. Chem. Phys.*, 11, 2863–2880, doi:10.5194/acp-11-2863-2011, 2011.
- Schumann, U. and Huntrieser, H.: The global lightning-induced nitrogen oxides source, *Atmos. Chem. Phys.*, 7, 3823–3907, doi:10.5194/acp-7-3823-2007, 2007.
- Shindell, D. T.: Evaluation of the absolute regional temperature potential, *Atmos. Chem. Phys.*, 12, 7955–7960, doi:10.5194/acp-12-7955-2012, 2012.
- Shindell, D. and Faluvegi, G.: Climate response to regional radiative forcing during the 20th century, *Nature Geosci.*, 2, 294–300, 2009.
- Shindell, D., Schmidt, G., Miller, R., and Rind, D.: Northern Hemisphere winter climate response to greenhouse gas, ozone, solar, and volcanic forcing, *J. Geophys. Res.-Atmos.*, 106, 7193–7210, doi:10.1029/2000JD900547, 2001.
- Shindell, D., Schulz, M., Ming, Y., Takemura, T., Faluvegi, G., and Ramaswamy, V.: Spatial scales of climate response to inhomogeneous radiative forcing, *J. Geophys. Res.*, 115, D19110, doi:10.1029/2010JD014108, 2010.
- Shindell, D., Faluvegi, G., Nazarenko, L., Bowman, K., Lamarque, J.-F., Voulgarakis, A., Schmidt, G., Pechony, O., and Reudy, R.: Attribution of historical ozone forcing to anthropogenic emissions, *Nature Climate Change*, doi:10.1038/NCLIMATE1835, 2013.
- Shindell, D. T., Walter, B. P., and Faluvegi, G.: Impacts of climate change on methane emissions from wetlands, *Geophys. Res. Lett.*, 31, L21202, doi:10.1029/2004gl021009, 2004.
- Shindell, D. T., G. Faluvegi, D. S. Stevenson, M. C. Krol, L. K. Emmons, J.-F. Lamarque, G. Petron, F. J. Dentener, K. Ellingsen, M.

- G. Schultz, O. Wild, M. Amann, C. S. Atherton, D. J. Bergmann, I. Bey, T. Butler, J. Cofala, W. J. Collins, R. G. Derwent, R. M. Doherty, J. Drevet, H. J. Eskes, A. M. Fiore, M. Gauss, D. A. Hauglustaine, L. W. Horowitz, I. S. A. Isaksen, M. G. Lawrence, V. Montanaro, J.-F. Muller, G. Pitari, M. J. Prather, J. A. Pyle, S. Rast, J. M. Rodriguez, M. G. Sanderson, N. H. Savage, S. E. Strahan, K. Sudo, S. Szopa, N. Unger, T. P. C. van Noije, and G. Zeng: Multi-model simulations of carbon monoxide: Comparison with observations and projected near-future changes, *J. Geophys. Res.*, 111, D19306, doi:10.1029/2006JD007100, 2006a.
- Shindell, D. T., Faluvegi, G., Unger, N., Aguilar, E., Schmidt, G. A., Koch, D. M., Bauer, S. E., and Miller, R. L.: Simulations of preindustrial, present-day, and 2100 conditions in the NASA GISS composition and climate model G-PUCCINI, *Atmos. Chem. Phys.*, 6, 4427–4459, doi:10.5194/acp-6-4427-2006, 2006b.
- Stevenson, D. S., F. J. Dentener, M. G. Schultz, K. Ellingsen, T. P. C. van Noije, O. Wild, G. Zeng, M. Amann, C. S. Atherton, N. Bell, D. J. Bergmann, I. Bey, T. Butler, J. Cofala, W. J. Collins, R. G. Derwent, R. M. Doherty, J. Drevet, H. J. Eskes, A. M. Fiore, M. Gauss, D. A. Hauglustaine, L. W. Horowitz, I. S. A. Isaksen, M. C. Krol, J.-F. Lamarque, M. G. Lawrence, V. Montanaro, J.-F. Muller, G. Pitari, M. J. Prather, J. A. Pyle, S. Rast, J. M. Rodriguez, M. G. Sanderson, N. H. Savage, D. T. Shindell, S. E. Strahan, K. Sudo, and S. Szopa: Multi-model ensemble simulations of present-day and near-future tropospheric ozone, *J. Geophys. Res.*, 111, D08301, doi:10.1029/2005JD006338, 2006.
- Stockwell, W. R., Kirchner, F., Kuhn, M., and Seefeld, S.: A new mechanism for regional atmospheric chemistry modeling, *J. Geophys. Res.*, 102, 25847–25879, 1997.
- Strahan, S., Douglass, A., Stolarski, R., Akiyoshi, H., Bekki, S., Braesicke, P., Butchart, N., Chipperfield, M., Cugnet, D., Dhomse, S., Frith, S., Gettelman, A., Hardiman, S., Kinnison, D., Lamarque, J., Mancini, E., Marchand, M., Michou, M., Morgenstern, O., Nakamura, T., Olivie, D., Pawson, S., Pitari, G., Plummer, D., Pyle, J., Scinocca, J., Shepherd, T., Shibata, K., Smale, D., Teysedre, H., Tian, W., and Yamashita, Y.: Using transport diagnostics to understand chemistry climate model ozone simulations, *J. Geophys. Res.-Atmos.*, 116, D17302, doi:10.1029/2010JD015360, 2011.
- Sun, S. and Bleck, R.: Multi-century simulations with the coupled GISS-HYCOM climate model: Control experiments, *Clim. Dynam.*, 26, 407–428, 2006.
- Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An Overview of CMIP5 and the experiment design, *B. Am. Meteorol. Soc.*, 93, 485–498, doi:10.1175/BAMS-D-11-00094.1, 2012.
- Thompson, A., Witte, J., Smit, H., Oltmans, S., Johnson, B., Kirchhoff, V., and Schmidlin, F.: Southern Hemisphere Additional Ozonesondes (SHADOZ) 1998–2004 tropical ozone climatology: 3. Instrumentation, station-to-station variability, and evaluation with simulated flight profiles, *J. Geophys. Res.-Atmos.*, 112, D03304, doi:10.1029/2005JD007042, 2007.
- Tsigaridis, K., Lathière, J., Kanakidou, M., and Hauglustaine, D. A.: Naturally driven variability in the global secondary organic aerosol over a decade, *Atmos. Chem. Phys.*, 5, 1891–1904, doi:10.5194/acp-5-1891-2005, 2005.
- Uppala, S., Kallberg, P., Simmons, A., Andrae, U., Bechtold, V., Fiorino, M., Gibson, J., Haseler, J., Hernandez, A., Kelly, G., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R., Andersson, E., Arpe, K., Balmaseda, M., Beljaars, A., Van De Berg, L., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Holm, E., Hoskins, B., Isaksen, L., Janssen, P., Jenne, R., McNally, A., Mahfouf, J., Morcrette, J., Rayner, N., Saunders, R., Simon, P., Sterl, A., Trenberth, K., Untch, A., Vasiljevic, D., Viterbo, P., and Woollen, J.: The ERA-40 re-analysis, *Q. J. Roy. Meteorol. Soc.*, 131, 2961–3012, doi:10.1256/qj.04.176, 2005.
- van der Werf, G. R., Randerson, J. T., Giglio, L., Collatz, G. J., Mu, M., Kasibhatla, P. S., Morton, D. C., DeFries, R. S., Jin, Y., and van Leeuwen, T. T.: Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009), *Atmos. Chem. Phys.*, 10, 11707–11735, doi:10.5194/acp-10-11707-2010, 2010.
- van Groenigen, K. J., Osenberg, C. W., and Hungate, B. A.: Increased soil emissions of potent greenhouse gases under increased atmospheric CO₂, *Nature*, 475, 214–216, 2011.
- van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G. C., Kram, T., Krey, V., Lamarque, J. F., Masui, T., Meinshausen, M., Nakicenovic, N., Smith, S. J., and Rose, S. K.: The representative concentration pathways: an overview, *Climatic Change*, 109, 5–31, 2011.
- Voulgarakis, A., Telford, P. J., Aghedo, A. M., Braesicke, P., Faluvegi, G., Abraham, N. L., Bowman, K. W., Pyle, J. A., and Shindell, D. T.: Global multi-year O₃-CO correlation patterns from models and TES satellite observations, *Atmos. Chem. Phys.*, 11, 5819–5838, doi:10.5194/acp-11-5819-2011, 2011.
- Voulgarakis, A., Naik, V., Lamarque, J.-F., Shindell, D. T., Young, P. J., Prather, M. J., Wild, O., Field, R. D., Bergmann, D., Cameron-Smith, P., Cionni, I., Collins, W. J., Dalsøren, S. B., Doherty, R. M., Eyring, V., Faluvegi, G., Folberth, G. A., Horowitz, L. W., Josse, B., McKenzie, I. A., Nagashima, T., Plummer, D. A., Righi, M., Rumbold, S. T., Stevenson, D. S., Strode, S. A., Sudo, K., Szopa, S., and Zeng, G.: Analysis of present day and future OH and methane lifetime in the ACCMIP simulations, *Atmos. Chem. Phys.*, 13, 2563–2587, doi:10.5194/acp-13-2563-2013, 2012.
- Wang, Y. H. and Jacob, D.: Anthropogenic forcing on tropospheric ozone and OH since preindustrial times, *J. Geophys. Res.*, 103, 31123–31135, 1998.
- Waugh, D. W., Primeau, F., DeVries, T., and Holzer, M.: Recent Changes in the Ventilation of the Southern Oceans, *Science*, 339, 568–570, 2013.
- Wild, O.: Modelling the global tropospheric ozone budget: exploring the variability in current models, *Atmos. Chem. Phys.*, 7, 2643–2660, doi:10.5194/acp-7-2643-2007, 2007.
- Worden, H. M., Bowman, K. W., Worden, J. R., Eldering, A., and Beer, R.: Satellite measurements of the clear-sky greenhouse effect from tropospheric ozone, *Nature Geosci.*, 1, 305–308, 2008.
- Worden, H. M., Bowman, K. W., Kulawik, S. S., and Aghedo, A. M.: Sensitivity of outgoing longwave radiative flux to the global vertical distribution of ozone characterized by instantaneous radiative kernels from Aura-TES, *J. Geophys. Res.*, 116, D14115, doi:10.1029/2010JD015101, 2011.
- Young, P. J., Archibald, A. T., Bowman, K. W., Lamarque, J.-F., Naik, V., Stevenson, D. S., Tilmes, S., Voulgarakis, A., Wild, O., Bergmann, D., Cameron-Smith, P., Cionni, I., Collins, W. J., Dalsøren, S. B., Doherty, R. M., Eyring, V., Faluvegi, G., Horowitz, L. W., Josse, B., Lee, Y. H., MacKenzie, I. A.,

Nagashima, T., Plummer, D. A., Righi, M., Rumbold, S. T., Skeie, R. B., Shindell, D. T., Strode, S. A., Sudo, K., Szopa, S., and Zeng, G.: Pre-industrial to end 21st century projections of tropospheric ozone from the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP), *Atmos. Chem. Phys.*, 13, 2063–2090, doi:10.5194/acp-13-2063-2013, 2013.

Ziemke, J. R., Chandra, S., Labow, G. J., Bhartia, P. K., Froidevaux, L., and Witte, J. C.: A global climatology of tropospheric and stratospheric ozone derived from Aura OMI and MLS measurements, *Atmos. Chem. Phys.*, 11, 9237–9251, doi:10.5194/acp-11-9237-2011, 2011.