



1 Evaluation of CESM1 (WACCM) free-running and 2 specified-dynamics atmospheric composition simulations 3 using global multi-species satellite data records

4 Lucien Froidevaux¹, Douglas E. Kinnison², Ray Wang³, John Anderson⁴, and Ryan A. Fuller¹

5 ¹ Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA

6 ² National Center for Atmospheric Research, Boulder, CO, USA

7 ³ School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA, USA

8 ⁴ School of Science, Hampton University, Hampton, VA, USA

9

10 Correspondence to: L. Froidevaux (lucienf@jpl.nasa.gov)

11 Abstract

12 We evaluate the recently delivered Community Earth System Model version 1 (CESM1) Whole Atmosphere Community
13 Climate Model (WACCM) using satellite-derived global composition datasets, focusing on the stratosphere. The simulations
14 include free-running (FR-WACCM) and specified-dynamics (SD-WACCM) versions of the model. Model evaluations are made
15 using global monthly zonal mean time series obtained by the Aura Microwave Limb Sounder (MLS), as well as longer-term
16 global data records compiled by the Global Ozone Chemistry and Related Trace gas Data Records for the Stratosphere
17 (GOZCARDS) project. A recent update (version 2.20) to the original GOZCARDS merged ozone (O₃) data set is used. We
18 discuss upper atmospheric climatology and zonal mean variability using O₃, hydrogen chloride (HCl), nitrous oxide (N₂O), nitric
19 acid (HNO₃), and water vapor (H₂O) data. There are a few significant model/data mean biases, such as for lower stratospheric
20 O₃, for which the models overestimate the mean observed values and seasonal amplitudes. Another clear difference occurs for
21 HNO₃ during recurring winter periods of strong HNO₃ enhancements at high latitudes; this stems from the known omission of
22 ion chemistry relating to particle precipitation effects, in the global models used here. In the lower stratosphere at high southern
23 latitudes, the variations in polar winter/spring composition observed by MLS are generally well matched by SD-WACCM, the
24 main exception being for the early winter rate of decrease in HCl, which is too slow in the model. In general, the
25 latitude/pressure distributions of annual and semi-annual oscillation amplitudes derived from the MLS data are properly captured
26 by the corresponding model values. Nevertheless, detailed aspects of the interactions between the quasi-biennial, annual, and
27 semi-annual ozone variations in the upper stratosphere are not as well represented by FR-WACCM as by SD-WACCM.

28 One of the evaluation diagnostics we use represents the closeness of fit between the model/data anomaly time series, and we
29 also consider the correlation coefficients. Not surprisingly, SD-WACCM, which is driven by realistic dynamics, generally
30 matches observed deseasonalized anomalies better than FR-WACCM does. Other results indicate that the root mean square
31 variability is sometimes found to be significantly smaller in FR-WACCM than in SD-WACCM and the observations. Most
32 notably, FR-WACCM underestimates the observed interannual variability for H₂O by ~30%, typically, and by as much as a
33 factor of two in some regions; this has some implications for the time needed to detect small trends.

34 We have derived trends using a multivariate linear regression (MLR) model, and there is a robust signal in both MLS
35 observations and WACCM of an upper stratospheric O₃ increase from 2005 to 2014 by ~0.2-0.4%/yr ($\pm 0.2\%/yr$, 2σ), depending
36 on which broad latitude bin (tropics or mid-latitudes) is considered. In the lower stratosphere, while some decreases are indicated
37 for 1998-2014 (based on merged GOZCARDS O₃), we find near-zero or positive trends when using MLS O₃ data alone for
38 2005-2014, albeit with no robust statistical significance. SD-WACCM results track such positive tendencies (albeit with no
39 statistical significance). For H₂O, the most statistically significant trend result for 2005-2014 is an upper stratospheric increase,
40 peaking at slightly more than 0.5%/yr in the lower mesosphere, in fairly close agreement with SD-WACCM trends, but with
41 smaller values in FR-WACCM. For HCl, while the lower stratospheric vertical gradients of MLS trends are duplicated to some
42 extent by SD-WACCM, the model trends (decreases) are always on the low side of the data trends. There is little model-based
43 indication (in SD-WACCM) of a significantly positive HCl trend derived from the MLS tropical series at 68 hPa; this deserves
44 further study. For N₂O, the MLS-derived trends (for 2005-2012) point to negative trends (of up to about -1%/yr) in the NH mid-
45 latitudes and positive trends (of up to about +3%/yr) in the SH mid-latitudes, in good agreement with the asymmetry that exists



1 in SD-WACCM trend results. The small observed positive N₂O trends of ~0.2%/yr in the 100 to 30 hPa tropical region are also
2 consistent with model results (SD-WACCM in particular), which in turn are very close to the known rate of increase in
3 tropospheric N₂O. In the case of HNO₃, MLS-derived lower stratospheric trend differences (for 2005-2014) between
4 hemispheres are opposite in sign to those from N₂O and in reasonable agreement with both WACCM results, despite large error
5 bars.

6 The data sets and tools discussed here for the evaluation of the models could be expanded to additional comparisons of
7 species not included here, as well as to model intercomparisons using a variety of CCMs, keeping in mind that there are different
8 parameterizations and approaches for both free-running and specified-dynamics simulations.
9

10 1 Introduction

11 State-of-the art chemistry climate models (CCMs) are known to reproduce the main features of stratospheric climatology and
12 change, although there has always been some variability between the models (e.g., Waugh and Eyring, 2008; SPARC, 2010;
13 Dhomse et al., 2018). Free-running CCMs are used to make long-term simulations of atmospheric composition, as well as
14 predictions of future changes, driven by a large set of photochemical reactions, as well as time-dependent boundary conditions
15 for surface concentrations of greenhouse gases and ozone depleting substances (ODSs), sea surface temperatures and sea ice
16 concentrations, 11-year solar variability, sulfate aerosol surface area density, as well as tropospheric ozone and aerosol precursor
17 emissions. In more recent years, modeling groups have implemented “specified-dynamics” versions that are constrained to
18 meteorological fields (e.g., surface pressure, temperature, and horizontal and meridional winds). Our main purpose here is to
19 evaluate these two types of model runs from CESM1 WACCM, using multi-species satellite-derived global composition
20 datasets; we will refer to these two types as FR-WACCM, for the free-running model version, and SD-WACCM, for the
21 specified dynamics version. The model simulations were based on scenarios defined by the Chemistry Climate Model Initiative
22 (CCMI) (Eyring et al., 2013; Morgenstern et al., 2017). Our evaluation focus is on monthly zonal mean time series from the
23 models versus satellite-derived global data sets. The main stratospheric (and mesospheric) time series used here are from Aura
24 MLS products (version 4.2 data) and from the Global OZone Chemistry And Related trace gas Data records for the Stratosphere
25 (GOZCARDS), which include MLS data from late 2004 onward. The GOZCARDS data records include merged multi-satellite
26 data files for O₃, H₂O, and HCl, and Aura MLS-derived data files for HNO₃ and N₂O; these 5 species are used for the model
27 evaluations herein. We also focus, in part, on the Aura MLS high quality data sets for 2005-2014 (with 2014 being the last year
28 of WACCM runs considered here). The regular and nearly uninterrupted daily global coverage of the MLS day and night
29 measurements leads to minimal sampling-related biases, both for climatological comparisons (see Toohey et al., 2013) and trend-
30 related studies (Millan et al., 2016). This data set also has a well characterized set of error bars (see Livesey et al., 2018, for the
31 latest update to the data quality documentation); however, we also note that there are some caveats to take into account regarding
32 long-term stability for some of the MLS species.

33 In terms of the model/data comparisons, we will analyze the climatological mean state and “goodness of fit” issues, as well as
34 variability. While one has the expectation that, in general, better fits to the data would be obtained for a specified-dynamics run
35 than for a less dynamically constrained run (FR-WACCM), one needs to demonstrate this quantitatively with diagnostics that can
36 provide enough differentiation between models that are sometimes found to track each other closely. There have been essentially
37 no published trend studies using the Aura MLS data set by itself. However, this data set now covers a sufficiently long time
38 period that it becomes useful to investigate such trends, as the analyses deal with one homogeneous data set, while removing the
39 potential issues associated with data merging prior to 2005, whether related to poorer sampling, or to uncertainties in the bias
40 removal between various data sets. On the downside, a shorter time series will also lead to larger uncertainties in the derived
41 trends.



1 We first provide (in Sect. 2) an overview of the global stratospheric data sets used for these comparisons. Brief descriptions
2 of FR-WACCM and SD-WACCM are provided in Sect. 3. Climatological comparisons between the models and the Aura MLS
3 data sets are provided in Sect. 4, in order to assess, for example, whether any obvious biases exist; these comparisons include an
4 overview of the main short-term variations, namely the annual oscillation (AO) and semi-annual oscillation (SAO). More
5 detailed comparisons of deseasonalized anomaly time series are provided in Sect. 5.1, where we evaluate how well the two
6 model versions fit the data sets, both in terms of closeness of fits and variability. Finally, trend comparisons are investigated in
7 Sect. 5.2.

8

9 2 Data sets

10 2.1 Aura MLS

11 The Microwave Limb Sounder (MLS) is one of four instruments on NASA's Aura satellite, launched on July 15, 2004. The
12 MLS antenna scans the atmospheric limb as Aura orbits the Earth in a near-polar sun-synchronous orbit, and the instrument
13 measures thermal emission (day and night) in narrow spectral channels, via microwave radiometers operating at frequencies near
14 118, 190, 240, and 640 GHz, as well as a 2.5 THz module to measure OH. MLS (see Waters et al., 2006) has been providing a
15 variety of daily vertical stratospheric temperature and composition profiles (~3500 profiles per day per product), with some
16 measurements extending down to the upper tropospheric region, and some into the upper mesosphere or higher. For more
17 information and access to the MLS data, the reader is referred to <http://disc.sci.gsfc.nasa.gov/Aura/data-holdings/MLS>; the
18 current data version is labeled 4.2x (with x varying between 0 and 3, depending on the date). Data users interested in MLS data
19 quality and characterization, estimated errors, and related information, should consult Livesey et al. (2018), the latest update to
20 the MLS data quality document (available from the MLS website at <http://mls.jpl.nasa.gov>).

21 2.2 GOZCARDS

22 The data considered here for longer-term model evaluation analyses is from GOZCARDS, a data record that was created
23 using high quality satellite-based Level 2 data as “source” data sets, which were merged together into global monthly zonal mean
24 records for O₃, H₂O, and HCl, going back in time before the (2004) launch of Aura. Readers are referred to the GOZCARDS
25 description and highlights provided by Froidevaux et al. (2015). In brief, for O₃, the original GOZCARDS version 1.01 (v1.01)
26 data record starts in 1979 with solar occultation measurements from the first Stratospheric Aerosol and Gas Experiment (SAGE
27 I), and continues with data from SAGE II, the Halogen Occultation Experiment (HALOE), the Upper Atmosphere Research
28 Satellite (UARS) MLS, the Atmospheric Chemistry Experiment Fourier Transform Spectrometer (ACE-FTS, using solar
29 occultation), and Aura MLS. Basically, the overlap time periods for different data sets are used to calculate offsets between zonal
30 mean time series in 10°-wide latitude bins at each pressure level, and the data sets are adjusted to a reference value (SAGE II
31 mean values, for O₃, or an average of satellite measurements for H₂O and HCl); monthly standard deviations are also provided,
32 along with other diagnostic quantities. GOZCARDS data extensions past 2012 were created simply by adding more recent MLS
33 data, appropriately adjusted to account for zonal mean differences between versions, once the MLS O₃ v2.2 data became
34 unavailable (for 2013 onward); the latest ACE-FTS data version was not included. GOZCARDS O₃ (v1.01) has been used for
35 past O₃ trend assessments and in comparisons to other data records (e.g., WMO, 2014; Nair et al., 2015; Tummon et al., 2015;
36 Harris et al., 2015; Ball et al., 2017, 2018).

37 For the GOZCARDS ozone data discussed here, unless otherwise noted, we use GOZCARDS v2.20, a recent improvement



1 and update to the original version. The ozone v2.20 data set was provided for an updated assessment of stratospheric ozone by
2 Steinbrecht et al. (2017), as well as for the assessment activities of the Long-term Ozone Trends and Uncertainties in the
3 Stratosphere (LOTUS) project and in preparation for the latest international report on the state of the ozone layer, led by the
4 World Meteorological Organization (WMO). GOZCARDS ozone data updates are also used as part of the yearly “State of the
5 Climate” stratospheric ozone-related summaries, produced for the Bulletin of the American Meteorological Society (BAMS). For
6 GOZCARDS O₃ v2.20, the stratospheric retrieval pressure grid is twice as fine as for v1.01; there are now 12 regularly-spaced
7 levels per decade change in log of pressure. The UARS MLS O₃ data were not included in v2.20, since these retrievals are not
8 readily available on the finer vertical grid (although approximations such as interpolation could be used); also, there is no easy
9 provision of UARS MLS retrieval uncertainties on a finer grid. The most significant change for the new merged ozone is the
10 effect of using the updated and more robust version 7 data from SAGE II (Damadeo et al., 2013). Version 7 uses National
11 Aeronautics and Space Administration (NASA) Global Modeling and Assimilation Office (GMAO) Modern-Era Retrospective
12 Analysis for Research and Applications (MERRA) temperature (T) profile data (Rienecker et al., 2011) in the retrievals, and
13 these values (rather than T from the National Centers for Environmental Prediction (NCEP)) also have a significant impact on
14 the conversion of SAGE II O₃ from its native density/altitude grid to the GOZCARDS mixing ratio/pressure grid. Also, Aura
15 MLS v4.2 O₃ data are now used (instead of MLS v2.2); HALOE v19 O₃ profiles are included, after interpolation to the finer
16 pressure grid before merging. As a result of these improvements in GOZCARDS O₃, we have observed closer agreement and
17 larger correlation coefficients between the Stratospheric Water and Ozone Satellite Homogenized (SWOOSH) data record (Davis
18 et al., 2016) and GOZCARDS v2.20 O₃ time series than between SWOOSH and GOZCARDS v1.01 (SWOOSH O₃ also uses
19 SAGE II v7 ozone data); more details regarding the impact of GOZCARDS v2.20 O₃ on trends are provided in Sect. 5 and in the
20 Appendix.

21

22 3 WACCM (CESM1) description and simulations

23 WACCM (CESM1) is a chemistry climate model of the Earth’s atmosphere, from the surface to the lower thermosphere
24 (Garcia et al., 2007; Kinnison et al., 2007; Marsh et al., 2013; Garcia et al., 2017). WACCM is a superset of the Community
25 Atmosphere Model, version 4 (CAM4), and includes all of the physical parameterizations of CAM4 (Neale et al., 2013) and a
26 finite volume dynamical core (Lin, 2004) for the tracer advection. The horizontal resolution is 1.9° latitude x 2.5° longitude. The
27 vertical resolution in the lower stratosphere ranges from 1.2 km near the tropopause to ~2 km near the stratopause; in the
28 mesosphere and thermosphere the vertical resolution is ~3km. Simulations used here are based on the guidelines from the
29 International Global Atmospheric Chemistry / Stratosphere-troposphere Processes And their role in Climate (IGAC/SPARC)
30 Chemistry Climate Model Initiative (CCMI) (Morgenstern et al., 2017). Improvements in CESM1 (WACCM) for CCMI include
31 a modification to the orographic gravity wave forcing, which reduced the cold bias in Antarctic polar temperatures (Garcia et al.,
32 2017; Calvo et al., 2017) and updates to the stratospheric heterogeneous chemistry, which improved the representation of polar
33 ozone depletion (Wegner et al., 2013; Solomon et al., 2015). In this work, there are two CCMI scenarios, spanning the 1990-
34 2014 period. The first scenario follows the CCMI REF-C1 definition and three ensemble members were completed; this falls
35 under the “free-running” scenario. We note that all the analyses herein are based on an average of these three simulations. We
36 have checked that the three representations’ departures from the average are small enough not to require separate comparisons
37 for each case, when pursuing average or root mean square (RMS) differences versus observations, in comparison to differences
38 using the 2nd model scenario (see below); this is also true for the model/data comparisons of RMS variability. This first model
39 scenario includes forcing from greenhouse gases (CH₄, N₂O, and CO₂), organic halogens, volcanic aerosol surface area density



1 and heating, and 11-year solar cycle variability. The sea surface temperatures are based on observations and the quasi-biennial
2 oscillation (QBO) is nudged to observed monthly mean tropical winds over 86–4 hPa, as described in Matthes et al. (2010).

3 The second scenario is based on the CCM1 REF-C1SD scenario and includes all the forcings of REF-C1, except for additional
4 external QBO nudging. This scenario uses the specified dynamics (SD) option in WACCM (Lamarque et al., 2012). Here,
5 temperature, zonal and meridional winds, and surface pressure are used to drive the physical parameterization controlling
6 boundary layer exchanges, advective and convective transport, and the hydrological cycle. The meteorological analyses are taken
7 from MERRA and the nudging approach is described in Kunz et al. (2011). The QBO circulation is inherent in the MERRA
8 meteorological fields and is therefore synchronized with that in the “real” atmosphere. The horizontal resolution is the same as
9 the REF-C1 version and the vertical resolution follows the MERRA reanalysis (from ~1 km resolution near the tropopause to
10 about 2 km near the stratopause). The model meteorological fields are nudged from the surface to 50 km; above 60 km, these
11 fields are fully interactive, with a linear transition in between. Both WACCM versions used here contain an identical
12 representation of tropospheric and stratospheric chemistry (Kinnison et al., 2007; Tilmes et al., 2016). The species included in
13 this mechanism are contained within the Ox, NOx, HOx, ClOx, and BrOx chemical families, along with CH₄ and its degradation
14 products. In addition, 20 primary non-methane hydrocarbons and related oxygenated organic compounds are represented, along
15 with their surface emissions. In total there are 183 species and 472 chemical reactions; this includes 17 heterogeneous reactions
16 on multiple aerosol types (i.e., sulfate, nitric acid trihydrate, and water-ice). For this work, the CESM1 (WACCM) REF-C1 and
17 REF-C1SD simulations will generally be referred to as FR-WACCM and SD-WACCM, respectively. While the runs were
18 originally designed to stop at the end of 2010, for this work, the forcing inputs have been extended through 2014.
19

20 4 Climatological comparisons and biases

21 We first describe some of the major climatological features for the various stratospheric species mentioned in the
22 Introduction. We focus here on the main differences between average model trace gas abundances in the FR-WACCM and SD-
23 WACCM runs and the corresponding abundances from Aura MLS for 2005 through 2014; this includes a sub-section on annual
24 and semi-annual variations. Further analyses of interannual variations and trends are discussed in Sect. 5.

25 4.1 Average abundances

26 We provide climatological latitude/pressure contour plots in the Supplement (Fig. S1 for O₃ and H₂O, and Fig. S2 for HCl,
27 HNO₃, and N₂O), showing Aura MLS and WACCM mixing ratio distributions, averaged over 2005 through 2014. Since such
28 plots do not easily allow one to quantify the largest areas of model/data disagreement, we show in Fig. 1 (left column, top two
29 panels) the percent differences between WACCM and the MLS climatologies; a positive value means that, on average, for 2005-
30 2014, the model values exceed the data values. Also, we show in the right column (top two panels) the absolute value of the
31 average model/data difference divided by systematic error estimates (2σ values) for MLS ozone. These error estimates have been
32 provided in past MLS validation and error characterization work as tabulated “typical” (global average) profiles as a function of
33 pressure; the latest update of such error estimates for version 4 MLS data is provided by the MLS team in Livesey et al. (2018).
34 The vertical profile of estimated systematic errors for MLS ozone is given in the Supplement (Fig. S3). Note that the MLS team
35 can provide further systematic error details for MLS data users (e.g., for pressure levels not listed in the standard Tables). Past
36 validation references for MLS ozone include Jiang et al. (2007), Froidevaux et al. (2008a), and Livesey et al. (2008), as well as
37 the more recent work covering many satellite (and other) instruments by Hubert et al. (2016). The original MLS data validation



1 work for H₂O is from Read et al. (2007) and Lambert et al. (2007), who also described N₂O validation, whereas HNO₃ validation
2 work was discussed by Santee et al. (2007).

3 The two bottom panels of Fig. 1 show a comparison between the two model runs (with percent differences on the left and the
4 ratio of the absolute differences versus data on the right). The main conclusion from Fig. 1 is that most of the model ozone
5 climatology falls within about 5 to 10% of the data climatology, except in the upper troposphere and lower stratosphere (UTLS),
6 where the SD-WACCM O₃ values are even larger than those from FR-WACCM (see bottom left panel). Our work focuses on the
7 stratosphere, but both FR-WACCM and SD-WACCM average O₃ values at low latitudes are lower than the observed mean
8 values from 215 to 261 hPa (MLS retrieval pressure levels that lie in the upper troposphere at these latitudes). While the right
9 column of Fig. 1 indicates that the difference to error ratio does not show a very significant systematic difference in this region,
10 there are known MLS positive biases versus tropical ozonesonde data, and this could account for at least part of the apparent
11 model low bias, at least for the 215 hPa level; between this level and the tropopause (near 100 hPa), the SD-WACCM model
12 values appear to be biased somewhat high. On the other hand, O₃ mid-latitude values from 100 to 215 hPa are biased high in SD-
13 WACCM in particular, with a difference to systematic error ratio larger than 2 to 3 in most of this region; otherwise, Fig. 1
14 shows that these ratios are typically less than 1 to 1.5. An illustration of the more significant differences is given in Fig. 2, for the
15 O₃ data and model time series at 215 hPa for 50°N-60°N (these abundances represent lower stratospheric values). This shows
16 that both FR-WACCM and SD-WACCM average values are larger than the data in this region, and more so for SD-WACCM. In
17 relation to this model overestimate of ozone, Imai et al. (2013) have shown that the SD-WACCM O₃ values are also larger than
18 those from the Superconducting Submillimeter-Wave Limb-Emission Sounder (SMILES) in the lowest portion of the
19 stratosphere (near 18-20 km). In Fig. 2, we also observe that the model ozone annual amplitude (AO) is larger than the observed
20 amplitude; we discuss the stratospheric AO more generally in the next section. If anything, MLS O₃ at mid-latitudes is biased
21 slightly high (by roughly 5%) with respect to a multi-instrument mean ozone field based on a large number of monthly zonal
22 mean satellite data sets from the Stratosphere-Troposphere Processes And their Role in Climate (SPARC) Data Initiative (DI), as
23 discussed by Tegtmeier et al. (2013), who also showed that the MLS O₃ seasonal cycle at mid-latitudes and 200 hPa is in very
24 good agreement with the multi-instrument mean. The limb emission measurements from Aura MLS for the species discussed
25 here provide generally strongly peaked vertical averaging kernels with a resolution of 2.5-4 km (see Livesey et al., 2018, for
26 sample plots of these kernels). We have confirmed that smoothing the model profiles using the MLS averaging kernels (and a
27 priori information) gives very little change (less than a few % for the Fig. 2 example, and even less at higher altitudes) in O₃
28 abundances and seasonal cycles, and we see no real need to use such a smoothing for the model/data comparisons herein; the
29 significant model/data differences in Figs. 1 and 2 are thus not caused by this sort of issue.

30 Figure 3 is the H₂O analog of Fig. 1, for pressures reaching up to 0.01 hPa, since the Aura MLS H₂O retrievals cover both the
31 stratosphere and the mesosphere; again, a typical profile of MLS systematic error estimates (Livesey et al., 2018) has been used
32 in deriving the results shown in the top two right panels, which test for the existence of a significant bias between model and
33 observations. We note that MLS H₂O version 3 stratospheric data generally exhibit a slight high bias (of a few to 5%) versus
34 multi-instrument mean values, with a somewhat larger positive bias (of ~10%) in the lower mesosphere (Hegglin et al., 2013).
35 Such biases are within the expected measurement systematic errors; MLS version 4 stratospheric H₂O data show essentially no
36 systematic change in comparison to version 3 (Livesey et al., 2018). FR-WACCM and SD-WACCM H₂O mean values are on the
37 low side (by ~5-15%) relative to MLS H₂O in the upper stratosphere and in most of the mesosphere (see Fig. 3), implying that
38 the models are in good agreement with the SPARC DI multi-instrument mean H₂O. There is independent evidence that MLS
39 H₂O has a dry bias near the hygropause (at the low end of the vertical range shown in Fig. 3, where the bottom level is 150 hPa);
40 this has been known for some time (see Read et al., 2007; Vömel et al., 2007). This is also consistent with the existence of a



1 model high bias relative to MLS near 150 hPa, assuming that the models come close to representing the H₂O climatology in this
2 region. In terms of the significance of the biases, the right top two panels of Fig. 3 indicate that the H₂O model/data comparisons
3 are generally in agreement within the estimated (2σ) systematic errors, and that this level of agreement is slightly better in the
4 case of SD-WACCM.

5 For HCl, the climatological comparisons of Fig. 4 show that both models exhibit a small (5-10%) low bias versus MLS HCl
6 in much of the stratosphere, with a stronger negative model bias in the tropical region between 100 and 150 hPa. The model/data
7 relative biases in stratospheric HCl are generally within the MLS HCl systematic errors (see Fig. 4, top two panels at right). MLS
8 HCl is slightly on the high side of the multi-instrument mean climatological results provided in the SPARC DI report (SPARC,
9 2017). The small negative model bias in the upper stratosphere could also arise from the lack of a sufficiently pronounced
10 decrease in upper stratospheric MLS HCl, as a result of the interruption in the main MLS HCl target band (band 13) data after
11 early 2006 (see Livesey et al., 2018). There is also a known strong positive systematic bias in MLS HCl at 150 hPa in the tropics
12 (see Froidevaux et al., 2008b), so model underestimates in this region are not a sign of model weakness. We also note that both
13 models exhibit a systematic difference versus HCl observations in the lower stratosphere (with larger differences for SD-
14 WACCM), as well as a downward sloping pattern (equator to pole) in the southern hemisphere (SH), and smaller mean
15 differences (for SD-WACCM) in the northern hemisphere (NH).

16 Figure 5 provides average comparisons for N₂O (a good dynamical tracer). While the mean lower stratospheric SH N₂O
17 values are larger for SD-WACCM than for FR-WACCM (bottom left panel of Fig. 5), the mean absolute fit for SD-WACCM
18 versus MLS is not significantly better. The most significant climatological differences with respect to the error bars (top two
19 right panels in Fig. 5) are in the upper stratosphere at low latitudes; in this region, SD-WACCM agrees somewhat better with
20 MLS. However, this is also where the mean abundances decline rapidly with height towards the limits of the MLS sensitivity.
21 Not too surprisingly, this is also where the SPARC DI results for N₂O show the largest scatter in terms of percent differences
22 (often exceeding 10-20%, see SPARC, 2017).

23 Finally, the climatological comparisons for HNO₃ (Fig. 6) reveal very few areas of mean model/data disagreements
24 significantly outside the systematic uncertainties. However, there is a general model underestimation, especially in the polar
25 upper stratosphere. We will see later that there are large recurring model/data differences in the upper stratosphere during certain
26 months. At high latitudes in the lower portion of the stratosphere, the models tend to overestimate the data. The upper
27 troposphere is where the satellite-based HNO₃ data have been validated the least, but there is some evidence for a high MLS bias
28 in this region, based on SPARC DI results (see SPARC, 2017). While this might explain, at least qualitatively, why the models
29 are observed to underestimate MLS tropical UT HNO₃ (see Fig. 6), more work is needed to better evaluate HNO₃ from models
30 and satellite-derived data sets in this particular region.

31 It is also interesting to point out modeled and observed seasonal changes in the polar lower stratosphere over Antarctica.
32 Figure 7 depicts average seasonal changes over the 70°S-80°S region at 46 hPa for 2005-2014, using Aura MLS observations
33 and model values for comparison. We wish to emphasize the slope of the early winter decline in HCl, and we see that the model
34 HCl values do not decline as fast as indicated by the data, even though SD-WACCM tracks the interannual variability better than
35 FR-WACCM does (see Fig. S4 for the relevant time series from 2005 through 2014). Many of the options and uncertainties
36 regarding lower stratospheric heterogeneous chemistry modeling for SD-WACCM at high latitudes in the polar winter/spring
37 have been discussed by Solomon et al. (2015), who have pointed, for one specific year (2011), most of the features depicted in
38 Fig. 7. For our broader time period, we see that the HCl rate of decline from May to July (dominated by nighttime conditions) is
39 slower in both the SD-WACCM and FR-WACCM results than the corresponding mean HCl rate of change from MLS (top left
40 panel of Fig. 7). Grooß et al. (2018) have recently discussed this HCl model/data discrepancy for dark polar vortex conditions,



1 including the potential impact of numerical diffusion issues in the Eulerian models, in comparison to their simulations using the
2 Chemical Lagrangian Model of the Stratosphere (CLAMS), which shows even larger HCl discrepancies. These authors discuss
3 some possible mechanisms and uncertainties, and they argue that additional decomposition of condensed-phase HNO₃ might
4 play a role, possibly via galactic cosmic ray impacts. Since this rapid decline in HCl occurs during polar night, these authors
5 point out that this early HCl issue does not lead to much difference in polar ozone loss rates, which only become significant
6 during sunlit conditions (early spring). Figure 7 confirms that, on average, the SD-WACCM O₃ decline and rise match the data
7 well. We also note that FR-WACCM shows smaller-than-observed declines in HNO₃ and H₂O, whereas SD-WACCM matches
8 these observations much better. The temperature panel (bottom center) gives a possible reason for these differences, as T from
9 FR-WACCM is higher by a few degrees during the coldest phase than T from SD-WACCM (MERRA-based), and larger than
10 the MLS-derived values. This would lead to less irreversible denitrification and dehydration. The general nature of the results
11 shown in Fig. 7 is similar at other lower stratospheric pressures (and for latitudes poleward of 80°S), although there is some
12 variability in the magnitude of the differences. Over the Arctic region (not shown here), temperature-related differences are not
13 as systematic or as large as over Antarctica, but similar model/data differences in the early winter rate of decline in HCl exist
14 there also (as mentioned by Grooß et al., 2018).

15 As an addendum regarding the evaluation of models in comparison to data sets, we provide in Appendix (A1) the results of a
16 model grading approach that has been used in the past (e.g., Douglass et al., 1999, Waugh and Eyring, 2008). We find (see
17 Appendix A1) that this grading method often leads to low grades (see Fig. A1), if applied using systematic uncertainty estimates
18 from the MLS data characterization work (Livesey et al., 2018). The model results, as good as they are in many respects, cannot
19 always match the data closely enough, at least based on such grades (although the grading formulation could also be
20 reconsidered). As mentioned in the Appendix, the multiplicative error factor (see Equation A1) can be increased (e.g., from 2 to
21 4) to force these grades (see Fig. A2) to span a more useful range (in the plots). Indeed, we observe similarities between Fig. A2
22 (top two panels) and Fig. 1 (middle panels): poorer SD-WACCM performance for pressures ≥ 100 hPa, better mid-latitude
23 results near 30-40 hPa, and poorer performance again near 3 to 5 hPa. Also, we observe the poorest H₂O grades near 1 hPa in
24 Fig. A2 (bottom panels), which is similar to the poorer performance in the middle panel of Fig. 3, while the best stratospheric
25 H₂O grades are found near 10-20 hPa, which matches the best performance (smallest values) in Fig. 3 (top two panels at right).
26 In the case of merged data records, we note that it is generally more difficult to estimate systematic errors. The GOZCARDS
27 data analyses led to conservative (i.e. possibly somewhat pessimistic) systematic error estimates as a function of latitude and
28 pressure (see Froidevaux et al., 2015) that are significantly larger than the systematic error estimates for MLS data only. There
29 are other methods that could lead to useful error estimates, through the use of multi-satellite data sets and the spread between
30 these (see SPARC, 2017), for some species at least. Moreover, when one considers relative variations such as anomaly time
31 series (as done in a subsequent section), it becomes even less clear how to best assign uncertainties in the context of “error-
32 weighted” grades; some data records may also drift with respect to others, or with respect to ground-based data, so that the actual
33 errors will change with time (and possibly with location as well), in a difficult to determine way. We do not pursue this more
34 traditional grading approach further here, especially as the two models we are comparing often lie fairly close together. For
35 multi-model comparisons (which is outside the scope of this work), one could consider how to best apply grading methods such
36 as the one in Appendix A1, along with other diagnostics such as those discussed here; in the end, the most important aspect of
37 such analyses possibly lies in the relative values of the grades or diagnostics for different models.



1 4.2 Annual and semi-annual cycles

2 Figure 8 displays the amplitudes of annual and semi-annual variations for MLS ozone and the corresponding FR-WACCM
3 and SD-WACCM runs for 2005-2014. We obtained these results from a simple regression fit to the monthly mean time series in
4 each latitude/pressure bin. The primary time dependence of the fitted function is given by additive sine and cosine terms (with
5 12-month and 6-month periods), in addition to constant and linear trend terms; the AO and SAO amplitudes are given by the
6 square root of the sum of the squares of the corresponding fitted coefficients. We see from Fig. 8 that the overall data and model
7 patterns of AO and SAO variability are quite similar. The ozone AO amplitudes peak (in ppmv) at mid- to upper stratospheric
8 levels, with high latitude variations also observed as a result of the effects of winter/spring polar chemistry and dynamics; the
9 lower stratospheric peak AO amplitudes are more prominent over the southern polar regions, where stronger O₃ depletion occurs
10 on a seasonal basis. These MLS O₃ AO patterns are very similar to those obtained by Schoeberl et al. (2008), using a much
11 shorter time period (Sep. 2004 to Dec.2006); the same holds for other species (H₂O and HCl) considered in that work and here.
12 The observed SAO amplitude for O₃ exhibits strong peaks in the upper stratosphere, both in the tropics and at high latitudes. The
13 anti-correlation between O₃ and temperature as a result of temperature-dependent photochemical production and loss terms for
14 O₃ has long been known to cause most of the O₃ variability in the upper stratosphere (see Perliski et al., 1989, in relation to the
15 AO and the SAO). The AO and SAO amplitudes obtained in that and other past studies (e.g., see Ray et al., 1994) are very
16 similar to the amplitude patterns shown here. If we look more closely (and based on AO amplitude ratio plots not shown here),
17 there are often O₃ AO amplitudes 20-80% larger than those derived from MLS data for both WACCM runs in the lower
18 stratosphere (from 50 hPa at low latitudes to 215 hPa at high latitudes); such a model overestimate of the AO amplitude was
19 shown in the time series example of Fig. 2. More generally, outside of this region, we observe somewhat closer fits to the MLS
20 AO amplitudes for the SD-WACCM version, but both models track the data and each other well, with AO amplitudes typically
21 within ~25% of the MLS AO amplitudes. For the Antarctic lower stratosphere, the timing and magnitude of the seasonal
22 recovery after the ozone hole plays a role, and in this respect, we have observed that SD-WACCM generally fits the MLS data
23 better than FR-WACCM does. In the tropical upper stratosphere, where the SAO is larger than the AO (see Fig. 8), the SD-
24 WACCM results match the observed SAO amplitudes slightly better than those from FR-WACCM. Despite the existence of a
25 few model/data differences, these AO and SAO amplitude comparisons, coupled with our examination of model/data amplitude
26 ratio plots as well as the time series (which include the phase information), do not elicit major concerns regarding the model
27 characterization of the primary processes expected to govern these modes of O₃ variability. The study of dynamical forcing
28 mechanisms in relation to such modes continues to be an active area of research (e.g., see Ern et al., 2015, for a discussion of
29 wave-driving and the SAO). Also, Smith et al. (2017) have recently shown that analyses of Aura MLS geopotential height data
30 lead to derived tropical zonal mean winds that agree well with those derived from Sounding of the Atmosphere using Broadband
31 Emission Radiometry (SABER) geopotential heights, and with direct wind data, thus enhancing our knowledge of tropical
32 atmospheric dynamics (including the SAO and QBO).

33 For H₂O, a similar overview of the AO and SAO amplitudes is given in Fig. 9, which covers the vertical range from 100 to
34 0.01 hPa. A peak in these amplitudes resulting from the seasonal downward transport before the winter, followed by wintertime
35 dehydration, is observed in the lower stratospheric southern polar region; we note that the SD-WACCM results match this
36 feature better than the FR-WACCM simulation does. Other interesting features include the southern hemisphere's upper
37 stratospheric AO peak in the extra-tropical region. This has been seen by many satellite-based measurements, as discussed in the
38 comparisons by Lossow et al. (2017a). Lossow et al. (2017b) have explained this "island" feature in more detail, with the help of
39 model simulations and analyses; they argue that vertical advection tied to the upper branch of the Brewer-Dobson circulation
40 largely explains the seasonal highs (lows), via downwelling (upwelling). They also show that an AO maximum is observed as



1 well in other species in roughly the same region, including in N₂O MIPAS data; we confirm this behavior (see also Fig. S5) from
2 the N₂O AO amplitude feature observed in MLS data, as well as in the WACCM runs (and more so in SD-WACCM). The
3 derived AO and SAO amplitude patterns in H₂O from Lossow et al. (2017a) are consistent with what we show here; this includes
4 the peak values in the upper stratosphere and mesosphere, attributed to the combined effects of photochemistry and vertical
5 transport. For ozone, a dominant feature in the SAO amplitude exists in the tropical upper stratosphere; see Lossow et al. (2017a)
6 for a brief review of past work explaining such dynamically-driven features for H₂O. While there is generally a good level of
7 model/data agreement in the main H₂O AO and SAO patterns, both WACCM comparisons tend to underestimate observed AO
8 and SAO amplitudes in the lower stratosphere and overestimate AO amplitudes in the SH mesosphere, while slightly
9 underestimating the mesospheric SAO amplitudes in both polar regions. The largest amplitude differences reach a factor of two,
10 in places, for the lower stratospheric model underestimates, which cannot be caused by slight model underestimates of average
11 MLS H₂O (as seen in Fig. 3); however, we note that the lower stratosphere is also the region where AO and SAO amplitudes are
12 smallest (typically < 0.1 ppmv).

13 Turning briefly to the other species discussed here, for N₂O, we have already mentioned the existence of the upper
14 stratospheric AO peak (the “island” feature described by Lossow et al., 2017b) in the southern hemispheric extra-tropical region.
15 This is similar to the H₂O AO amplitude maximum feature, but the N₂O seasonal variations are anti-correlated with H₂O, as
16 demonstrated by Lossow et al. (2017b), using MIPAS data (and as is also apparent in the MLS time series, not shown here).
17 Furthermore, we observe in Fig. S5 a somewhat better match in the AO and SAO N₂O amplitude patterns for SD-WACCM than
18 for FR-WACCM versus MLS, in particular for tropical to southern mid-latitudes. We note that this is also manifested in better
19 time series fits for SD-WACCM, besides the closer match in average values; as mentioned in the Introduction, this is what one
20 would generally expect for the two model versions. We saw also in Fig. 5 that FR-WACCM overestimates the average values of
21 upper stratospheric tropical N₂O. For HCl and HNO₃, the AO and SAO amplitudes are dominated by the variations at high
22 latitudes (see Figs. S6 and S7). The model HCl upper stratospheric AO and SAO amplitudes match up fairly well with the
23 observed amplitudes, despite the aforementioned issues relating to MLS upper stratospheric HCl trends. The NH lower
24 stratospheric polar variations are somewhat underestimated by the models, and more so by FR-WACCM run, which (based on
25 time series plots) exhibits larger values at the low end and smaller values at the high end of the HCl summer to winter cycle. For
26 HNO₃, the main AO and SAO model features follow the MLS patterns, although there is a model underestimation of the
27 amplitudes in the upper stratospheric polar regions, because these models do not properly capture the observed recurrences of
28 enhanced HNO₃, as mentioned earlier (see also Sec. 4).

29

30 5 Time series comparisons

31 5.1 Anomaly time series: fits and variability

32 5.1.1 Fits

33 We wish to evaluate which of the two WACCM models (free-running or specified dynamics) provides a better match, or fit,
34 to the temporal variations in observed deseasonalized anomalies. Again, we would expect SD-WACCM to generally fit these
35 anomalies better than FR-WACCM. We analyze the model and data deseasonalized anomaly time series, obtained by
36 differencing each month’s zonal mean value from the long-term averages for the same month. We then calculate a diagnostic of
37 model fit to the data, by using the RMS differences between these deseasonalized model and data series, and normalize by
38 dividing this quantity by the RMS of the data anomalies themselves. Thus, a diagnostic value that is much less than unity means



1 that the match to the time series is much smaller than the typical variability; this also implies a good fit to the observed
2 anomalies. In the Appendix (A2), we provide the mathematical expression for this “RMS difference diagnostic”. A better model
3 fit to the observed anomaly series will be represented by a smaller “RMS difference diagnostic” value. We also calculate the
4 standard (Pearson) correlation coefficients, R , between model and data anomalies, and we use R^2 as another measure of
5 “goodness of fit” for the models. The first diagnostic is unitless and does not depart too much from the 0 to 1 range; R^2 is limited
6 to the 0 to 1 range, with larger values indicating a higher degree of linear correlation. We also use these two diagnostics together,
7 by calculating the ratio of R^2 over the RMS difference diagnostic to obtain a “combined diagnostic”. This diagnostic could have
8 a large value (and a good model result) from both a large R^2 value (in the numerator), meaning a high correlation with observed
9 anomalies, and a small RMS difference (in the denominator), implying a good fit to these observations. This also tends to
10 amplify the differences between two model comparisons to the same data series. An ideal model fit would correlate tightly in
11 time to observed oscillations, but also exhibit the right magnitude for these variations by “hugging” the anomaly series without
12 being off in magnitude. Indeed, two model series could have oscillations in phase with data variations and thus the same R^2
13 values, but with different amplitudes (and different overall fits); conversely, two model series could have different R^2 values
14 versus observations, if one is more out of phase than the other, but they could still produce similar RMS difference fits.

15 In Fig. 10, we display latitude/pressure contour plots of the above diagnostics for FR-WACCM and SD-WACCM O_3
16 anomalies in relation to MLS anomalies for 2005-2014. We immediately see from the top two panels that the SD-WACCM RMS
17 difference diagnostic values (in the 0.2 to 1 range in the stratosphere) are smaller than those from FR-WACCM (with typical
18 values between 0.8 and 1.2). Values of R^2 (middle panels) show that SD-WACCM also correlates very well with the
19 observations from 2005-2014, with values typically between 0.7 and 0.95 in most of the stratosphere, and somewhat poorer
20 correlations in the UTLS. The FR-WACCM ozone series tend to correlate fairly well with the data at low to mid-latitudes for
21 pressure levels between about 70 and 7 hPa, which is also where the FR-WACCM RMS difference diagnostic shows better
22 performance than in other regions, and (as an explanation) where the dynamics are nudged (to tropical winds) in a similar way as
23 for SD-WACCM. However, FR-WACCM shows poor performance (almost zero correlation) at high latitudes and in all of the
24 uppermost stratosphere.

25 Figure 11 shows sample O_3 series for the upper stratosphere (at 2.2 hPa) for 0° - 10° N as well as 40° N- 50° N. Although the
26 tropical series show good correlations overall for both models versus data, the details of observed semi-annual peaks and the
27 interplay between the AO, SAO, and the quasi-biennial oscillation (QBO) are better followed by the SD-WACCM curve. The
28 differences in O_3 amplitude and phase are more clearly displayed in the bottom (left) panel, which shows the deseasonalized
29 anomalies. Diagnostic values provided in this panel for both models distinctly show that SD-WACCM performs better than FR-
30 WACCM here, with a much larger R^2 value and a smaller RMS difference diagnostic value, and hence, a much better (larger)
31 combined diagnostic value. The same comments apply to the right two panels of Fig. 11, which showcase the NH mid-latitudes
32 at 2.2 hPa. In the high latitude lower stratosphere, the poorer FR-WACCM results in Fig. 10 are generally caused by time series
33 that are observed to be less in-phase with the polar winter/spring variations, as well as by more departures in the magnitude of
34 such variations; we will return later to sample polar time series for ozone and other species. The bottom two panels in
35 Fig. 10 amplify the differences between the models, with values of the combined diagnostic below 1 in most regions for FR-
36 WACCM, but values between 1 and 3 for the whole stratosphere in the SD-WACCM case. The more realistic dynamics in SD-
37 WACCM, coupled with the same chemistry as FR-WACCM, allow for better SD-WACCM fits to the data, as shown
38 quantitatively in Figs. 10 and 11. These plots also point to poor results for both models in the upper troposphere, albeit with
39 somewhat better results for SD-WACCM, although this is not the focus of this paper.



1 We now turn to Fig. 12 for a description of the same diagnostics as above, but for model/data H₂O comparisons, and a top
2 pressure level at 0.01 hPa. We observe, again, that the diagnostics of fit are usually much better for SD-WACCM, which yields
3 R² values of 0.6 to 0.9 and RMS difference diagnostic values below 1 for most of the stratosphere and lower mesosphere, and
4 therefore better combined diagnostic results than FR-WACCM as well. The SD-WACCM diagnostics themselves worsen in the
5 upper mesosphere and in the high latitude regions near 215 hPa. In the upper mesosphere, the better diagnostic results for SD-
6 WACCM are seen in the time series (not shown here) as a better match versus the MLS H₂O series anomalies in terms of the
7 interannual variability at all latitudes, as well as some seasonal peaks at high latitudes; we interpret this as the result of a better
8 dynamical representation of the mesosphere for SD-WACCM. We note that the high-quality representation of mesospheric
9 composition by SD-WACCM is also demonstrated in comparisons to measurements of CO profiles above Kiruna, Sweden, by
10 MLS and the Kiruna Microwave Radiometer (Ryan et al., 2018). For H₂O near 200 hPa, poor fits at high latitudes occur where
11 MLS is known to have a significant dry bias versus sonde and Atmospheric Infrared Sounder (AIRS) data, as discussed
12 previously in MLS data documentation (Livesey et al., 2018), as well as by Vömel et al. (2007) and Davis et al. (2016). MLS
13 H₂O is low by a factor of several here versus the WACCM runs (which show values of 10–60 ppmv). The data variability may
14 also be affected by the dry bias retrieval issues at the lowest altitudes for these high latitude regions, where observed anomalies
15 are more poorly tracked by the models; a planned future update to the MLS H₂O retrievals might help to mitigate this
16 discrepancy.

17 Figure 13 displays results similar to Fig. 10 but for stratospheric HCl. SD-WACCM HCl results versus MLS are also
18 generally superior to those from FR-WACCM and show high correlations (R² > 0.7) in most of the lower stratosphere, with
19 somewhat poorer results in the upper stratosphere, where the RMS difference fits as well as the combined diagnostic are poor for
20 both models. We trace the upper stratospheric issue back to a known data problem in this region, where MLS HCl trends are too
21 flat and the models depart more from the observations (down to about 10 hPa), notably in the RMS difference diagnostic. Poorer
22 correlations are observed for FR-WACCM at high latitudes, which we will return to later for the lower stratosphere, but even
23 fairly subtle differences in the timing (phase) can lead to significantly poorer correlations. The poorer fits at 100 hPa in the deep
24 tropics are related to a large underestimate of the mean data, which may be caused, at least in part, by an MLS high bias
25 (Froidevaux et al., 2008b), but more so, for R², by some out-of-phase variability as well. It will be difficult to resolve this issue
26 until more realistic (lower) MLS HCl values are obtained, and as the phasing may also change with new retrievals.

27 For the dynamical tracer N₂O, we also observe in Fig. 14 (showing model comparisons to the 68 to 1 hPa observations from
28 the 190 GHz MLS N₂O band for 2005–2014) that SD-WACCM fits the data better than FR-WACCM, in both the R² and the
29 RMS difference categories. FR-WACCM exhibits poor results in the upper stratosphere and at high latitudes in the lower
30 stratosphere. Both models exhibit poorer RMS fits and poorer correlations in the tropical lower stratosphere. Partly, this appears
31 to be caused by a model underestimation of the MLS N₂O variability in this region, with some QBO phasing differences as well.

32 HNO₃ results (see Fig. 15) show, again, better fits to the stratospheric MLS data from SD-WACCM than from FR-WACCM,
33 and poor performance from FR-WACCM at high latitudes. Both models do poorly in the upper stratosphere, and Fig. 16
34 illustrates the magnitude of this discrepancy in the region (3.2 hPa and 70°S–80°S) where it reaches its maximum, in terms of
35 mixing ratio values. Since its launch, MLS has been observing very large values of HNO₃ in the upper stratosphere, mostly in the
36 polar regions during winter. The WACCM runs used here do not have the right chemistry in the mesosphere and upper
37 stratosphere to adequately represent such variations; implementation of the many necessary missing chemical reactions has not
38 made its way into most CCMs. The solution is believed to be tied to ion cluster chemistry during energetic particle precipitation
39 (EPP) events, which includes large solar proton events (SPEs) as well as more regular auroral-type activity. The study of upper
40 stratospheric NO_x enhancements tied to auroral activity and other EPP events has a long history based on other satellite



1 measurements and modeling (e.g., Kawa et al., 1995; Callis and Lambeth, 1998; Siskind et al., 2000; Orsolini et al., 2005;
2 Randall et al., 2007; Reddmann et al., 2010). Yearly upper stratospheric enhancements in ground-based microwave retrievals of
3 HNO₃ profiles over Antarctica were discussed by de Zafra et al. (1997) and de Zafra and Smyshlaev (2001). Direct high altitude
4 EPP effects enhance NO_x, which can propagate downward in polar winter and increase stratospheric NO_x and HNO₃ via this
5 indirect effect and conversion of N₂O₅ on ion water clusters (Böhringer et al., 1983). Large polar enhancements in upper
6 stratospheric HNO₃ (and other species) were observed by the Michelson Interferometer for Passive Atmospheric Sounding
7 (MIPAS) after significant SPE activity in 2003, as presented in several papers (Orsolini et al., 2005; von Clarmann et al., 2005;
8 Lopez-Puertas et al., 2005; Stiller et al., 2005). More complex modeling using modified chemistry and transport-related effects
9 (e.g., Jackman et al., 2008; Funke et al., 2011; Verronen et al., 2011; Kvissel et al., 2012; Andersson et al., 2016) has produced
10 EPP-induced enhancements in high latitude HNO₃, with related improvements in model/data comparisons into the mesosphere,
11 and in comparisons of other species. Regarding the low latitude upper stratospheric HNO₃ comparisons, the poorer model fits
12 (even for SD-WACCM) seem to be caused at least in part by more noisy and variable MLS data, under low HNO₃ conditions.
13 Finally, tropical MLS HNO₃ data at 147 and 215 hPa are not fit well by either model, as the data exhibit significant seasonal
14 oscillations between 0.2 and 0.5 ppbv (with larger amplitudes occurring at 147 hPa), whereas model values are smaller than 0.1
15 ppbv. There have been very few tropical UT validation comparisons for HNO₃ (see Santee et al., 2007), but in situ HNO₃ data
16 from an airborne chemical ionization mass spectrometer have indicated that UT HNO₃ tropical mixing ratios are mostly below
17 0.1-0.2 ppbv (Popp et al., 2007, 2009).

18 For a more in-depth look at the lower stratosphere over Antarctica, Fig. 17 displays anomaly time series comparisons for O₃
19 and temperature at 68 hPa and 70°S-80°S, along with the diagnostic quantities that we are using. We note (top panel) the poorer
20 O₃ correlations for FR-WACCM ($R^2 = 0.22$) than for SD-WACCM ($R^2 = 0.89$), as well as the poorer RMS difference values
21 (0.89 for FR-WACCM versus 0.33 for SD-WACCM), with correspondingly poorer results in the FR-WACCM combined
22 diagnostic (0.25) versus SD-WACCM (2.67). Similar differences in the diagnostics are obtained in the bottom panel for
23 temperature (T) anomalies, showing excellent agreement between SD-WACCM and retrieved temperature anomaly time series
24 from MLS, both for the R^2 and RMS fit diagnostics. One should not expect the free-running model, even if it has certain tropical
25 QBO-related constraints that mimic those from the SD version, to perform as well in terms of predicted high latitude
26 temperatures as the SD-WACCM run, driven by realistic (MERRA) meteorological winds, as well as temperatures, which match
27 up closely to the observed temperatures from MLS. These plots also show that springtime anomalies dominate the variability,
28 with warmer than usual springs (during October in particular), such as 2012 and 2013, leading to more positive ozone anomalies,
29 i.e. less ozone depletion; conversely, years (2006, 2008, 2010, 2011) with colder than usual springtime conditions are correlated
30 with negative ozone anomalies and more depleted conditions. Other factors (besides local mean temperatures) can significantly
31 influence interannual variability and longer-term ozone loss over Antarctica; this includes the strength of the vortex, total
32 chlorine abundances, the phase of the QBO, tropospheric wave driving, the timing of warming events, and the impact of aerosols
33 (e.g., Scaife et al, 2005; Parrondo et al., 2014; Strahan et al., 2015; Langematz et al., 2016; Solomon et al., 2016). We saw in
34 Sect. 4 that there are generally good comparisons between SD-WACCM and MLS variations over the Antarctic region during
35 polar winter/spring, except for the rate of HCl decline during early winter; also, poorer results are obtained by FR-WACCM. We
36 also find (not too surprisingly) that interannual differences in lower stratospheric chemical evolution over Antarctica are not as
37 faithfully reproduced by FR-WACCM as by SD-WACCM, although we do not show more related details here.



1 5.1.2 Variability

2 Given our expectations that SD-WACCM would match up better than FR-WACCM to the observed time series of multiple
3 species, and having demonstrated this in the previous section, we turn to what should be a more fair comparison between the two
4 sets of model results, namely the variability aspect. We calculate the ratio of model to data interannual variability, as obtained
5 from the root mean square values of deseasonalized monthly anomaly time series, expressed as a percent of climatological (full
6 period) means; a simple linear trend is first subtracted from the time series, so that the variability comparisons remove any
7 significant trend differences. We do this for the MLS data considered previously (starting in 2005), but also for longer-term time
8 series, using the GOZCARDS data records. The models are sampled monthly following the monthly sampling of the data sets
9 (but not at the daily sampling level of detail); sampling plays a role for the longer-term (merged) GOZCARDS data, which are
10 comprised of some unevenly-sampled occultation data records (depending on latitude and pressure). Figure 18 compares the
11 ozone variability ratios for models versus data using as a reference the MLS 2005-2014 data (top two panels), and the
12 GOZCARDS merged ozone (1992-2003) data (bottom two panels). To first order, we observe similar patterns for both time
13 periods. The SD-WACCM variability is generally within 10-20% of the data variability (ratio values between 0.8 and 1.2). The
14 FR-WACCM variability is generally somewhat smaller than the data variability in the polar regions and in the upper
15 stratosphere. In a recent study, Bandoro et al. (2017) also found that the free-running version of WACCM displays somewhat
16 smaller ozone variability in the upper stratosphere, both for shorter-term and longer-term variabilities, than the observed
17 variability, based on the merged SWOOSH O₃ data record. In the lower stratosphere at low to mid-latitudes, we observe that FR-
18 WACCM exhibits slightly larger variability than the data, whereas SD-WACCM shows slightly smaller variability than the data;
19 Bandoro et al. (2017) found that FR-WACCM slightly overestimates the decadal variability in this region. If a free-running
20 model exhibits significant differences versus observed variability, this has some implications regarding predictions of trend
21 detection feasibility, as free-running models are used for such predictions; indeed, trend error bars increase if the variability
22 increases. We also see in Fig. 18 that at high latitudes, FR-WACCM tends to underestimate the actual variability, which means
23 that interannual swings in lower stratospheric ozone, in particular, are more muted in the model, whereas this is less of an issue
24 for SD-WACCM, with its more realistic representation of the dynamics; as an example, refer to Fig. 16, for model and data
25 anomalies at 68 hPa and 70°S-80°S.

26 A similar overview of model/data variability ratios is provided for H₂O in Fig. 19, which covers both the mesosphere and the
27 stratosphere. In this case, while the variability from SD-WACCM is somewhat closer than that from FR-WACCM to the data
28 variability during both time periods, the tendency for both models is to underestimate the observed variability, with FR-
29 WACCM exhibiting a stronger underestimate in the upper stratosphere and mesosphere. A variability underestimate for FR-
30 WACCM implies that any trend detection in the future will require more years of data in the real atmosphere, if H₂O continues to
31 have larger variability than model expectations. Also, this will translate into smaller estimated uncertainties in the model-derived
32 trends in comparison to the observations (as we will see in the next section on trends). The FR-WACCM underestimate of the
33 variability is sometimes by as much as a factor of two, although it is more typically by ~30% (see Fig. 19). For a time series with
34 RMS variability about the fit represented by σ , the number of years needed to statistically detect a trend is proportional to $\sigma^{2/3}$
35 (Weatherhead et al., 1998), and thus, an increase of σ by factors of 1.3, 1.5, and 2.0, for example, will lead to an increase in the
36 number of years for trend detection by factors close to 1.2, 1.3, and 1.6, respectively. In the tropical lower stratospheric case,
37 H₂O and temperature values and anomalies for 1992-2014 are shown for 100 hPa and 10°S-20°S in Fig. 20. Again, we note the
38 smaller-than-observed variability in the model H₂O oscillations, with SD-WACCM tracking the data better. This correlates with
39 the temperature series, where smaller variability is seen in FR-WACCM, in comparison to SD-WACCM (which follows the
40 MERRA temperatures); we also note that FR-WACCM temperatures are somewhat larger (by ~1K on average) than SD-



1 WACCM temperatures in this region. It is well known that stratospheric entry level H₂O is governed by temperatures near the
2 tropopause “cold trap”; the monthly average variations shown here are similar to what has been shown in past H₂O work (e.g.,
3 Randel et al., 2004, 2006; Randel and Jensen, 2013). Brinkop et al. (2016) used model runs from both free-running and nudged
4 simulations to analyze the impacts of different constraints, including sea surface temperatures (SST) and meteorological fields,
5 on “sudden” drops in water vapor; they found that several of these factors play a role in the H₂O variations, including the timing
6 of ENSO and SST variability, the phasing with the QBO, cold point temperatures, as well as the correct dynamical model state.
7 Many other analyses of the relation between entry level H₂O, tropopause temperatures, transport, and convection have been
8 carried out previously (e.g., Holton and Gettelman, 2001; Jensen and Pfister, 2004; Fueglistaler and Haynes, 2005; Rosenlof and
9 Reid, 2008; Read et al., 2008; Schoeberl et al., 2013). Our point here is that the WACCM H₂O anomaly series underestimate the
10 observed variability. We note that this model underestimate exists if we calculate relative variability using a maximum minus
11 minimum range from yearly average anomalies rather than monthly averages; the resulting variability ratio patterns (not shown)
12 are similar, overall, to the contour plots of Fig. 19. We provide a global view of lower stratospheric variability differences
13 (models versus data) in the anomaly time series comparisons at 83 hPa for all latitude bins in Fig. S8. This also shows that the
14 observed interannual changes in H₂O are clearly better followed by the SD-WACCM time series than by FR-WACCM,
15 including the drop in H₂O after 2011 (see Urban et al., 2014). While lower stratospheric H₂O variability is underestimated by
16 SD-WACCM by ~20%, the actual correlation between SD-WACCM and the observed anomalies is very good (as was shown in
17 Fig. 12).

18 For HCl, we show the (detrended) variability ratios in Fig. 21. The observed HCl variability is fairly well matched (within
19 ~20%) by both models in the MLS time period, with an edge given to SD-WACCM. The observed variability is often
20 underestimated (by ~30%) by both FR-WACCM and SD-WACCM in the earlier period (1992-2003). We believe that the
21 HALOE sampling plays a role in this, i.e. even if we limit the model comparison (as we do here) to just the same months as
22 when HALOE observations occurred, incomplete sampling in latitude and time can lead to differences versus a fully sampled
23 model (see Toohey et al., 2013), and more so in the polar regions where the HCl variability is large. In the upper stratosphere, the
24 variability ratios are comparable to or somewhat smaller than those in the middle stratosphere, and there is a 20-30%
25 underestimate of the observed variability, which is based on HALOE HCl observations for the 1992-2003 period. For 2005-
26 2014, SD-WACCM actually matches the upper stratospheric MLS variability fairly well, although these variability values are
27 small. There have been difficulties in fully understanding (or modeling) observed upper stratospheric HCl variations before the
28 declining phase that started after about 2000 (Vaughn et al., 2001); see also Sect. 5.

29 To complete this discussion of variability, we show the ratios of model to data variability for stratospheric HNO₃ (2005-2014)
30 and N₂O (2005-2012) in Fig. 22. We already discussed the issues with missing model chemistry for upper stratospheric HNO₃
31 variability, as well as the low signal-to-noise issue for HNO₃ data at low latitudes in this region (see bottom two panels, showing
32 low variability ratios there). There is reasonably good agreement in the HNO₃ variability between SD-WACCM and MLS for the
33 lower to mid-stratosphere, while FR-WACCM generally overestimates the HNO₃ variability in this region. Fig. 22 shows N₂O
34 results extending down to 100 hPa. Here, the MLS N₂O-640 data (from the 640 GHz radiometer) for 2005-2012 are used; these
35 retrievals were curtailed in the first half of 2013 as a result of degradation in the 640 GHz radiometer signal chain. Based on
36 results shown in SPARC (2017), there appears to be good agreement in the tropical interannual variability comparisons for N₂O
37 at 100 hPa between MLS and other satellite-derived results. The lower stratospheric N₂O time series behave more smoothly at
38 low latitudes in the models than in the observations. The interannual variability in the MLS N₂O measurements there is
39 somewhat smaller than the standard deviations in monthly mean N₂O values (of order 20-30 ppbv). The MLS N₂O measurement
40 noise itself for a monthly zonal mean (made up of about 5000-6000 profiles) should be less than 1 ppbv. Smoothing the model in



1 the vertical domain to better match the MLS vertical resolution would not lead to a better fit to the observed variability.
2 However, we should keep in mind that the MLS-derived N₂O variability is a small percentage (< 3%) of the monthly zonal mean
3 N₂O abundances. In summary, SD-WACCM shows some underestimate of the observed lower stratospheric tropical variability
4 for all the species considered here, except for HNO₃; FR-WACCM does so also for 3 species (H₂O, HCl, and N₂O). It may be
5 that some of the larger variability in the measurements arises from effects not tied just to MLS radiance noise issues, or from
6 variability caused by the proximity to the tropopause for measurements with finite vertical resolution; WACCM could also be
7 genuinely underestimating the actual atmospheric variability near the tropopause (for unknown reasons).

8 5.2 Trends

9 In this section, we discuss how the WACCM runs compare to stratospheric observations when it comes
10 to trends, from fairly short-term trends (from the Aura MLS time period) to longer-term trends based on comparisons with O₃,
11 H₂O, and HCl GOZCARDS data records (see Sect. 2 and Froidevaux et al., 2015). Trend analyses have their own complexities
12 in terms of analysis methods and uncertainty estimates. For example, O₃ trend assessments have had to deal with trend estimates
13 from different long-term data records, each with its own characteristics (Tummon et al., 2015; WMO, 2014; Harris et al., 2015;
14 Steinbrecht et al., 2017; Ball et al., 2017, 2018). This kind of analysis is especially difficult when investigating trends from time
15 series with high variability compared to the size of the change over time, which is an issue for the lower stratosphere in
16 particular. Global modeling efforts have also led to improved characterizations of the expected long-term impact of different
17 forcings, including the combined and separate impacts on ozone profiles of changes in halogen source gases and greenhouse
18 gases (WMO, 2014).

19 Here, we focus mostly on whether we obtain significant differences between trend results from WACCM and from
20 observations, given the application of the same analysis methods for the three sets of time series, for ozone and for other species.
21 We have applied multiple (or multivariate) linear regression (MLR) to the time series of deseasonalized anomalies from the data,
22 FR-WACCM, and SD-WACCM. In the Appendix (A3), we provide more details regarding the regression model, which includes
23 commonly used additive functional terms, namely a linear trend and a constant term, cosine and sine functions with annual and
24 semi-annual periodicities, as well as functions describing well known variations arising from the QBO and the El Niño southern
25 oscillation (ENSO); the same functions are applied to fit the model anomaly time series as well. Examples of observational time
26 series from merged ozone observations for 1998 through 2014 are provided in Fig. A3, along with the fits to the series and the
27 linear components (trends). Given the use of fairly short-term time series here (e.g., Aura MLS data alone), we have not included
28 a solar cycle component in the fits, as it can be highly correlated with a linear trend, and more than one 11-yr cycle would be
29 useful to better enable a separation of this signal. We also discuss our methodology for trend error evaluations in the Appendix
30 (A3). We use a block bootstrap method, like the approach of Bourassa et al. (2014) for their trend analyses of ozone from the
31 OSIRIS retrievals. We use random resampling of the residuals in yearly blocks (with 20,000 samples) in this Monte Carlo
32 approach to estimating errors, and we display the trend error bars as 2σ values (which is very close to the 95% bounds on the
33 distribution of linear trend results). Such calculations often lead to significantly larger error bars than more standard methods,
34 which neglect the autocorrelation of residuals. Again, we use the same regression model fits to extract trends from both
35 WACCM and observed time series; error bar calculations are also applied the same way to all the time series.

36 For ozone, we give an overview in Fig. 23 of deseasonalized anomaly time series (expressed as a percent of long-term means)
37 for 3 latitude bins (northern mid-latitudes, tropics, and southern mid-latitudes) and 2 pressure levels (3.2 hPa for upper
38 stratosphere, 68 hPa for lower stratosphere). The series were first deseasonalized in 10° latitude bins and then averaged. The
39 GOZCARDS data record used here (version 2.20) is an update to the original (version 1.01) record (Froidevaux et al., 2015), as



1 mentioned in Sect. 2. Fig. 23 shows generally good agreement between the various time series, although if one looks carefully,
2 SD-WACCM is generally closer to the observational time series than FR-WACCM is, as one might expect from previous
3 considerations of goodness of fit and variability; also, percent variability is larger in the lower stratosphere than in the upper
4 stratosphere (note the y-axis range difference for these regions), thus rendering trend detection more difficult at lower altitudes.
5 We compare in Fig. 24 the ozone profile trend results from MLS data alone, from the period 2005 through 2014, to those from
6 FR-WACCM and SD-WACCM, for the 3 aforementioned latitude bins. We show the error bars as 2σ estimates, as we find this
7 to be an easy way to visualize if there are really significant differences between models and data, or indeed, between the models.
8 Figure 24 shows that there is a robust indication from the MLS data that the upper stratospheric ozone values have been on the
9 upswing in the past decade, at a rate of about 0.2 to 0.4%/yr, depending on latitude region, with 2σ uncertainties of $\sim 0.2\%$ /yr.
10 While the 2σ error bars (obtained using the bootstrap method mentioned earlier) are fairly large, there are several independent
11 latitude regions and pressure levels with similar results, and thus, this positive trend is a robust near-global upper stratospheric
12 result. These results are broadly consistent with O₃ trends obtained by Steinbrecht et al. (2017), who use MLS as part of the
13 longer-term merged data records, although they studied a longer time period (2000-2016). All things being equal, the errors in
14 these trends should diminish as more years of data are added to the MLS O₃ record, which, for the middle and upper stratosphere,
15 has been characterized as “very stable”, namely within 0.1 to 0.2%/yr versus sonde and lidar network ozone data (Hubert et al.,
16 2016); it currently seems difficult to quantify “absolute stability” to much better than this, especially in the lower stratosphere. In
17 the lower stratosphere, trend results tend to be closer to zero, with larger variability and error bars (in %/yr), and unambiguous
18 detection of post-1997 ozone trends in this region remains elusive (WMO, 2014; Harris et al., 2015). The 2005-2014 trends in
19 Fig. 24 show good broad agreement between model and data, with a tendency for SD-WACCM to agree better than FR-
20 WACCM with MLS, albeit not significantly so, given the size of the error bars, especially in the lower stratosphere; the lower
21 stratospheric tropical results from FR-WACCM are negative, in contrast to both the observational result as well as SD-WACCM,
22 but with large overlapping error bars (given the variability and fairly short time period). In the tropical upper stratosphere, both
23 models exhibit a somewhat more positive trend than observed for this period, although, again, these trend differences are not
24 statistically significant. Nevertheless, as a function of latitude (in 10°-wide bins) and pressure, the 2005-2014 SD-WACCM O₃
25 trends do follow the MLS trend results quite well; this is shown in Fig. 25 for central latitudes from 55°S to 55°N.

26 For a consideration of longer time periods, we use the merged O₃ data record from GOZCARDS (version 2.20) to display
27 NH mid-latitude and tropical trends from the models and data for 3 different periods in Fig. 26. The top panels for 1985-1997
28 focus on the main “declining phase”, while the middle panels (1998-2014) show results for the (expected) “early recovery” stage;
29 the 2005-2014 (bottom panels) results are essentially the same trends as those from Aura MLS (Fig. 24) for that period. The
30 largest differences between the two GOZCARDS data versions occur in the tropical upper stratosphere for the declining ozone
31 phase; Fig. S9 displays the tropical trend differences that we obtain for the same three periods as in Fig. 24. In agreement with
32 this are the trend differences provided by Ball et al. (2017), who showed results for the original (version 1.01) GOZCARDS data
33 and for SWOOSH. GOZCARDS version 2.20 data are now in better agreement with the merged SWOOSH O₃ product (as both
34 use SAGE II version 7 data); also, Steinbrecht et al. (2017) showed that the two merged records lead to similar (post-2000) trend
35 results. As mentioned in Sect. 2, the improvements in GOZCARDS version 2.20 ozone are a result of the incorporation of the
36 SAGE II v7 retrievals (Damadeo et al., 2013), and the use of the MERRA temperatures (used in v7) for the conversion from
37 density/altitude to the GOZCARDS mixing ratio/pressure grid. We note, however, that the lower stratospheric region exhibits
38 interannual variability that is several times larger than that in the upper stratosphere, as seen in Fig. A3 for tropical 1998-2014
39 data versus SD-WACCM anomaly time series. Even fairly subtle differences in time series over a few years can lead to a sign
40 change in the trends, although there is no statistical significance in the resulting trend differences (see Fig. A3 and Fig. 26,



1 middle right panel at 68 hPa, for the data and SD-WACCM trend results).

2 To put in perspective what a statistically significant trend difference looks like, we show in Fig. S10 the O₃ anomaly time
3 series for 1998-2014 at 1 hPa for 30°N-60°N, where the SD-WACCM and GOZCARDS trend results lie outside their respective
4 2σ uncertainties (based on Fig. 26, middle left panel); in this case, the FR-WACCM results happen to be in better agreement
5 with the data. One aspect that could impact model/data differences is that the models are not sampled, here, following the
6 sparser (occultation) viewing, neither in latitude nor in time (time within each month and local time also, since model values
7 represent 24-hr averages). Much denser spatial and temporal sampling is obtained during the MLS observation period, with very
8 regular sampling; while small systematic differences may affect comparisons between average model and MLS abundances, such
9 differences should be very consistent from year to year, thus minimizing the impact on derived trend differences. Also, some of
10 the differences in the upper stratosphere might arise because the averaging of sunset and sunrise occultation data is not as robust
11 for 1998-2004 as for pre-1998, when SAGE II was operating more continuously in both sunset and sunrise modes (and O₃ varies
12 more strongly with local time at 1 hPa than at lower altitudes). Also, HALOE had decreasing spatio-temporal coverage in the
13 later years, but these upper stratospheric trend differences would require further investigation. For the 30°S-60°S region (not
14 shown here), we observe similar results (within error bars) as in Fig. 26 for 30°N-60°N; this includes small negative trends in the
15 lower stratosphere for all 3 time periods, although with no statistical significance (i.e., consistent with a zero or slightly positive
16 trend). We also see the same sort of vertical shift in the SD-WACCM profile trends for 1985-1997, as compared to the observed
17 trend profile, which reaches its most negative value at a slightly higher altitude (lower pressure), but the reason for this shift is
18 unknown; theoretically, this could be tied to model/data differences in the ClO peak, and its evolution. Our ozone trend results
19 are largely consistent with other previous work (references mentioned above), which (for records including MLS) typically used
20 merged O₃ from GOZCARDS or from SWOOSH (Davis et al., 2016). We find statistically significant trends (meaning that,
21 assuming Gaussian statistics, a zero trend is not included in the 2σ error bar range) mostly in the upper stratosphere, both pre-
22 1997 and after 1998. While we observe some small O₃ decreases in the lower stratosphere post-1998, as obtained recently also
23 by the novel analyses of Ball et al. (2018), our study finds little statistical significance there, and a fair level of sensitivity to the
24 starting year or to the data sets used, with a swing to more positive (but marginally significant) results, if the starting year is 2005
25 and one just uses MLS data. Past work (e.g., Harris et al., 2015) has also shown sensitivity to the series starting and end points;
26 different regression analysis methods can also lead to some non-negligible differences, as shown for example by Nair et al. (2013)
27 and Kuttippurath et al. (2015). We also note that past analyses of lower stratospheric tropical O₃ data have shown positive
28 tendencies, based not just on satellite data as indicated here (with marginal significance) from MLS data alone, but also based on
29 earlier SCIAMACHY data (Gebhardt et al., 2014); in this reference, a positive trend was also obtained from averaged tropical
30 ozonesonde data. This will continue to require further study, towards a longer-term result.

31 For H₂O, there have been somewhat conflicting past results on stratospheric trends, depending on whether one investigates
32 sonde or satellite data (e.g., Oltmans et al., 2000; Scherer et al., 2008), and regarding mechanisms that could account for more
33 than a few tenths of a %/yr increase in H₂O, as CH₄ increases do not appear to be large enough for this. Beyond the potential
34 impact of CH₄ trends on upper stratospheric H₂O, changes in cold point temperatures or in the circulation need to be invoked in
35 order to account for significant decadal-scale trends in H₂O (e.g., Randel et al., 2000; Rohs et al., 2006; Tian and Chipperfield,
36 2006; Hegglin et al., 2014). Based on our analyses for 2005-2014 (MLS data versus models), we find in Fig. 27 that this recent
37 decade shows a positive H₂O trend both in the data and in the SD-WACCM result, which tracks the observations (versus latitude
38 as well as pressure) better than FR-WACCM does, as already seen in terms of quality of fits and variability. FR-WACCM
39 exhibits systematically smaller H₂O trend values than both MLS and SD-WACCM at all pressures except near 100 hPa, although
40 the FR-WACCM and SD-WACCM trend error bars (2σ) overlap. This overlap for FR-WACCM and MLS is more marginal in



1 the lower mesosphere, where the impact on H₂O from CH₄ decomposition should be at its maximum, and cold point temperature
2 variability issues are smaller than near the tropical tropopause. Also, the FR-WACCM trends have smaller error bars, given the
3 lower variability typically found in this model. With such lower variability, detection of a given trend would take less time than
4 with the observed variability; however, if the trend is larger (as shown by SD-WACCM or MLS), it also becomes easier to
5 detect. An analysis of H₂O HALOE profiles and ground-based microwave profiles over Hawaii (Nedoluha et al., 2009) showed
6 that changes in upper stratospheric and mesospheric H₂O are sensitive to the solar cycle (see also the mesospheric GOZCARDS
7 H₂O time series in Froidevaux et al., 2015), but show only negligible overall trends between 1992 and 2008.

8 We should also note some drifts have been detected between coincident MLS and sonde H₂O data, mostly since about 2010,
9 implying that MLS-derived trends are more positive than those from frost point profiles at several sites (Hurst et al., 2016). This
10 relative drift (of as much as 0.5-1.6%/yr for 2010-2015, depending on altitude and location) could therefore play a role in the
11 (small, in comparison) discrepancy between model (SD-WACCM) and MLS trends. The SD-WACCM results agree quite well
12 with MLS in Fig. 27, but they would become larger than MLS (adjusted) results if one were to subtract more than 0.1-0.2%/yr
13 from these MLS trends. Possible causes of observed drifts between MLS and sonde H₂O data are still being investigated, with
14 only a small part of this discrepancy currently attributable to a known instrumental degradation issue for MLS H₂O, which
15 probably also impact other MLS data from the 190 GHz spectral region (N₂O in particular). We note from Fig. 27 that H₂O trend
16 values from both models and MLS data agree better at 100 hPa. Also, we see that H₂O in the tropical region at 80 to 100 hPa
17 (near the stratospheric water vapor entry level) does not display much of a trend. Other recent studies of entry-level H₂O using
18 large-scale satellite data and longer-term analyses (starting in the mid-1980s) have concluded that no significant long-term trend
19 is discernible (Hegglin et al., 2014; Dessler et al., 2014). The former study led to slightly negative lower stratospheric H₂O trends
20 (although with no statistical significance).

21 For the 1992-2014 period and using GOZCARDS H₂O data (not shown here), we also find negative, although not statistically
22 significant, trends in the lower stratosphere, with small (< 0.2%/yr) positive trends in the upper stratosphere and lower
23 mesosphere; there is close agreement between the FR-WACCM and SD-WACCM trends (typically within 0.1%/yr) for that
24 period. This GOZCARDS record, however, does not include SAGE II data back to the late 1980s, as was the case for the
25 analysis by Hegglin et al. (2014), who also obtained positive longer-term (1980-2010) trends in the upper stratosphere from
26 satellite-derived H₂O anomaly records, merged using a global CCM as a transfer function. As found by others, especially when
27 dealing with relatively large decadal-type variability (including the QBO), the choice of start and end points, as well as the length
28 of period studied, can significantly influence the trend values, whether it be for H₂O or for O₃. Long-term lower stratospheric
29 H₂O trend detection is rendered difficult by such variability, including significant short-term water vapor changes (Randel et al.,
30 2006; Hurst et al., 2011; Fueglistaler, 2012; Urban et al., 2014), as noted here also (for the relatively short time series of Fig. S8).

31 For HCl, the changes in stratospheric abundances have been non-linear, with a rapid rise prior to 1998, and a slower rate of
32 decrease after 2004, as expected from the time-shifted abundances of total chlorine at the surface (Froidevaux et al., 2015).
33 Focusing first on 2005-2014, we show the corresponding model and data HCl trend results for the lower stratosphere in Fig. 28.
34 The agreement between the SD-WACCM and MLS trends is quite good, especially for the 30°S-60°S bin, although the error
35 bars are fairly large. However, the (negative) HCl trend results from both models practically always lie below the observed
36 trends for the three latitude bins shown in Fig. 28. The upper portion of this model/data bias follows what we observe as well in
37 the upper stratosphere (not shown), where MLS-derived HCl trends are clearly too flat (shallow) compared to expectations (from
38 model and surface-derived chlorine trends), whereas the upper stratospheric (negative) trends from the original MLS HCl
39 product were more negative (see Froidevaux et al., 2006; Livesey et al., 2018). As a reminder, the MLS team recommendation is
40 for data users not to include upper stratospheric MLS HCl data (post-2006) in any trend studies. For the lower stratosphere,



1 where the HCl line is broader, there is less concern about the inability to track the HCl trend; also, the near-zero drifts (i.e., drifts
2 < 0.1%/yr) obtained between two separate MLS ozone band retrievals (not shown here), one from the same radiometer as HCl,
3 and one from the main (very stable) standard MLS ozone product (see Hubert et al., 2016), provide some confidence regarding
4 the stability of lower stratospheric HCl trends. At low latitudes, MLS HCl shows a positive trend (largest and statistically
5 significant at 68 hPa, per Fig. 28). The vertical gradient structure in the observed HCl trends is duplicated to some extent by the
6 SD-WACCM results, although the model trends are always less than those derived from MLS. Latitude/pressure trend
7 variability, including positive tendencies, could be related to circulation changes, as implied by analyses of short-term increases
8 in lower stratospheric HCl seen in both ground-based and GOZCARDS data (Mahieu et al., 2014). Given the rapid rise in
9 chlorine prior to 1998 and the non-linear changes near the peak period, we show in Fig. 29 some of the lower stratospheric time
10 series (for 3 latitude bins and 3 pressures) from GOZCARDS merged HCl (Froidevaux et al., 2015) and the WACCM runs for
11 1992-2014. There is fairly good agreement in the non-linear behavior observed in both data and model lower stratospheric series.
12 The scatter in HCl data decreases after 2005, and the earlier time series suffer from more inhomogeneous sampling, which may
13 at least in part explain the larger scatter and model/data differences (there is no attempt here to sample the models within each
14 month like the data, and this would be difficult for a merged data set). There are also regions and periods of slow increases in
15 HCl in both data and models (Fig. 29), as well as hemispheric differences in the short-term tendencies, as discussed before by
16 Mahieu et al. (2014) and Froidevaux et al. (2015). The HCl time series are tracked fairly well by SD-WACCM, which generally
17 matches the data better than FR-WACCM; this is consistent with the understanding that dynamically-driven variations are better
18 captured by the incorporation in SD-WACCM of realistic meteorological fields (MERRA). Stolarski et al. (2018) have recently
19 investigated the removal of dynamical variability from MLS lower stratospheric HCl time series by using MLS N₂O
20 measurements as a fitting parameter in the regression analysis; this led to retrieved HCl trends that generally match expectations
21 based on the rates of decrease in surface total chlorine. The search for detailed explanations of such short-term increases and
22 variability in lower stratospheric HCl (and other changes in composition) continues to be an interesting area of investigation.

23 In the upper stratosphere, it has been difficult to explain the details of the observed HCl variations between 1998 and 2002,
24 including the dip between these years (Waugh et al., 2001). We show in Fig. S11 the near-global (60°S-60°N) HCl time series
25 from GOZCARDS at 1 hPa. This helps to underscore that there is a systematic model underestimate of HCl in the uppermost
26 stratosphere; the model/data difference is much smaller if one moves to a pressure close to 5 hPa. While the data uncertainties (of
27 ~0.2 ppbv, based on Froidevaux et al., 2015) encompass the model values, the model total should actually be increased by the
28 chlorine contribution from very short-lived substances (VSLs) to the stratosphere; although this contribution is believed to be
29 less than 0.1 ppbv (Carpenter et al., 2014), recent evidence suggests that there could be a somewhat larger stratospheric chlorine
30 contribution from VSLs (Oram et al., 2017). Nevertheless, the historical maximum for total tropospheric chlorine abundance
31 was about 3.65 ppbv (WMO, 2014), and this should be the maximum total chlorine expected in the uppermost stratosphere.
32 Upper stratospheric HCl should see a somewhat broader peak than at the surface, with a smaller and time-delayed maximum,
33 depending on transport-related effects (and age of air spectrum, e.g., see WMO, 2010). While WACCM includes the proper
34 abundance and evolution of chlorine source gases at the surface, maximum WACCM HCl in the upper stratosphere (and lower
35 mesosphere) is just under 3.4 ppbv. It is also interesting that the gap between the models (both versions) and the data worsens
36 from 1992 to 2000, with the HCl peak occurring later in the data (with a broader peak than in the models). After about 2002, the
37 decrease in near-global HCl roughly follows the model decrease; additional years of HCl data from ACE-FTS should help refine
38 this comparison.

39 For N₂O and HNO₃, lower stratospheric model trends are compared to the corresponding MLS data trends in Fig. 30. We note
40 that the MLS standard product right after launch was N₂O-640 (retrieved from the 640 GHz radiometer band data), but it was



1 discontinued after mid-2013, as mentioned earlier, as a result of a rapid hardware degradation issue affecting that band (N_2O
2 only). The current MLS standard product, N_2O -190, is retrieved from the 190 GHz band. Figure S12 provides evidence of
3 negative drifts in lower stratospheric N_2O -190, apparently accelerating in the last few years, since the SD-WACCM and actual
4 N_2O values would be expected to continue to rise slowly after the end date on this plot, notably in the tropical lower stratosphere,
5 where N_2O should follow tropospheric trends. Indeed, tropospheric N_2O has been increasing at a fairly steady rate of $\sim 0.26\%/yr$
6 (WMO, 2014), consistent with the underlying model N_2O and MLS N_2O -640 lower stratospheric increases at low latitudes (see
7 Fig. S12 and especially the tropical trends obtained in Fig. 30 at 100 hPa). The FR-WACCM N_2O trends show slightly poorer
8 agreement than SD-WACCM versus N_2O -640, although this is not statistically significant. The tropical lower stratospheric MLS
9 N_2O -190 trend results (not shown here) are negative (albeit with error bars that encompass small positive trends), but show some
10 differences versus expectations and the N_2O -640 results. As for HCl, interhemispheric differences in lower stratospheric N_2O
11 trends are interesting in terms of their implications for effects relating to transport (age of air) and changes in the circulation. At
12 lower pressure values, the N_2O trends do not mirror the tropospheric N_2O trends (in $\%/yr$), and other factors play a role (age of
13 air and changes in circulation, as well as N_2O photodissociation). The asymmetric trend pattern between hemispheres, even if it
14 is not a long-term trend, may well point primarily to short-term effects tied to asymmetries in the age of air, and therefore, in the
15 circulation. The asymmetry in age of air results obtained by analyses of (2002-2012) MIPAS SF_6 data (Haenel et al., 2015) could
16 also be related to asymmetries in the N_2O tendencies. They found relatively older air in the northern extra-tropics and younger air
17 in the southern extra-tropics. This could also imply opposite trends for N_2O in the southern and northern lower stratosphere.
18 However, Bönisch et al. (2011) have pointed out that different tracers, like O_3 and N_2O , can be impacted in different ways by
19 both vertical and quasi-horizontal transport effects, depending on their relative gradients in both vertical and meridional domains.
20 Moreover, their work indicates that detailed attribution of tracer variations to structural changes in the Brewer-Dobson
21 circulation is a complex matter, and short-term and longer-term changes may well have different characteristics. Our work here
22 mainly identifies some similarities between some of the trend patterns versus pressure and latitude from SD-WACCM, in
23 particular, and the observed trends, for certain time periods. For the HNO_3 trends (Fig. 30), we also see good agreement between
24 models and data for 2005-2014; latitudinal tendencies and interhemispheric differences therein are similar for model and data.
25 The spatial gradients of these species are different in the lower stratosphere (HCl and HNO_3 increase with height, in contrast to
26 N_2O), and we see that the decreasing HCl trends for 2005-2014 at 30°S - 60°S (Fig. 28), in particular, are qualitatively similar to
27 those from HNO_3 in this region. For lower stratospheric HNO_3 , there is an underlying trend part caused by the slow increases in
28 N_2O , as we can observe in longer-term (1980 to present) model time series (not shown here). N_2O and H_2O (source gases for
29 HNO_3) are significantly affected by the QBO and there is a strong related variability in lower stratospheric HNO_3 . Furthermore,
30 substantial increases in stratospheric aerosols after large volcanic eruptions have influenced lower stratospheric HNO_3 via
31 heterogeneous hydrolysis of N_2O_5 (Arnold et al., 1990; Rinsland et al., 1994), and this will impact HNO_3 trends that include
32 volcanically-perturbed periods. We saw that seasonal enhancements in NO_x coming down from the mesosphere can also affect
33 HNO_3 at high latitudes. Some observed short-term trend patterns in HCl, HNO_3 , and N_2O are better captured by the SD-
34 WACCM model overall than by FR-WACCM, as we show in Fig. 31 for the 2005-2010 period, relevant to the results from
35 Mahieu et al. (2014), who emphasized short-term HCl increases during this time. We note the correlation in these short-term
36 trend results for HNO_3 and HCl, but an anti-correlation for N_2O versus HCl (and HNO_3).

37 Finally, to re-emphasize how difficult it can be to detect small underlying trends in the lower stratosphere, in particular, Fig.
38 32 shows deseasonalized model anomalies from SD-WACCM for the 25-year period between 1990 and the end of 2014, for
39 HNO_3 , as well as O_3 , H_2O , and HCl; on this scale, N_2O variability in this region is much smaller (not shown). As noted
40 previously, and from past work on this topic, fairly sharp drops in water vapor in this region occurred shortly after 2000 and



1 2011, but we also note the significant decadal-type variability in this region, besides the expected links to QBO- and ENSO-type
2 variations. There are also non-negligible radiative implications surrounding such variations for H₂O (Solomon et al., 2010;
3 Gilford et al., 2016; Wang et al., 2016); however, a slow underlying long-term evolution would take time to detect, given the
4 variability. Moreover, the percent variability is even larger for other species, which correlate well with H₂O in this region;
5 indeed, the correlation coefficients between these various time series are all between 0.64 and 0.86, although this multi-species
6 agreement is poorer at many other pressure levels or latitudes. In Fig. S13, we show that the observed variability in this same
7 region for O₃ and H₂O (for 1992-2014) is fairly well matched by the models, although this is somewhat less true for H₂O than for
8 O₃. Besides the importance of circulation effects on tracers in this region, tropopause temperatures will also affect water vapor;
9 this adds some complexity in terms of exactly modeling its variations in the tropical lower stratosphere.

10 6 Summary and discussion

11 The climatological averages from FR-WACCM and SD-WACCM for O₃, H₂O, HCl, N₂O, and HNO₃ generally compare
12 favorably (within the 2 σ estimated systematic errors) with the Aura MLS data averages for 2005-2014. Model ozone values are
13 usually within ± 5 -10% of the average observations, except in the UTLS. In the lowest portion of the stratosphere, SD-WACCM
14 generally exceeds the observed ozone means by about 30-50%, with FR-WACCM showing a smaller overestimate; both models
15 also tend to overestimate the amplitude of the annual cycle in this region. There is also a model low bias (by ~ 10 -30%) from 215
16 to 261 hPa at low latitudes, which could largely be caused by a known MLS high ozone bias in this region. For H₂O, there is a
17 model low bias (by 5-15%) versus MLS data in the upper stratosphere and most of the mesosphere, although some of this arises
18 from a small high bias in MLS H₂O versus other satellite data sets (see Hegglin et al., 2013). Also, the models significantly
19 underestimate the average HNO₃ abundances in the upper stratosphere, notably at high latitudes; this largely appears to stem
20 from missing model ion chemistry, as it relates to particle precipitation effects in the mesosphere, followed by downward
21 wintertime polar transport of enhanced NO_x, and subsequent seasonal increases in HNO₃. There is also some model
22 overestimation by SD-WACCM of MLS HNO₃ (by about 40%) at high latitudes for pressures larger than 100 hPa, although
23 there is a need for further validation of the HNO₃ data in this region. In the lower stratosphere at high southern latitudes, the
24 variations in polar winter/spring composition observed by MLS are generally well matched by SD-WACCM, the main exception
25 being for the early winter rate of decrease in HCl, which is too slow in the model. Grooß et al. (2018) have provided further
26 discussion of this discrepancy, which should have little impact on winter/spring polar ozone depletion; indeed, we find good or
27 better agreement between the seasonal high latitude observations and SD-WACCM for ozone and other species.

28 Regarding the fitted variability tied to the AO and SAO, there are a few discrepancies between model-derived amplitude
29 patterns and the corresponding MLS climatology features, but FR-WACCM and SD-WACCM appear to properly capture the
30 primary processes governing these modes of variability. SD-WACCM generally matches the data sets slightly better than FR-
31 WACCM does. The O₃ AO stratospheric amplitudes are within ~ 25 % of the MLS AO amplitudes. For H₂O, both WACCM
32 versions exhibit AO and SAO patterns that are generally consistent with the observations, and with recently published satellite-
33 derived results (Lossow et al., 2017a); we also note the WACCM underestimation of H₂O AO and SAO amplitudes in the lower
34 stratosphere, although this is the region with the smallest amplitudes (< 0.1 ppmv).

35 We have provided diagnostics for the fits between the WACCM runs and the MLS deseasonalized anomaly time series. These
36 consist of the correlation coefficient (R^2 diagnostic) as well as a diagnostic of RMS differences (model versus data), divided by
37 the RMS variability in the data; a combined diagnostic (the ratio of the above two diagnostics) is also used to help differentiate
38 between the two model runs, which are often not too far apart. Not too surprisingly, SD-WACCM, which is driven by realistic
39 dynamics versus time, generally matches the observed zonal monthly mean anomalies significantly better than does FR-
40 WACCM. This holds for all five species that we considered, with larger values of R^2 and smaller values of the RMS difference



1 diagnostic. In the tropical lower stratosphere, where there is some nudging to equatorial winds for FR-WACCM (and even more
2 so for SD-WACCM, see Sect. 2), the FR-WACCM fits to the data are generally improved. However, details of the observed
3 interplay between SAO, AO, and QBO variations in tropical upper stratospheric ozone are better matched by SD-WACCM
4 variations than by FR-WACCM. Also, FR-WACCM shows poorer agreement with observed seasonal polar winter/spring lower
5 stratospheric variations than does SD-WACCM. Finally, in the mesosphere, the water vapor anomalies are better matched by
6 SD-WACCM than by FR-WACCM.

7 Variability comparisons represent a more fair and useful metric in terms of the characterization of model quality, in particular
8 for a free-running model. To this end, we have compared the RMS interannual variability from the anomalies in both WACCM
9 models and observations; we have used both MLS data (2005-2014) and data from longer-term series based on GOZCARDS
10 data records for O₃, H₂O, and HCl. One of the main features from these comparisons is that the H₂O variability from the lower
11 stratosphere to the upper mesosphere is underestimated by both model runs used here; this underestimate can reach a factor of
12 two, although more typically, it is of order 30%. This implies that a larger number of years would be needed to detect an actual
13 trend in H₂O than if one uses a model-based prediction (from FR-WACCM); this number of years would be increased by a factor
14 of 1.2 to 1.6, if one uses the two variability factors mentioned above. Apart from the WACCM underestimate of observed H₂O
15 variability, the observed lower stratospheric variations, including significant drops in the H₂O abundance, are better tracked by
16 SD-WACCM than by FR-WACCM. The ozone variability is better represented by the WACCM models, with model/data
17 variability ratios typically within a factor of 0.8 to 1.2. Observed HCl variability is underestimated somewhat by FR-WACCM
18 for the 1992-2003 period, but not for the later (2005-2014) period; the sparser HALOE sampling, compared to MLS, could
19 explain some of the underestimate for the early period, especially in the polar regions. For N₂O, there is also a model
20 underestimate (from both FR-WACCM and SD-WACCM) of the lower stratospheric low latitude variability observed by MLS,
21 although this variability is a small percentage of the mean values.

22 Regarding trends, the model comparisons versus the longer-term ozone data record from GOZCARDS (version 2.20 being
23 used here) show generally good qualitative agreement in the time series in different latitude bins for both upper and lower
24 stratospheric change. It is clear from such time series that the larger percent variability in the lower stratosphere will continue to
25 render trend detection in this region more difficult than in the upper stratosphere. Based on the Aura MLS O₃ data record itself,
26 which has been deemed very stable (Hubert et al., 2016), we observe robust evidence, considering the 2σ error bars of ~0.2%/yr
27 (estimated using a block bootstrap method), that there is a positive upper stratospheric ozone trend for 2005-2014, at a rate of
28 ~0.2-0.4%/yr. This is true for all three broad mid-latitude and tropical regions considered here (30°N-60°N, 30°S-60°S, and
29 20°S-20°N), although the evidence is more marginal in the SH mid-latitudes. The WACCM trends estimated using the same
30 regression model as used for the MLS data (anomaly) series show generally good agreement with the data trends, although the
31 error bars are fairly large (for both data and models). Furthermore, the observed trend (relative) dependence on latitude and
32 pressure is well matched by the SD-WACCM trend results. We have not considered the high latitudes in detail in this work, in
33 part because of the significant dynamical variability in that region. In this regard, Stone et al. (2018) recently analyzed model
34 results at high latitudes in the upper stratosphere, and showed that the large variability in that region, including the effects of
35 solar proton events, is likely to mask detection of recovery (for now), although autumn and winter should exhibit the strongest
36 recovery signals. In the lower stratosphere, where larger variability exists, the trends we deduce from the data sets and models
37 agree within fairly large error bars, although there is generally no statistical significance. While there is a tendency for the
38 GOZCARDS merged O₃ record to show small decreasing trends for the 1998-2014, the trend results reverse to near-zero or
39 slightly positive tendencies (albeit with no robust statistical significance) if one considers the MLS data alone for 2005-2014.
40 SD-WACCM trend results seem to track these positive tendencies, although with no robust statistical significance, based on our



1 analyses. The recent work by Ball et al. (2018) indicates a net O₃ decrease in the lower stratosphere from about 1998 to the
2 recent few years; this does not contradict the possibility of a turn-around towards a more positive rate of change in this region
3 during the more recent 10-12 years. The positive tendency noted here may get more robust with the analyses of more years of
4 high quality global ozone profiles, and possibly more aligned with longer-term model expectations. Future detailed analyses of
5 these issues with different regression models and other methods are certainly indicated.

6 For H₂O, the most statistically significant trend result is an upper stratospheric increase for the post-2005 time period, peaking
7 at slightly more than 0.5%/yr in the lower mesosphere, with MLS and SD-WACCM results agreeing fairly well, and FR-
8 WACCM showing significantly smaller increases. The fact that the last decade has seen more of an upper stratospheric and
9 mesospheric H₂O increase than the previous decade appears to correlate with the very shallow maximum that occurred in the last
10 cycle (number 24) of the solar flux, which seems tied to the shallower dip, and broader overall maximum, in upper mesospheric
11 H₂O (see the H₂O and solar flux time series in Fig. 16 of Froidevaux et al., 2015). However, the non-linear influence of recent
12 changes in methane, which include a plateau from 1999-2006, with a return to rising abundances after that (Schaefer et al., 2016;
13 Nisbet et al., 2016), would also need to be considered for the upper stratosphere. There is also a caveat regarding MLS-derived
14 H₂O trends, given the existence of non-negligible drifts between sonde and MLS H₂O data (Hurst et al., 2016), at least since
15 about 2010. Such drifts can only be partially explained by currently known instrumental degradation issues affecting the MLS
16 retrievals of H₂O, with some impact on other data from the 190 GHz radiometer (N₂O, in particular). Thus, the MLS-derived
17 H₂O trend results obtained here are likely to be upper limits; this can probably explain why the model trends (at least from SD-
18 WACCM) currently lie on the low side of the observed trends. An upcoming update to the MLS retrievals will lead to a
19 reduction (but not an elimination) of the aforementioned drifts between MLS and sonde H₂O data. There is a continued need for
20 cross-comparison of the various (diminishing number of) satellite H₂O data sets, as well as H₂O profiles from satellites and
21 sondes, hopefully leading to a better understanding and mitigation of instrumental issues and drifts between different water vapor
22 observations.

23 Our HCl trend analyses reveal broad agreement between the lower stratospheric MLS data (2005-2014) and the models, but
24 with some systematic differences. As mentioned in the past, upper stratospheric MLS HCl data are not deemed to be reliable
25 enough for trend studies, since the cessation (in 2006) of the primary target band retrievals for MLS HCl. While lower
26 stratospheric HCl decreases are generally indicated for 2005-2014, there are some hemispheric differences, and a significant
27 increase is suggested in the tropical data at 68 hPa; however, there is only a slight positive trend there from the SD-WACCM
28 result (with no statistical significance). While the lower stratospheric vertical gradients of trend results from MLS HCl are
29 duplicated to some extent by SD-WACCM, the model trends are always on the low side. There is no clear preference for SD-
30 WACCM or FR-WACCM trend results, in comparison to the observed lower stratospheric HCl trends. There have also been past
31 indications of short-term increases in lower stratospheric HCl (Mahieu et al., 2014); the study of such short-term tendencies for
32 implications regarding circulation changes are worth pursuing further (but outside the scope of this work). There is a need for
33 more comparisons of the various HCl measurements, satellite-based and ground-based, as well as models, in order to better
34 understand circulation influences on stratospheric composition, as well as potential measurement-related issues (e.g., from
35 sampling differences or potential drifts).

36 For N₂O, the asymmetry in MLS-derived trends (for 2005-2012) between hemispheres, with negative trends (of up to about -
37 1%/yr) in the NH mid-latitudes and positive trends (of up to 3%/yr) in the SH mid-latitudes, is in agreement with the asymmetry
38 that exists in SD-WACCM results. The small observed positive trends of ~0.2%/yr in the 100 to 30 hPa tropical region are also
39 consistent with model results (SD-WACCM in particular), which in turn are very close to the known rate of increase in
40 tropospheric N₂O (at a rate of about +0.26%/yr, see WMO, 2014). In the case of HNO₃, the MLS-derived lower stratospheric



1 trend differences (for 2005-2014) between hemispheres are opposite in sign to those from N₂O (whose spatial gradients are
2 largely of a sign opposite to those from HNO₃) and in reasonable agreement with both WACCM results, despite large error bars
3 compared to the size of the trends. More detailed analyses would be needed to try to relate such trend asymmetries to changes in
4 age of air, or circulation.

5 Overall, the models and observations show good agreement in the trends, with somewhat better results for SD-WACCM,
6 which displays good correlations in the trend behavior versus latitude and pressure. However, the error bars are non-negligible,
7 and the choice of start and end dates can have a significant impact on trends or tendencies. Given the existence of significant
8 short-term and decadal-type variations for several lower stratospheric species, one should be cautious not to assign, or
9 extrapolate, a tendency based on even a decade of data, to an underlying longer-term trend.

10 The diagnostics provided in this WACCM model evaluation can help distinguish even fairly subtle differences between
11 models and observations, as well as between models. The generally improved fits to observations from a specified dynamics
12 model versus a free-running model are to be expected, but also need to be documented (part of the aim of this work). We are also
13 reminded that observations have their own systematic issues, and close collaboration between modeling groups and instrument
14 teams can help untangle issues that might be more driven by, or at least influenced by, species-dependent instrumental effects.
15 Especially when comparing longer-term model time series to observations, even small systematic effects such as measurement
16 drifts, or data merging issues, can become important for trend diagnostics. Finally, independent CCMs are not created in the
17 same exact way, and nudging approaches for free-running and specified dynamics models can vary; for example, some models
18 have an internally-generated QBO, but most do not. While this study focused on (CESM1) WACCM runs, further studies of the
19 differences between high quality observations and various international models of atmospheric composition would be useful, to
20 put this work in perspective. This could also be expanded to include some species not considered in this work, and/or with more
21 of a focus on the upper troposphere.
22



1 Appendix A

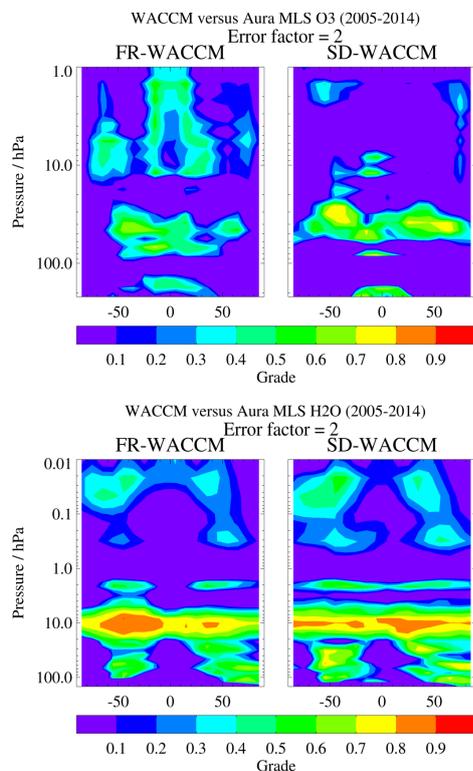
2 A1 Examples of model grades

3 A grading method that has been applied in some previous comparisons (e.g., see Douglass et al., 1999; Waugh and Eyring,
4 2008) between atmospheric model values (M_n) and observed values (O_n) utilizes Eq. (A1) below to arrive at grades between 0
5 and 1 (and if a grade is < 0 , it can be set to 0):

$$6 \text{ grade} = 1 - \sum_1^N \frac{|M_n - O_n|}{E_f \times \sigma_n} \quad (\text{A1})$$

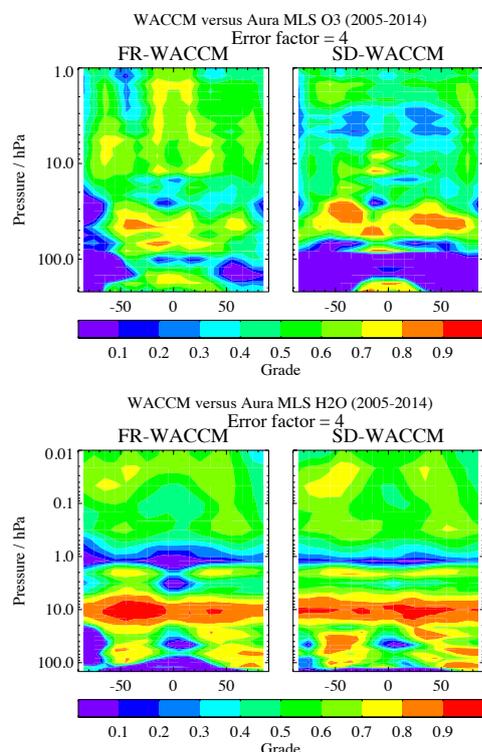
7 with index n (in a given time series) varying between 1 and N (the total number of monthly values being compared for a given
8 latitude/pressure bin), and σ_n representing the error in the observations. While the error factor E_f should probably be set to 2 or
9 3, this gives grades that are too small (close to zero or negative) if one applies such a formula to the MLS O_3 or H_2O time series,
10 specifically to data sets with pretty well defined total measurement errors (provided as 2σ error estimates, per Sect. 4, meaning
11 an error factor of 2). The grades shown here in Figs. A1 and A2 correspond to error factors (E_f) of 2 and 4, respectively. Figure
12 A2 leads to O_3 and H_2O grades that are more useful than Fig. A1; it also shows similarities with the diagnostic based on the RMS
13 of the differences between model and data, as shown in Sect. 4 (see the description of this diagnostic in Appendix A2).

14



15
16
17
18
19
20

Figure A1. Examples of grades for model evaluations of O_3 (top two panels) and H_2O (bottom two panels), using a grading methodology that has been used in the past (see Eq. A1), applied to both FR-WACCM (left panels) and SD-WACCM (right panels) time series versus Aura MLS time series from 2005 through 2014. These grades are for an error factor of 2 (in Eq. A1).



1
 2
 3
 4
 5

Figure A2. As in Fig. A1, for model evaluation grades of O₃ and H₂O (WACCМ versus Aura MLS data), but for a value of 4 (rather than 2) for the error factor (see Eq. A1).

6 A2 RMS difference diagnostic

7
 8 Given a model time series $M_i(t)$ and an observational time series (both series here representing deseasonalized anomalies)
 9 $O_i(t)$, the difference values between the two anomaly series are simply given by

$$10 \Delta_i(t) = M_i(t) - O_i(t) \quad (\text{A2})$$

11 and the root mean square (RMS) of these anomaly differences (RMSdif) is expressed as

$$12 \text{RMSdif} = \sqrt{\frac{1}{N} \sum_i \Delta_i^2} \quad (\text{A3})$$

13 The RMS value (variability) of the observational series (of anomalies) in this case is

$$14 \text{RMS}_O = \sqrt{\frac{1}{N} \sum_i O_i^2} \quad (\text{A4})$$

15 One of the diagnostics that we use in Sect. 4 to compare how well different models match up with the observed time series is
 16 given by

$$17 D_{\text{RMSdif}} = \frac{\text{RMSdif}}{\text{RMS}_O} \quad (\text{A5})$$

19 A3 Regression model

20 *Functional form:* The MLR model and fitting methodology used here is similar to the methods used by many others in the
 21 past, with a linear model description that uses annual, semi-annual, QBO, and ENSO terms. Thus, the model function to be fitted
 22 for coefficients a , b , c_n , d_n , f_1 , f_2 , and f_3 has the familiar form:



1
2
$$y(t) = a + b(t - t_m) + \sum_n (c_n \sin 2\pi t/P_n + d_n \cos 2\pi t/P_n)$$

3
$$+ f_1 QBO_1(t) + f_2 QBO_2(t) + f_3 ENSO(t)$$

4 (A6)

5 with the (monthly series) time variable expressed by t , and t_m chosen as the series mid-point; the linear trend term is coefficient b
6 above. The *sine* and *cosine* functions provide for periodic variations with periods P_n . For our work, we use the two primary
7 shorter-term periodic oscillations, annual (12-month period) and semi-annual (6-month period), in Equation (A6). The QBO is
8 also a major source of variability in stratospheric composition time series. As a QBO proxy, we include the variability in
9 monthly mean tropical wind series, as is commonly done; we use the linear combination of (roughly orthogonal) equatorial wind
10 series at 50 hPa and 30 hPa as the QBO_1 and QBO_2 functions above, to account for phase shifts in the series at different
11 locations. Monthly mean zonal equatorial wind data are made available by the Freie Universität Berlin (see <http://www.geo.fu-berlin.de/en/met/ag/strat/produkte/qbo/> for data access information and references). We have also tested fits with the zonal mean
12 wind vertical shear (gradient) rather than the wind itself as a proxy, but this did not make significant changes in the trends (or
13 improvements in the residuals). The ENSO proxy follows the monthly mean multivariate ENSO Index (MEI), which combines
14 data from six main geophysical variables over the tropical Pacific (see Wolter and Timlin, 1993, 1998;
15 <https://www.esrl.noaa.gov/psd/enso/mei/index.html>). Also, since the solar cycle 11-yr term can be highly correlated with the
16 linear term, especially for shorter-term records like the time series from MLS data alone, we have not added a proxy solar term in
17 this study. For further discussions of alternate fitting methods (e.g., methods using effective equivalent chlorine time series as a
18 proxy), the reader is referred to the abundant literature on trend assessments (see WMO, 2014 and references therein). Our main
19 goal here is to retrieve trends and trend errors from the data and the models in the same way. An example of deseasonalized
20 ozone time series in the tropics at two pressure levels is provided in Fig. A3, which shows MLS data and SD-WACCM time
21 series, along with the fits and the linear trends.

22
23 *Trend errors:* For the evaluation of error bars in the linear trends, we have used the method of bootstrap resampling (Efron
24 and Tibshirani, 1986). As others have done for ozone trend analyses (Randel and Thompson, 2011, Bourassa et al., 2014), we
25 have applied this using block bootstrapping (using yearly blocks of data), thereby preserving some of the dependency in the time
26 series. Basically, one samples and (randomly) replaces blocks of yearly data for a large number of resampling cases (on the
27 residuals), and then calculates the standard deviation of the large number of trend results (linear fits) to arrive at the trend
28 uncertainties; note that we use 2σ values as error bars in our trend comparison plots (which is very close to the 95% bounds). We
29 have used 20,000 samples in our bootstrap analyses; changing this number (e.g., by several thousand) does not alter the results
30 significantly, as long as one chooses a large enough total number of resampled cases. An alternative method is to attempt to
31 correct trend uncertainties for the autocorrelation of the residuals after the regression fit (Tiao et al., 1990; Weatherhead et al.,
32 1998; Santer et al., 2000). The existence of non-random residuals effectively implies that the number of independent data points
33 is less than the number making up the original time series. The end result is that trend uncertainties are (and should be) larger
34 than if one neglects these effects. We find that trend errors from this bootstrap method for our time series examples are often
35 larger than the more simplistic/standard calculations by factors ranging from about 1.2 to 2 or more. We have checked our trend
36 error calculations with the OSIRIS team, based on a sample time series, as they have used the same block bootstrap approach
37 (Bourassa et al., 2014).

38

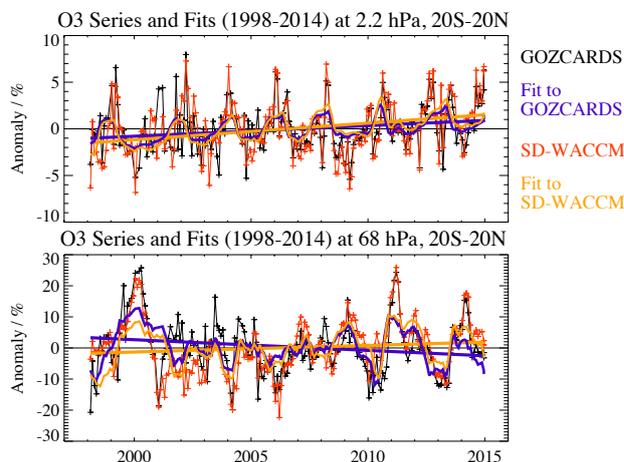


Figure A3. Deseasonalized monthly mean anomaly time series for O₃ (%) at 2.2 hPa (top panel) and 68 hPa (bottom panel), for 1998 through 2014, for averages over 20°S–20°N. Note that the y-axis range for 68 hPa (bottom) is 3 times larger than for 2.2 hPa (top). The data set (black) is from the GOZCARDS ozone record, with the SD-WACCM (simulated) series (red) also shown. Fits to the observational series are in purple, and fits to the model series are in orange; the fitted linear components (straight lines) and the fitted time series functions (curves) are shown.

Data availability. Aura MLS data used in this work are monthly zonal means derived from Level 2 MLS data, which are accessible from the Goddard Earth Sciences Data and Information Services Center (GES DISC), funded by NASA's Science Mission Directorate; a link for MLS Level 2 data access can be found at <https://mls.jpl.nasa.gov/data/>. GOZCARDS data sets can be obtained (by entering GOZCARDS in the search) at <http://disc.gsfc.nasa.gov>. More recent years (and version updates) will be made available at this site, or can be obtained by request to the first author. The WACCM model output used here is provided in some of the references and is available from the NCAR Earth System Grid at <https://www.earthsystemgrid.org/search.html?Project=CCMI1>. For WACCM, we thank NASA Goddard Space Flight Center for the MERRA data (accessed freely online at <http://disc.sci.gsfc.nasa.gov/>).

Author contributions. L. Froidevaux produced the majority of this manuscript, including the Figures. D. Kinnison provided the model runs used in these comparisons to observational data from Aura MLS and GOZCARDS; he also described the models used here and provided substantial guidance and comments for this manuscript. H.-J. Wang and J. Anderson, along with L. Froidevaux, were key participants in the development and creation of the GOZCARDS data sets, including the recently updated version (2.20) for merged ozone used herein. R. Fuller was another key participant in the GOZCARDS data production, and he also provided programming support for these model intercomparisons.

Competing interests. The authors declare that they have no conflicts of interest.

Special issue statement. This manuscript is intended as part of the Chemistry-Climate Modelling Initiative (CCMI) inter-journal special issue (ACP/AMT/ESSD/GMD SI).

Acknowledgements. Work at the Jet Propulsion Laboratory, California Institute of Technology, was performed under contract with the National Aeronautics and Space Administration (NASA). GOZCARDS data were initially produced under the NASA Making Earth System Data Records for Use in Research Environments (MEaSUREs) program, with data continuity/updates now provided under MLS funding. We acknowledge insights and comments regarding aspects of this work from Rolando Garcia, Nathaniel Livesey, Jessica Neu, and Michelle Santee. We also appreciate discussions and comparison work regarding trends and error bars with Doug Degenstein and Chris Roth, as well as Emmanuel Mahieu (and via some of his European contacts). The National Center for Atmospheric Research (NCAR) is sponsored by the U.S. National Science Foundation (NSF). WACCM is a component of NCAR's Community Earth System Model (CESM), which is supported by the NSF and the Office of Science of



1 the U.S. Department of Energy. Computing resources for WACCM were provided by NCAR's Climate Simulation Laboratory,
2 sponsored by NSF and other agencies; we also acknowledge the computational and storage resources of NCAR's Computational
3 and Information Systems Laboratory (CISL).

4

5 References

6

7 Andersson, M. E., Verronen, P. T., Marsh, D. R., Pälvärinta, S.-M., and Plane, J. M. C.: WACCM-D-Improved modeling of
8 nitric acid and active chlorine during energetic particle precipitation, *J. Geophys. Res. Atmos.*, 121, 10,328–10,341,
9 doi:10.1002/2015JD024173, 2016.

10

11 Arnold, F., Buhrke, T., and Qiu, S.: Evidence for stratospheric ozone-depleting heterogeneous chemistry on volcanic aerosols
12 from El Chichon, *Nature*, 348, 49–50, 1990.

13

14 Ball, W. T., Alsing, J., Mortlock, D. J., Rozanov, E. V., Tummon, F., and Haigh, J. D.: Reconciling differences in stratospheric
15 ozone composites, *Atmos. Chem. Phys.*, 17, 12269–12302, <https://doi.org/10.5194/acp-17-12269-2017>, 2017.

16

17 Ball, W. T., Alsing, J., Mortlock, D. J., Staehelin, J., Haigh, J. D., Peter, T., Tummon, F., Stübi, R., Stenke, A., Anderson, J.,
18 Bourassa, A., Davis, S. M., Degenstein, D., Frith, S., Froidevaux, L., Roth, C., Sofieva, V., Wang, R., Wild, J., Yu, P., Ziemke,
19 J. R., and Rozanov, E. V.: Continuous decline in lower stratospheric ozone offsets ozone layer recovery, *Atmos. Chem. Phys.*,
20 18, 1379–1394, <https://doi.org/10.5194/acp-18-1379-2018>, 2018.

21

22 Bandoro, J., Solomon, S., Santer, B. D., Kinnison, D. E., and Mills, M. J.: Detectability of the impacts of ozone-depleting
23 substances and greenhouse gases upon stratospheric ozone accounting for nonlinearities in historical forcings, *Atmos. Chem.*
24 *Phys.*, 18, 143–166, <https://doi.org/10.5194/acp-18-143-2018>, 2018.

25

26 Böhringer, H., Fahey, D. W., Fehsenfeld, F. C., and Ferguson, E. E.: The role of ion–molecule reactions in the conversion of
27 N₂O₅ to HNO₃ in the stratosphere, *Planet. Space. Sci.*, 31, 185–191, 1983.

28

29 Bönisch, H., Engel, A., Birner, Th., Hoor, P., Tarasick, D. W., and Ray, E. A.: On the structural changes in the Brewer–Dobson
30 circulation after 2000, *Atmos. Chem. Phys.*, 11, 3937–3948, doi:10.5194/acp-11-3937-2011, 2011.

31

32 Bourassa, A. E., Degenstein, D. A., Randel, W. J., Zawodny, J. M., Kyrölä, E., McLinden, C. A., Sioris, C. E., and Roth, C. Z.:
33 Trends in stratospheric ozone derived from merged SAGE II and Odin-OSIRIS satellite observations, *Atmos. Chem. Phys.*, 14,
34 6983–6994, doi:10.5194/acp-14-6983-2014, 2014.

35

36 Brinkop, S., Dameris, M., Jockel, P., Garny, H., Lossow, S., and Stiller, G.: The millennium water vapour drop in chemistry-
37 climate model simulations, *Atmos. Chem. Phys.*, 16, 8125–8140, www.atmos-chem-phys.net/16/8125/2016, doi:10.5194/acp-16-
38 8125-2016, 2016.

39

40 Callis, L. B., and Lambeth, J. D.: NO_y formed by precipitating electron events in 1991 and 1992: Descent into the stratosphere as
41 observed by ISAMS, *Geophys. Res. Lett.*, 25, 1875–1878, 1998.

42

43 Calvo, N., Garcia, R. R., and Kinnison, D. E.: Revisiting Southern Hemisphere polar stratospheric temperature trends in
44 WACCM: The role of dynamical forcing, in press, *Geophys. Res., Lett.*, 44, 3402–3410, doi:10.1002/2017GL072792, 2017.

45

46 Carpenter, L. J., Reimann, S., Burkholder, J. B., Clerbaux, C., Hall, B. D., Hossaini, R., Laube, J. C., and Yvon-Lewis, S. A.:
47 Ozone-depleting substances (ODSs) and other gases of interest to the Montreal Protocol, chap. 1, in: *Scientific Assessment of*
48 *Ozone Depletion: 2014*, Global Ozone Research and Monitoring Project – Report No. 55, World Meteorological Organization,
49 Geneva, Switzerland, 2014.

50

51 Damadeo, R. P., Zawodny, J. M., Thomason, L. W., and Iyer, N.: SAGE version 7.0 algorithm: application to SAGE II, *Atmos.*
52 *Meas. Tech.*, 6, 3539–3561, <https://doi.org/10.5194/amt-6-3539-2013>, 2013.

53

54 Davis, S. M., Rosenlof, K. H., Hassler, B., Hurst, D. F., Read, W. G., Vömel, H., Selkirk, H., Fujiwara, M., and Damadeo, R.:
55 The Stratospheric Water and Ozone Satellite Homogenized (SWOOSH) database: a long-term database for climate studies, *Earth*
56 *Syst. Sci. Data*, 8, 461–490, <https://doi.org/10.5194/essd-8-461-2016>, 2016.

57



- 1 de Zafra, R., V. Chan, S. Crewell, C. Trimble, and J. M. Reeves, Millimeter wave spectroscopic measurements over the South
2 Pole, 3, The behavior of stratospheric nitric acid through polar fall, winter and spring, *J. Geophys. Res.*, 102, 1399-1410, 1997.
3
- 4 de Zafra, R. L. and Smyshlyaev, S. P.: On the formation of HNO₃ in the Antarctic mid to upper stratosphere in winter, *J.*
5 *Geophys. Res.*, 106, 23115–23125, doi:10.1029/2000JD000314, 2001.
6
- 7 Dessler, A. E., Schoeberl, M. R., Wang, T., Davis, S. M., Rosenlof, K. H., and Vernier, J.-P.: Variations of stratospheric water
8 vapor over the past three decades, *J. Geophys. Res. Atmos.*, 119, 12,588-12,598, doi:10.1002/2014JD021712, 2014.
9
- 10 Dhomse, S., et al.: Estimates of Ozone Return Dates from Chemistry-Climate Model Initiative Simulations, *Atmos. Chem. Phys.*,
11 in review, 2018.
12
- 13 Douglass, A. R., Prather, M. J., Hall, T. M., S. E. Strahan, S. E., Rasch, P. J., Sparling, L. C.,
14 Coy, L., and Rodriguez, J. M.: Choosing meteorological input for the global modeling initiative assessment of high-speed
15 aircraft, *J. Geophys. Res.*, 104, 27,545-27,564, 1999.
16
- 17 Ern, M., Preusse, P., and Riese, M.: Driving of the SAO by gravity waves as observed from satellite, *Ann. Geophys.*, 33, 483–
18 504, www.ann-geophys.net/33/483/2015/, doi:10.5194/angeo-33-483-2015, 2015.
19
- 20 Eyring, V., et al.: Overview of IGAC/SPARC Chemistry-Climate Model Initiative (CCMI) Community Simulations in Support
21 of Upcoming Ozone and Climate Assessments, *SPARC Newsletter*, 40, 48-66, January 2013, 2013.
22
- 23 Froidevaux, L., Livesey, N. J., Read, W. G., Salawitch, R. J., Waters, J. W., Drouin, B., MacKenzie, I. A., Pumphrey, H. C.,
24 Bernath, P., Boone, C., Nassar, R., Montzka, S., Elkins, J., Cunnold, D., and Waugh, D.: Temporal decrease in upper
25 atmospheric chlorine, *Geophys. Res. Lett.*, 33, L23813, doi:10.1029/2006GL027600, 2006.
26
- 27 Froidevaux, L., Jiang, Y. B., Lambert, A., Livesey, N. J., Read, W. G., Waters, J. W., Fuller, R. A., Marcy, T. P., Popp, P. J.,
28 Gao, R. S., Fahey, D. W., Jucks, K. W., Stachnik, R. A., Toon, G. C., Christensen, L. E., Webster, C. R., Bernath, P. F., Boone,
29 C. D., Walker, K. A., Pumphrey, H. C., Harwood, R. S., Manney, G. L., Schwartz, M. J., Daffer, W.H., Drouin, B. J., Cofield, R.
30 E., Cuddy, D. T., Jarnot, R. F., Knosp, B. W., Perun, V. S., Snyder, W. V., Stek, P. C., Thurstans, R. P., and Wagner, P. A.:
31 Validation of Aura Microwave Limb Sounder HCl measurements, *J. Geophys. Res.*, 113, doi:10.1029/2007JD009025, D15S25,
32 2008a.
33
- 34 Froidevaux, L., Jiang, Y. B., Lambert, A., Livesey, N. J., Read, W. G., Waters, J. W.,
35 Browell, E. V., Hair, J. W., Avery, M. A., McGee, T. J., Twigg, L. W., Sumnicht, G. K., Jucks, K. W., Margitan, J. J., Sen, B.,
36 Stachnik, R. A., Toon, G. C., Bernath, P. F., Boone, C. D., Walker, K. A., Filipiak, M. J., Harwood, R. S., Fuller, R. A., Manney,
37 G. L., Schwartz, M. J., Daffer, W.H., Drouin, B. J., Cofield, R. E., Cuddy, D. T., Jarnot, R. F., Knosp, B. W., Perun, V. S.,
38 Snyder, W. V., Stek, P. C., Thurstans, R. P., and Wagner, P. A.: Validation of Aura Microwave Limb Sounder stratospheric and
39 mesospheric ozone measurements, *J. Geophys. Res.*, 113, doi:10.1029/2007JD008771, D15S20, 2008b.
40
- 41 Froidevaux, L., Anderson, J., Wang, H.-J., Fuller, R. A., Schwartz, M. J., Santee, M. L., Livesey, N. J., Pumphrey, H. C.,
42 Bernath, P. F., Russell III, J. M., and McCormick, M. P.: Global Ozone Chemistry And Related trace gas Data records for the
43 Stratosphere (GOZCARDS): methodology and sample results with a focus on HCl, H₂O, and O₃, *Atmos. Chem. Phys.*, 15,
44 10471–10507, <https://doi.org/10.5194/acp-15-10471-2015>, 2015.
45
- 46 Fueglistaler, S.: Step-wise changes in stratospheric water vapor? *J. Geophys. Res.*, 117, D13302, doi:10.1029/2012JD017582,
47 2012.
48
- 49 Fueglistaler, S. and Haynes, P. H.: Control of interannual and longer-term variability of stratospheric water vapor, *J. Geophys.*
50 *Res.*, 110, 24 108, D24 108, doi:10.1029/2005JD006019, 2005.
51
- 52 Funke, B., Baumgaertner, A., Calisto, M., Egorova, T., Jackman, C. H., Kieser, J., Krivolutsky, A., Lopez-Puertas, M., Marsh,
53 D. R., Reddmann, T., Rozanov, E., Salmi, S.-M., Sinnhuber, M., Stiller, G. P., Verronen, P. T., Versick, S., von Clarmann, T.,
54 Vyushkova, T. Y., Wieters, N., and Wissing, J. M.: Composition changes after the “Halloween” solar proton event: The High-
55 Energy Particle Precipitation in the Atmosphere (HEPPA) model versus MIPAS data intercomparison study, *Atmos. Chem.*
56 *Phys.*, 11(3), 9089–9139, doi:10.5194/acp-11-9089-2011, 2011.
57
- 58 Garcia, R. R., D. Marsh, D. E. Kinnison, B. Boville, and F. Sassi, Simulations of secular trends in the middle atmosphere, 1950-
59 2003, *J. Geophys. Res.*, 112, D09301, doi:10.1029/2006JD007485, 2007.
60



- 1 Garcia, R. R., Smith, A. K., Kinnison, D. E., de la Camara, A. and Murphy, D.: Modification of the gravity wave
2 parameterization in the Whole Atmosphere Community Climate Model: Motivation and results, *J. Atmos. Sci.*, 74, 275-291,
3 doi:10.1175/JAS-D-16-0104.1, 2017.
- 4
- 5 Gebhardt, C., Rozanov, A., Hommel, R., Weber, M., Bovensmann, H., Burrows, J. P., Degenstein, D., Froidevaux, L., and
6 Thompson, A. M.: Stratospheric ozone trends and variability as seen by SCIAMACHY from 2002 to 2012, *Atmos. Chem. Phys.*,
7 14, 831–846, doi:10.5194/acp-14-831-2014, 2014.
- 8
- 9 Gilford, D. M., Solomon, S., and Portmann, R. W.: Radiative impacts of the 2011 abrupt drops in water vapor and ozone in the
10 tropical tropopause layer, *J. Climate*, 29, 595-612, doi:10.1175/JCLI-D-15-0167.1, 2016.
- 11
- 12 Grooß, J.-U., Müller, R., Spang, R., Tritscher, I., Wegner, T., Chipperfield, M. P., Feng, W., Kinnison, D. E., and Madronich, S.:
13 On the discrepancy of HCl processing in the dark polar vortices, *Atmos. Chem. Phys. Discuss.*, [https://doi.org/10.5194/acp-2018-](https://doi.org/10.5194/acp-2018-202)
14 [202](https://doi.org/10.5194/acp-2018-202), 2018.
- 15
- 16 Haenel, F. J., Stiller, G. P., von Clarmann, T., Funke, B., Eckert, E., Glatthor, N., Grabowski, U., Kellmann, S., Kiefer, M.,
17 Linden, A., and Reddmann, T.: Reassessment of MIPAS age of air trends and variability, *Atmos. Chem. Phys.*, 15, 13,161–
18 13,176, doi:10.5194/acp-15-13161-2015, 2015.
- 19
- 20 Harris, N. R. P., Hassler, B., Tummon, F., Bodeker, G. E., Hubert, D., Petropavlovskikh, I., Steinbrecht, W., Anderson, J.,
21 Bhartia, P. K., Boone, C. D., Bourassa, A., Davis, S. M., Degenstein, D., Delcloo, A., Frith, S. M., Froidevaux, L., Godin-
22 Beekmann, S., Jones, N., Kurylo, M. J., Kyrölä, E., Laine, M., Leblanc, S. T., Lambert, J.-C., Liley, B., Mahieu, E., Maycock,
23 A., de Maziere, M., Parrish, A., Querel, R., Rosenlof, K. H., Roth, C., Sioris, C., Staehelin, J., Stolarski, R. S., Stubi, R.,
24 Tamminen, J., Vigouroux, C., Walker, K., Wang, H. J., Wild, J., and Zawodny, J. M.: Past changes in the vertical distribution of
25 ozone - Part 3: Analysis and interpretation of trends, *Atmos. Chem. Phys.*, 15, 9965-9982, [https://doi.org/10.5194/acp-15-9965-](https://doi.org/10.5194/acp-15-9965-2015)
26 [2015](https://doi.org/10.5194/acp-15-9965-2015), 2015.
- 27
- 28 Hegglin, M. I., Tegtmeier, S., Anderson, J., Froidevaux, L., Fuller, R., Funke, B., Jones, A., Lingenfelser, G., Lumpe, J.,
29 Pendlebury, D., Remsberg, E., Rozanov, A., Toohey, M., Urban, J., von Clarmann, T., Walker, K. A., Wang, R., and Weigel, K.:
30 SPARC Data Initiative: Comparison of water vapor climatologies from international satellite limb sounders, *J. Geophys. Res.*
31 *Atmos.*, 118, 11,824–11,846, doi: 10.1002/jgrd.50752, 2013.
- 32
- 33 Hegglin, M. I., Plummer, D. A., Shepherd, T. G., Scinocca, J. F., Anderson, J., Froidevaux, L., Funke, B., Hurst, D. and
34 Rozanov, A., Urban, J., von Clarmann, T., Walker, K. A., Wang, H. J., Tegtmeier, S., and Weigel, K.: Vertical structure of
35 stratospheric water vapour trends derived from merged satellite data, *Nat. Geosci.*, 7, 768–776,
36 <https://doi.org/10.1038/ngeo2236>, 2014.
- 37
- 38 Holton, J. R., and Gettelman, A.: Horizontal transport and the dehydration of the stratosphere, *Geophys. Res. Lett.*, 28, 2799–
39 2802, 2001.
- 40
- 41 Hubert, D., Lambert, J.-C., Verhoelst, T., Granville, J., Keppens, A., Baray, J.-L., Cortesi, U., Degenstein, D. A., Froidevaux, L.,
42 Godin-Beekmann, S., Hoppel, K. W., Kyrölä, E., Leblanc, T., Lichtenberg, G., McElroy, C. T., Murtagh, D., Nakane, H., Russell
43 III, J. M., Salvador, J., Smit, H. G. J., Stebel, K., Steinbrecht, W., Strawbridge, K. B., Stubi, R., Swart, D. P. J., Taha, G.,
44 Thompson, A. M., Urban, J., van Gijsel, J. A. E., von der Gathen, P., Walker, K. A., Wolfram, E., and Zawodny, J. M.: Ground-
45 based assessment of the bias and long-term stability of fourteen limb and occultation ozone profile data records, *Atmos. Meas.*
46 *Tech.*, 9, 2497-2534, <https://doi.org/10.5194/amt-9-2497-2016>, 2016.
- 47
- 48 Hurst, D. F., Oltmans, S. J., Vomel, H., Rosenlof, K. H., Davis, S. M., Ray, E. A., Hall, E. G., and Jordan, A.: Stratospheric
49 water vapor trends over Boulder, Colorado: Analysis of the 30 year Boulder record, *J. Geophys. Res.*, 116, D02306,
50 doi:10.1029/2010JD015065, 2011.
- 51
- 52 Hurst, D. F., Read, W. G., Vömel, H., Selkirk, H. B., Rosenlof, K. H., Davis, S. M., Hall, E. G., Jordan, A. F., and Oltmans, S. J.:
53 Recent divergences in stratospheric water vapor measurements by frost point hygrometers and the Aura Microwave Limb
54 Sounder, *Atmos. Meas. Tech.*, 9, 4447–4457, www.atmos-meas-tech.net/9/4447/2016/, doi:10.5194/amt-9-4447-2016, 2016.
- 55
- 56 Imai, K., Manago, N., Mitsuda, C., Naito, Y., Nishimoto, E., Sakazaki, T., Fujiwara, M., Froidevaux, L., von Clarmann, T.,
57 Stiller, G. P., Murtagh, D. P., Rong, P.-P., Mlynczak, M. G., Walker, K. A., Kinnison, D. E., Akiyoshi, H., Nakamura, T.,
58 Miyasaka, T., Nishibori, T., Mizobuchi, S., Kikuchi, K.-I., Ozeki, H., Takahashi, C., Hayashi, H., Sano, T., Suzuki, M.,
59 Takayanagi, M., and Shiotani, M.: Validation of ozone data from the Superconducting Submillimeter-Wave Limb-Emission
60 Sounder (SMILES), *J. Geophys. Res.-Atmos.*, 118, 5750–5769, doi:10.1002/jgrd.50434, 2013.



- 1
2 Jackman, C. H., Marsh, D. R., Vitt, F. M., Garcia, R. R., Fleming, E. L., Labow, G. J., Randall, C. E., Lopez-Puertas, M., Funke,
3 B., von Clarmann, T., and Stiller, G. P.: Short- and medium-term atmospheric constituent effects of very large solar proton
4 events, *Atmos. Chem. Phys.*, 8, 765–785, doi:10.5194/acp-8-765-2008, 2008.
- 5
6 Jensen, E. and Pfister, L.: Transport and freeze-drying in the tropical tropopause layer, *J. Geophys. Res. Atmos.*, 109, D02207,
7 <http://dx.doi.org/10.1029/2003JD004022>, 2004.
- 8
9 Jiang, Y. B., Froidevaux, L., Lambert, A., Livesey, N. J., Read, W. G., Waters, J. W., Bojkov, B., Leblanc, T., McDermid, I. S.,
10 Godin-Beekmann, S., Filipiak, M. J., Harwood, R. S., Fuller, R. A., Daffer, W. H., Drouin, B. J., Cofield, R. E., Cuddy, D. T.,
11 Jarnot, R. F., Knosp, B. W., Perun, V. S., Schwartz, M. J., Snyder, W. V., Stek, P. C., Thurstans, R. P., Wagner, P. A., Allaart,
12 M., Andersen, S. B., Bodeker, G., Calpini, B., Claude, H., Coetzee, G., Davies, J., De Backer, H., Dier, H., Fujiwara, M.,
13 Johnson, B., Kelder, H., Leme, N. P., Koenig-Langlo, G., Kyro, E., Laneve, G., Fook, L. S., Merrill, J., Morris, G., Newchurch,
14 M., Oltmans, S., Parrondos, M. C., Posny, F., Schmidlin, F., Skrivankova, P., Stubi, R., Tarasick, D., Thompson, A., Thouret, V.,
15 Viatte, P., Vömel, H., von Der Gathen, P., Yela, M., and Zablocki, G.: Validation of the Aura Microwave Limb Sounder Ozone
16 by Ozone-sonde and Lidar Measurements, *J. Geophys. Res.*, 112, D24S34, doi:10.1029/2007JD008776, 2007.
- 17
18 Kawa, S. R., Kumer, J. B., Douglass, A. R., Roche, A. E., Smith, S. E., Taylor, F. W., and
19 Allen, D. J.: Missing chemistry of reactive nitrogen in the upper stratospheric polar winter, *Geophys. Res. Lett.*, 22, 2629-2632,
20 1995.
- 21
22 Kinnison, D. E., Brasseur, G. P., Walters, S., Garcia, R. R., Sassi, F., Boville, B. A., Marsh, D., Harvey, L., Randall, C., Randel,
23 W., Lamarque, J.-F., Emmons, L. K., Hess, P., Orlando, J., Tyndall, J., and Pan, L.: Sensitivity of chemical tracers to
24 meteorological parameters in the MOZART-3 chemical transport model, *J. Geophys. Res.*, 112, D20302,
25 doi:10.1029/2006JD007879, 2007.
- 26
27 Kunz, A., Pan, L., Konopka, P., Kinnison, D., and Tilmes, S.: Chemical and dynamical discontinuity at the extratropical
28 tropopause based on START08 and WACCM analysis, *J. Geophys. Res.*, 116, D24302, doi:10.1029/2011JD016686, 2011.
- 29
30 Kuttippurath, J., Bodeker, G. E., Roscoe, H. K., and Nair, P. J.: A cautionary note on the use of EESC-based regression analysis
31 for ozone trend studies, *Geophys. Res. Lett.*, 42, 162-168, doi:10.1002/2014GL062142, 2015.
- 32
33 Kvissel, O.-K., Orsolini, Y. J., Stordal, F., Isaksen, I. S. A., and Santee, M. L.: Formation of stratospheric nitric acid by a
34 hydrated ion cluster reaction: Implications for the effect of energetic particle precipitation on the middle atmosphere, *J. Geophys.*
35 *Res.*, 117, D16301, doi:10.1029/2011JD017257, 2012.
- 36
37 Lamarque, J.-F., Emmons, L. K., Hess, P. G., Kinnison, D. E., Tilmes, S., Vitt, F., Heald, C. L., Holland, E. A., Lauritzen, P. H.,
38 Neu, J., Orlando, J. J., Rasch, P. J., and Tyndall, G. K.: CAM-chem: description and evaluation of interactive atmospheric
39 chemistry in the Community Earth System Model, *Geosci. Model Dev.*, 5, 369–411, <https://doi.org/10.5194/gmd-5-369-2012>,
40 2012.
- 41
42 Lambert, A., Read, W. G., Livesey, N. J., Santee, M. L., Manney, G. L., Froidevaux, L., Wu, D. L., Schwartz, M. J., Pumphrey,
43 H. C., Jimenez, C., Nedoluha, G. E., Cofield, R. E., Cuddy, D. T., Daffer, W. H., Drouin, B. J., Fuller, R. A., Jarnot, R. F.,
44 Knosp, B. W., Pickett, H. M., Perun, V. S., Snyder, W. V., Stek, P. C., Thurstans, R. P., Wagner, P. A., Waters, J. W., Jucks, K.
45 W., Toon, G. C., Stachnik, R. A., Bernath, P. F., Boone, C. D., Walker, K. A., Urban, J., Murtagh, D., Elkins, J. W., and Atlas,
46 E.: Validation of the Aura Microwave Limb Sounder stratospheric water vapour and nitrous oxide measurements, *J. Geophys.*
47 *Res.*, 112, D24S36, doi:10.1029/2007JD008724, 2007.
- 48
49 Langematz, U., Schmidt, F., Kunze, M., Bodeker, G. E., and Braesicke, P.: Antarctic ozone depletion between 1960 and 1980 in
50 observations and chemistry-climate model simulations, *Atmos. Chem. Phys.*, 16, 15,619-15,627, [www.atmos-chem-](http://www.atmos-chem-phys.net/16/15619/2016/)
51 [phys.net/16/15619/2016/](http://www.atmos-chem-phys.net/16/15619/2016/), doi:10.5194/acp-16-15619-2016, 2016.
- 52
53 Lin, S.-J., A “vertically-Lagrangian” finite-volume dynamical core for global atmospheric models, *Mon. Wea. Rev.*, 132, 2293-
54 2307, 2004.
- 55
56 Livesey, N. J., Filipiak, M. J., Froidevaux, L., Read, W. G., Lambert, A., Santee, M. L., Jiang, J. H., Waters, J. W., Cofield, R. E.,
57 Cuddy, D. T., Daffer, W. H., Drouin, B. J., Fuller, R. A., Jarnot, R. F., Jiang, Y. B., Knosp, B. W., Li, Q. B., Perun, V. S.,
58 Schwartz, M. J., Snyder, W. V., Stek, P. C., Thurstans, R. P., Wagner, P. A., Pumphrey, H. C., Avery, M., Browell, E. V.,
59 Cammas, J.-P., Christensen, L. E., Edwards, D. P., Emmons, L. K., Gao, R.-S., Jost, H.-J., Loewenstein, M.,
60 Lopez, J. D., Nedelec, P., Osterman, G. B., Sachse, G. W., and Webster, C. R.: Validation of Aura Microwave Limb Sounder O₃



- 1 and CO observations in the upper troposphere and lower stratosphere, *J. Geophys. Res.*, 113, D15S02,
2 doi:10.1029/2007JD008805, 2008.
- 3
- 4 Livesey, N. J., Read, W. G., Froidevaux, L., Lambert, A., Manney, G. L., Pumphrey, H. C.,
5 Santee, M. L., Schwartz, M. J., Wang, S., Cofield, R. E., Cuddy, D. T., Fuller, R. A., Jarnot, R. F., Jiang, J. H., Knosp, B. W.,
6 Stek, P. C., Wagner, P. A., and Wu, D. L.: EOS MLS Version 4.2x Level 2 data quality and description document, Tech. rep., Jet
7 Propulsion Laboratory D-33509 Rev. D, available from <http://mls.jpl.nasa.gov/>, 2018.
- 8
- 9 Lopez-Puertas, M., Funke, B., Gil-Lopez, S., von Clarmann, T., Stiller, G. P., Höpfner, M., Kellmann, S., Mengistu Tsidu, G.,
10 Fischer, H., and Jackman, C. H.: HNO₃, N₂O₅ and ClONO₂ Enhancements after the October–November 2003 Solar Proton
11 Events, *J. Geophys. Res.*, 110, A09S44, doi:10.1029/2005JA011051, 2005b.
- 12
- 13 Lossow, S., Khosrawi, F., Nedoluha, G. E., Azam, F., Bramstedt, K., Burrows, J. P., Dinelli, B. M., Eriksson, P., Espy, P. J.,
14 García-Comas, M., Gille, J. C., Kiefer, M., Noël, S., Raspollini, P., Read, W. G., Rosenlof, K. H., Rozanov, A., Sioris, C. E.,
15 Stiller, G. P., Walker, K. A., and Weigel, K.: The SPARC water vapour assessment II: comparison of annual, semi-annual and
16 quasi-biennial variations in stratospheric and lower mesospheric water vapour observed from satellites, *Atmos. Meas. Tech.*, 10,
17 1111–1137, <https://doi.org/10.5194/amt-10-1111-2017>, 2017a.
- 18
- 19 Lossow, S., Garny, H., and Jöckel, P.: An “island” in the stratosphere – on the enhanced annual variation of water vapour in the
20 middle and upper stratosphere in the southern tropics and subtropics, *Atmos. Chem. Phys.*, 17, 11521–11539,
21 <https://doi.org/10.5194/acp-17-11521-2017>, 2017b.
- 22
- 23 Mahieu, E., Chipperfield, M. P., Notholt, J., Reddmann, T., Anderson, J., Bernath, P. F., Blumenstock, T., Coffey, M. T.,
24 Dhomse, S. S., Feng, W., Franco, B., Froidevaux, L., Griffith, D. W. T., Hannigan, J. W., Hase, F., Hossaini, R., Jones, N. B.,
25 Morino, I., Murata, I., Nakajima, H., Palm, M., Paton-Walsh, C., Russell III, J. M., Schneider, M., Servais, C., Smale, D., and
26 Walker, K. A.: Recent Northern Hemisphere stratospheric HCl increase due to atmospheric circulation changes, *Nature*, 515,
27 104–107, doi:10.1038/nature13857, 2014.
- 28
- 29 Marsh, D. R., Mills, M. J., Kinnison, D. E., Lamarque, J.-F., Calvo, N., and Polvani, L. M.: Climate change from 1850 to 2005
30 simulated in CESM1 (WACCM), *Journal of Climate*, 26 (19), doi:10.1175/JCLI-D-12-00558.1, 2013.
- 31
- 32 Matthes, K., Marsh, D. R., Garcia, R. R., Kinnison, D. E., Sassi, F., and Walters, S.: Role of the QBO in modulating the
33 influence of the 11-year solar cycle on the atmosphere using constant forcings, *Journal of Geophysical Research*, 115, D18110.
34 <https://doi.org/10.1029/2009JD013020>, 2010.
- 35
- 36 Millán, L. F., Livesey, N. J., Santee, M. L., Neu, J. L., Manney, G. L., and Fuller, R. A.: Case studies of the impact of orbital
37 sampling on stratospheric trend detection and derivation of tropical vertical velocities: solar occultation vs. limb emission
38 sounding, *Atmos. Chem. and Phys.*, 16, 11 521–11 534, doi:10.5194/acp-16-11521-2016, 2016.
- 39
- 40 Morgenstern, O., Hegglin, M. I., Rozanov, E., O’Connor, F. M., Abraham, N. L., Akiyoshi, H., Archibald, A. T., Bekki, S.,
41 Butchart, N., Chipperfield, M. P., Deushi, M., Dhomse, S. S., Garcia, R. R., Hardiman, S. C., Horowitz, L. W., Joeckel, P., Josse,
42 B., Kinnison, D., Lin, M., Mancini, E., Manyin, M. E., Marchand, M., Marecal, V., Michou, M., Oman, L. D., Pitari, G.,
43 Plummer, D. A., Revell, L. E., Saint-Martin, D., Schofield, R., Stenke, A., Stone, K., Sudo, K., Tanaka, T. Y., Tilmes, S.,
44 Yamashita, Y., Yoshida, K., and Zeng, G.: Review of the global models used within phase 1 of the Chemistry-Climate Model
45 Initiative (CCMI), *Geosci. Model Dev.*, 10, 639–671, <https://doi.org/10.5194/gmd-10-639-2017>, 2017.
- 46
- 47 Nair, P., Godin-Beekmann, S., Kuttippurath, J., Ancellet, G., Goutail, F., Pazmino, A., Froidevaux, L., Zawodny, J. M., Evans,
48 R. D., Wang, H. J., Anderson, J., and Pastel, M.: Ozone trends derived from the total column and vertical profiles at a northern
49 mid-latitude station, *Atmos. Chem. Phys.*, 13, 10373–10384, www.atmos-chem-phys.net/13/10373/2013/, doi:10.5194/acp-13-
50 10373-2013, 2013.
- 51
- 52 Nair, P. J., Froidevaux, L., Kuttippurath, J., Zawodny, J. M., Russell III, J. M., Steinbrecht, W., Claude, H., Leblanc, T., van
53 Gijssels, J. A. E., Johnson, B., Swart, D. P. J., Thomas, A., Querel, R., Wang, R., and Anderson, J.: Subtropical and midlatitude
54 ozone trends in the stratosphere: Implications for recovery, *J. Geophys. Res.*, 120, 7247–7257, doi:10.1002/2014JD022371, 2015.
- 55
- 56 Nedoluha, G. E., Gomez, R. M., Hicks, B. C., Wrotny, J. E., Boone, C., and Lambert, A.: Water vapor measurements in the
57 mesosphere from Mauna Loa over solar cycle 23, *J. Geophys. Res.*, 114, D23303, doi:10.1029/2009JD012504, 2009.
- 58
- 59 Nisbet, E. G., Dlugokencky, E. J., Manning, M. R., Lowry, D., Fisher, R. E., France, J. L.,
60 Michel, S. E., Miller, J. B., White, J. W. C., Vaughn, B., Bousquet, P., Pyle, J. A., Warwick, N. J., Cain, M., Brownlow, R.,



- 1 Zazzeri, G., Lanoisellé, M., Manning, A. C., Gloor, E., Worthy, D. E. J., Brunke, E.-G., Labuschagne, C., Wolff, E. W., and
2 Ganesan, A. L.: Rising atmospheric methane: 2007–2014 growth and isotopic shift, *Global Biogeochem. Cycles*, 30, 1356–1370,
3 doi:10.1002/2016GB005406, 2016.
- 4
- 5 Oltmans, S. J., Vömel, H., Hofmann, D. J., Rosenlof, K. H., and Kley, D.: The increase in stratospheric water vapor from
6 balloonborne, frostpoint hygrometer measurements at Washington, D.C., and Boulder, Colorado, *Geophys. Res. Lett.*, 27, No.
7 21, 3453–3456, 2000.
- 8
- 9 Oram, D. E., et al.: A growing threat to the ozone layer from short-lived anthropogenic chlorocarbons, *Atmos. Chem. Phys.*, 17,
10 11929–11941, <https://doi.org/10.5194/acp-17-11929-2017>, 2017.
- 11
- 12 Orsolini, Y. J., Manney, G. L., Santee, M. L., and Randall, C. E.: An upper stratospheric layer of enhanced HNO₃ following
13 exceptional solar storms, *Geophys. Res. Lett.*, 32, L12S01, doi:10.1029/2004GL021588, 2005.
- 14
- 15 Parrondo, M. C., Gil, M., Yela, M., B. J. Johnson, B. J., and Ochoa, H. A.: Antarctic ozone variability inside the polar
16 vortex estimated from balloon measurements, *Atmos. Chem. Phys.*, 14, 217–229, www.atmos-chem-phys.net/14/217/2014/,
17 doi:10.5194/acp-14-217-2014, 2014.
- 18
- 19 Perliski, L. M., Solomon, S., and London, J.: On the interpretation of seasonal variations of stratospheric ozone, *Planet. Space*
20 *Sci.*, 37, 12, 1527–1538, 1989.
- 21
- 22 Popp, P. J., Marcy, T. P., Watts, L. A., Gao, R. S., Fahey, D. W., Weinstock, E. M., Smith, J. B., Herman, R. L., Troy, R. F.,
23 Webster, C. R., Christensen, L. E., Baumgardner, D. G., Voigt, C., Kärcher, B., Wilson, J. C., Mahoney, M. J., Jensen, E. J., and
24 Bui, T. P.: Condensed-phase nitric acid in a tropical subvisible cirrus cloud, *Geophys. Res. Lett.*, 34, L24812,
25 doi:10.1029/2007GL031832, 2007.
- 26
- 27 Popp, P. J., et al.: Stratospheric correlation between nitric acid and ozone, *J. Geophys. Res.*, 114, D03305,
28 doi:10.1029/2008JD010875, 2009.
- 29
- 30 Randall, C. E., Harvey, V. L., Singleton, C. S., Bailey, S. M., Bernath, P. F., Codrescu, M., Nakajima, H., and Russell III, J. M.:
31 Energetic particle precipitation effects on the southern hemisphere stratosphere in 1992–2005, *J. Geophys. Res.*, 112, D08308,
32 doi:10.1029/2006JD007696, 2007.
- 33
- 34 Randel, W. J., Wu, F., and Gaffen, D. J.: Interannual variability of the tropical tropopause derived from radiosonde data and
35 NCEP reanalysis, *J. Geophys. Res.*, 105, 15,509–15,523, 2000.
- 36
- 37 Randel, W. J., Wu, F., Oltmans, S. J., Rosenlof, K., and Nedoluha, G. E.: Interannual changes of stratospheric water vapor and
38 correlations with tropical tropopause temperatures, *J. Atmos. Sci.*, 61, 2133–2148, 2004.
- 39
- 40 Randel, W. J., Wu, F., Vömel, H., Nedoluha, G. E., and Forster, P.: Decreases in stratospheric water vapor after 2001: Links to
41 changes in the tropical tropopause and the Brewer–Dobson circulation, *J. Geophys. Res.*, 111, D12312,
42 doi:10.1029/2005JD006744, 2006.
- 43
- 44 Randel, W. J., and Thompson, A. M.: Interannual variability and trends in tropical ozone derived from SAGE II satellite data and
45 SHADOZ ozonesondes, *J. Geophys. Res.*, 116, D07303, doi:10.1029/2010JD015195, 2011.
- 46
- 47 Randel, W. J., and Jensen, E. J.: Physical processes in the tropical tropopause layer and their roles in a changing climate, *Nat.*
48 *Geosci.*, 6, 169–176, doi:10.1038/ngeo1733, 2013.
- 49
- 50 Ray, E. A., Holton, J. R., Fishbein, E., Froidevaux, L., and Waters, J.: The tropical semiannual oscillations in temperature and
51 ozone as observed by the MLS, *J. Atmos. Sci.*, 51, 3045–3052, doi:10.1175/1520-0469(1994)051<3045:TTSOIT>2.0.CO;2,
52 1994.
- 53
- 54 Read, W. G., Lambert, A., Bacmeister, J., Cofield, R. E., Christensen, L. E., Cuddy, D. T., Daffer, W. H., Drouin, B. J., Fetzer,
55 E., Froidevaux, L., Fuller, R., Herman, R., Jarnot, R. F., Jiang, J. H., Jiang, Y. B., Kelly, K., Knosp, B. W., Kovalenko, L. J.,
56 Livesey, N. J., Liu, H.-C., Manney, G. L., Pickett, H. M., Pumphrey, H. C., Rosenlof, K. H., Sabouchi, X., Santee, M. L.,
57 Schwartz, M. J., Snyder, W. V., Stek, P. C., Su, H., Takacs, L. L., Thurstans, R. P., Vömel, H., Wagner, P. A., Waters, J. W.,
58 Webster, C. R., Weinstock, E. M., and Wu, D. L.: Aura Microwave Limb Sounder upper tropospheric and lower stratospheric
59 H₂O and relative humidity with respect to ice validation, *J. Geophys. Res.*, 112, D24S35, doi:10.1029/2007JD008752, 2007.
- 60



- 1 Read, W. G., Schwartz, M. J., Lambert, A., Su, H., Livesey, N. J., Daffer, W. H., and
2 Booe, C. D.: The roles of convection, extratropical mixing, and in-situ freeze-drying in the Tropical Tropopause Layer, Atmos.
3 Chem. Phys., 8, 6051–6067, doi:10.5194/acp-8-6051-2008, 2008.
- 4
5 Reddmann, T., Ruhnke, R., Versick, S., and Kouker, W.: Modeling disturbed stratospheric chemistry during solar-induced NO_x
6 enhancements observed with MIPAS/ENVISAT, J. Geophys. Res., 115, D00I11, doi:10.1029/2009JD012569, 2010.
- 7
8 Rienecker, M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M. G., Schubert, S. D., Takacs, L.,
9 Kim, G.-K., Bloom, S., Chen, J., Collins, D., Conaty, A., da Silva, A., Gu, W., Joiner, J., Koster, R. D., Lucchesi, R., Molod, A.,
10 Owens, T., Pawson, S., Pegion, P., Redder, C. R., Reichle, R., Robertson, J., F. R., Ruddick, A. G., Sienkiewicz, M., and
11 Woollen, J.: MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications, J. Climate, 24, 3624–3648,
12 doi:10.1175/JCLI-D-11-00015.1, 2011.
- 13
14 Rinsland, C. P., Gunson, M. R., Abrams, M. C., Lowes, L. L., Zander, R., Mahieu, E., Goldman, A., Ko, M. K W., Rodriguez, J.
15 M., and Sze, N. D.: Heterogeneous conversion of N₂O₅ to HNO₃ in the post-Mount Pinatubo eruption stratosphere, J. Geophys.
16 Res., 99, 8213-8219, 1994.
- 17
18 Rohs, S., Schiller, C., Riese, M., Engel, A., Schmidt, U., Wetter, T., Levin, I., Nakazawa, T., and Aoki, S.: Long-term changes of
19 methane and hydrogen in the stratosphere in the period 1978–2003 and their impact on the abundance of stratospheric water
20 vapor, J. Geophys. Res., 111, D14315, doi:10.1029/2005JD006877, 2006.
- 21
22 Rosenlof, K. H., and Reid, G. R.: Trends in the temperature and water vapor content of the tropical lower stratosphere: Sea
23 surface connection, J. Geophys. Res., 113, D06107, doi:10.1029/2007JD009109, 2008.
- 24
25 Ryan, N. J., Kinnison, D. E., Garcia, R. R., Hoffmann, C. G., Palm, M., Raffalski, U., and Notholt, J.: Assessing the ability to
26 derive rates of polar middle-atmospheric descent using trace gas measurements from remote sensors, Atmos. Chem. Phys., 18,
1457–1474, <https://doi.org/10.5194/acp-18-1457-2018>, 2018.
- 27
28 Santee, M. L., Lambert, A., Read, W. G., Livesey, N. J., Cofield, R. E., Cuddy, D. T., Daffer, W. H., Drouin, B. J., Froidevaux,
29 L., Fuller, R. A., Jarnot, R. F., Knosp, B. W., Manney, G. L., Perun, V. S., Snyder, W. V., Stek, P. C., Thurstans, R. P., Wagner,
30 P. A., Waters, J. W., Muscari, G., de Zafra, R. L., Dibb, J. E., Fahey, D. W., Popp, P. J., Marcy, T. P., Jucks, K. W., Toon, G. C.,
31 Stachnik, R. A., Bernath, P. F., Boone, C. D., Walker, K. A., Urban, J., and Murtagh, D.: Validation of the Aura Microwave Limb
32 Sounder HNO₃ measurements, J. Geophys. Res., 112, D24S40, doi:10.1029/2007JD008, 2007.
- 33
34 Santer, B. D., Wigley, T. M. L., Boyle, J. S., Gaffen, D. J., Hnilo, J. J., Nychka, D., Parker, D. E., and Taylor, K. E.: Statistical
35 significance of trends and trend differences in layer-average atmospheric temperature time series, J. Geophys. Res., 105, 7337-
36 7356, 2000.
- 37
38 Scaife, A. A., Jackson, D. R., R. Swinbank, R., Butchart, N., Thornton, H. E., Keil, M., and Henderson, L.: Stratospheric
39 Vacillations and the Major Warming over Antarctica in 2002, J. Atmos. Sci., 62, 629-639, 2005.
- 40
41 Schaefer, H., Mikaloff Fletcher, S. E., Veidt, C., Lassey, K. R., Brailsford, G. W., Bromley, T. M., Dlugokencky, E. J., Michel,
42 S. E., Miller, J. B., Levin, I., Lowe, D. C., Martin, R. J., Vaughn, B. H., and White, J. W. C.: A 21st century shift from fossil-
43 fuel to biogenic methane emissions indicated by ¹³CH₄, Science, doi:10.1126/science.aad2705, 2016.
- 44
45 Scherer, M., Vömel, H., Fueglistaler, S., Oltmans, S. J., and Staehelin, J.: Trends and variability of midlatitude stratospheric
46 water vapour deduced from the re-evaluated Boulder balloon series and HALOE, Atmos. Chem. Phys., 8, 1391-1402,
47 <https://doi.org/10.5194/acp-8-1391-2008>, 2008.
- 48
49 Schoeberl, M. R., Douglass, A. R., Newman, P. A., Lait, L. R., Lary, D., Waters, J., Livesey, N., Froidevaux, L., Lambert, A.,
50 Read, W., Filipiak, M. J., and Pumphrey, H. C.: QBO and annual cycle variations in tropical lower stratosphere trace gases from
51 HALOE and Aura MLS observations, J. Geophys. Res., 113, D05301, doi:10.1029/2007JD008678, 2008.
- 52
53 Schoeberl, M. R., Dessler, A. E., and Wang, T.: Modeling upper tropospheric and lower stratospheric water vapor anomalies,
54 Atmos. Chem. Phys., 13, 7783–7793, <https://doi.org/10.5194/acp-13-7783-2013>, 2013.
- 55
56 Siskind, D. E., Nedoluha, G. E., Randall, C. E., Fromm, M., and Russell III, J. M.: An assessment of Southern Hemisphere
57 stratospheric NO_x enhancements due to transport from the upper atmosphere, Geophys. Res. Lett., 27, 329–332, 2000.



- 1 Smith, A. K., Garcia, R. R., Moss, A. C., and Mitchell, N. J.: The semiannual oscillation of the tropical zonal wind in the middle
2 atmosphere derived from satellite geopotential height retrievals, *J. Atmos. Sci.*, 74, 2413-2425, doi:10.1175/JAS-D-17-0067.1,
3 2017.
- 4
- 5 Solomon, S., Rosenlof, K., Portmann, R., Daniel, J., Davis, S., Sanford, T., and Plattner, G.-K.: Contributions of Stratospheric
6 Water Vapor to Decadal Changes in the Rate of Global Warming, *Science*, 237, 1219-1223, doi:10.1126/science.1182488, 2010.
- 7
- 8 Solomon, S., Kinnison, D. E., Bandoro, J., and Garcia, R.: Simulations of Polar Ozone Depletion: An Update, *J. Geophys. Res.*,
9 120, 7958-7974, doi:10.1002/2015JD0233652015, 2015.
- 10
- 11 Solomon, S., Ivy, D. J., Kinnison, D., Mills, M. J., Neely III, R. R., and Schmidt, A.: Emergence of healing in the Antarctic
12 ozone layer, *Science*, 353, 269-274, 2016.
- 13
- 14 SPARC, 2010: SPARC CCMVal Report on the Evaluation of Chemistry-Climate Models, Eyring, V., Shepherd, T., and Waugh,
15 D. (Eds.), SPARC Report No. 5, WCRP-30/2010, WMO/TD – No. 40, available at www.sparc-climate.org/publications/sparc-reports/, 2010.
- 16
- 17
- 18 SPARC, 2017: The SPARC Data Initiative: Assessment of stratospheric trace gas and aerosol climatologies from satellite limb
19 sounders, M. I. Hegglin and S. Tegtmeier (Eds.), SPARC Report No. 8, WCRP-05/2017, doi:10.3929/ethz-a-010863911,
20 available at www.sparc-climate.org/publications/sparc-reports/, 2017.
- 21
- 22 Steinbrecht, W., Froidevaux, L., Fuller, R., Wang, R., Anderson, J., Roth, C., Bourassa, A., Degenstein, D., Damadeo, R.,
23 Zawodny, J., Frith, S., McPeters, R., Bhartia, P., Wild, J., Long, C., Davis, S., Rosenlof, K., Sofieva, V., Walker, K., Rahpoe, N.,
24 Rozanov, A., Weber, M., Laeng, A., von Clarmann, T., Stiller, G., Kramarova, N., Godin-Beekmann, S., Leblanc, T., Querel, R.,
25 Swart, D., Boyd, I., Hocke, K., Kämpfer, N., Maillard Barras, E., Moreira, L., Nedoluha, G., Vigouroux, C., Blumenstock, T.,
26 Schneider, M., Garcia, O., Jones, N., Mahieu, E., Smale, D., Kotkamp, M., Robinson, J., Petropavlovskikh, I., Harris, N.,
27 Hassler, B., Hubert, D. and Tummon, F.: An update on ozone profile trends for the period 2000 to 2016, *Atmos. Chem. Phys.*
28 *Discuss.*, 2017, 1–24, doi:10.5194/acp-2017-391, 2017.
- 29
- 30 Stiller, G. P., Mengistu Tsidu, G., von Clarmann, T., Glatthor, N., Höpfner, M., Kellmann, S., Linden, A., Ruhnke, R., Fischer,
31 H., Lopez-Puertas, M., Funke, B., and Gil-Lopez, S.: An enhanced HNO₃ second maximum in the Antarctic midwinter upper
32 stratosphere 2003, *J. Geophys. Res.*, 110, D20303, doi:10.1029/2005JD006011, 2005.
- 33
- 34 Stolarski, R., Douglass, A., and Strahan, S. E.: Using satellite measurements of N₂O to remove dynamical variability from HCl
35 measurements, *Atmos. Chem. Phys.*, 18, 5691–5697, <https://doi.org/10.5194/acp-18-5691-2018>, 2018.
- 36
- 37 Stone, K. A., Solomon, S., and Kinnison, D. E.: On the identification of ozone recovery, *Geophys. Res. Lett.*, in press,
38 doi:10.1029/2018GL077955, 2018.
- 39
- 40 Strahan, S. E., Oman, L. D., Douglass, A. R., and L. Coy, L.: Modulation of Antarctic vortex composition by the quasi-biennial
41 oscillation, *Geophys. Res. Lett.*, 42, 4216–4223, doi:10.1002/2015GL063759, 2015.
- 42
- 43 Tegtmeier, S., Hegglin, M. I., Anderson, J., Bourassa, A., Brohede, S., Degenstein, D.,
44 Froidevaux, L., Fuller, R., Funke, B., Gille, J., Jones, A., Kasai, Y., Krüger, K., Kyrölä, E., Lingenfelter, G., Lumpe, J., Nardi,
45 B., Neu, J., Pendlebury, D., Remsberg, E., Rozanov, A., Smith, L., Toohey, M., Urban, J., von Clarmann, T., Walker, K. A., and
46 Wang, H. J.: The SPARC Data Initiative: A comparison of ozone climatologies from international satellite limb sounders, *J.*
47 *Geophys. Res. Atmos.*, 118, 12,229–12,247, doi: 10.1002/2013JD019877, 2013.
- 48
- 49 Tian, W., and Chipperfield, M. P.: Stratospheric water vapor trends in a coupled chemistry-climate model, *Geophys. Res. Lett.*,
50 33, L06819, doi:10.1029/2005GL024675, 2006.
- 51
- 52 Tiao, G. C., Reinsel, G. C., Xu, D., Pedrick, J. H., Zhu, X., Miller, A. J., DeLuisi, J. J., Mateer, C. L., and Wuebbles, D. J.:
53 Effects of autocorrelation and temporal sampling schemes on estimates of trend and spatial correlation, *J. Geophys. Res.*, 95,
54 20,507-20,517, 1990.
- 55
- 56 Tilmes, S., Lamarque, J.-F., Emmons, L. K., Kinnison, D., Marsh, D., Garcia, R. R., Smith, A. K., Neeley, R. R., Conley, A.,
57 Vitt, F., Val Martin, M., Tanimoto, H., Simpson, I., Blake, D. R., and Blake, N.: Representation of the Community Earth System
58 Model (CESM1) CAM4-Chem within the Chemistry-Climate Model Initiative, *Geosci. Model Dev.*, 9, 1853-1890,
59 doi:10.5194/gmd-9-1853-2016, 2016.
- 60

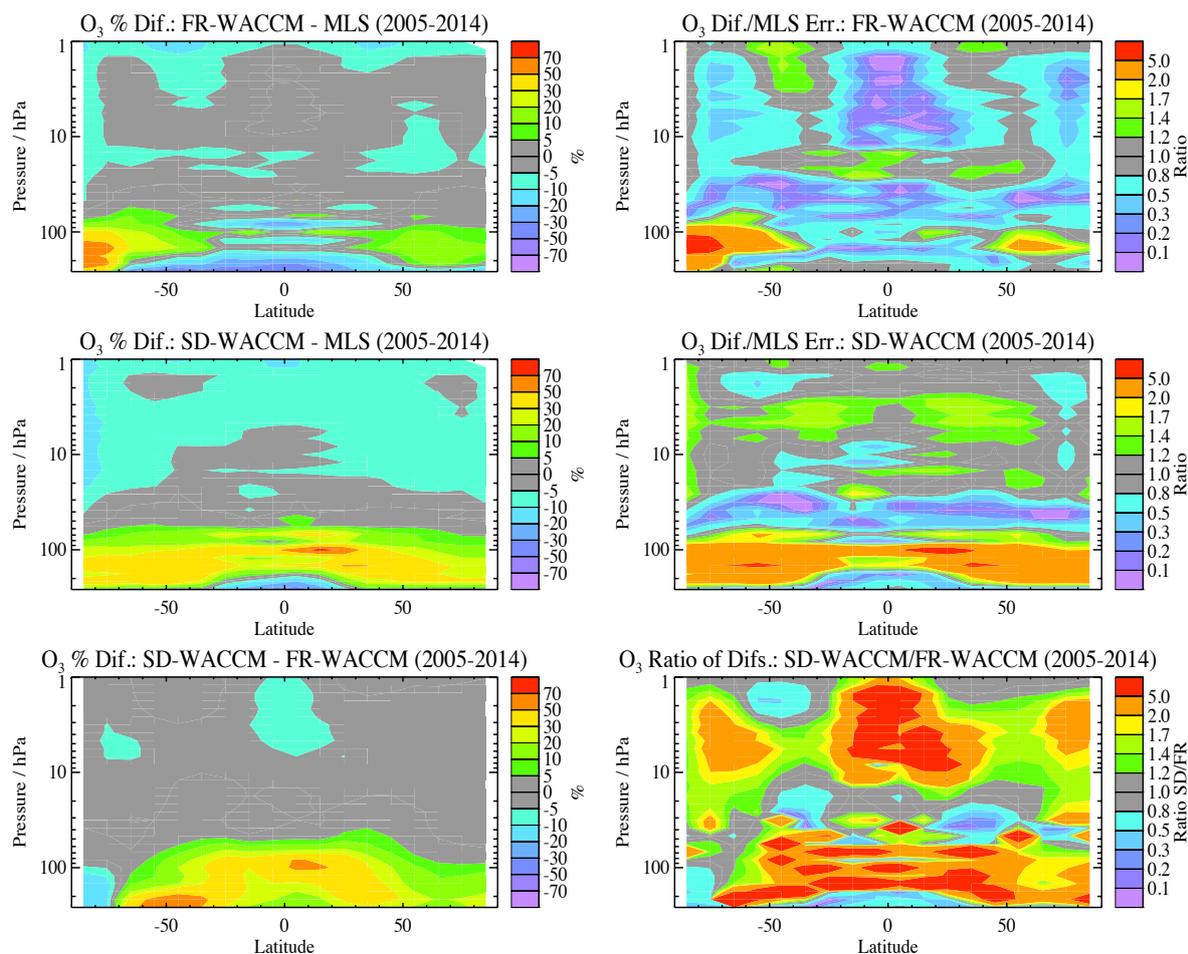


- 1 Toohy, M., Hegglin, M. I., Tegtmeier, S., Anderson, J., Añel, J. A., Bourassa, A., Brohede, S., Degenstein, D., Froidevaux, L.,
2 Fuller, R., Funke, B., Gille, J., Jones, A., Kasai, Y., Krüger, K., Kyrölä, E., Neu, J. L., Rozanov, A., Smith, L., Urban, J., von
3 Clarmann, T., Walker, K. A., and Wang, R.: Characterizing sampling bias in the trace gas climatologies of the SPARC Data
4 Initiative, *J. Geophys. Res. Atmos.*, 118, 11,847–11,862, doi: 10.1002/jgrd.5087, 2013.
- 5
6 Tummon, F., Hassler, B., Harris, N. R. P., Staehelin, J., Steinbrecht, W., Anderson, J., Bodeker, G. E., Bourassa, A., Davis, S.
7 M., Degenstein, D., Frith, S. M., Froidevaux, L., Kyrölä, E., Laine, M., Long, C., Penckwitt, A. A., Sioris, C. E., Rosenlof, K.
8 H., Roth, C., Wang, H.-J., and Wild, J.: Intercomparison of vertically resolved merged satellite ozone data sets: interannual
9 variability and long-term trends, *Atmos. Chem. Phys.*, 15, 3021-3043, <https://doi.org/10.5194/acp-15-3021-2015>, 2015.
- 10
11 Urban, J., Lossow, S., Stiller, G., and Read, W.: Another drop in water vapor, *EOS Transactions, American Geophysical Union*,
12 95, 27, 245-252, doi:10.1002/2014EO270001, 2014.
- 13
14 Verronen, P. T., Santee, M. L., Manney, G. L., Lehmann, R., Salmi, S.-M. and A. Seppälä.: Nitric acid enhancements in the
15 mesosphere during the January 2005 and December 2006 solar proton events, *J. Geophys. Res.*, 116, D17301,
16 doi:10.1029/2011JD016075, 2011.
- 17
18 von Clarmann, T., Glatthor, N., Höpfner, M., Kellmann, S., Ruhnke, R., Stiller, G. P., and Fischer, H.: Experimental evidence of
19 perturbed odd hydrogen and chlorine chemistry after the October 2003 solar proton events, *J. Geophys. Res.*, 110, A09S45,
20 doi:10.1029/2005JA011053, 2005.
- 21
22 Vömel, H., Barnes, J. E., Forno, R. N., Fujiwara, M., Hasebe, F., Iwasaki, S., Kivi, R., Komala, N., Kyrölä, E., Leblanc, T.,
23 Morel, B., Ogino, S.-Y., Read, W. G., Ryan, S. C., Saraspriya, S., Selkirk, H., Shiotani, M., Valverde Canossa, J., and
24 Whiteman, D. N.: Validation of Aura Microwave Limb Sounder water vapor by balloon-borne Cryogenic Frost point
25 Hygrometer measurements, *J. Geophys. Res.*, 112, D24S37, doi:10.1029/2007JD008698, 2007.
- 26
27 Wang, Y., Su, H., Jiang, J. H., Livesey, N. J., Santee, M. L., Froidevaux, L., Read, W. G., and Anderson, J.: The linkage between
28 stratospheric water vapor and surface temperature in an observation-constrained coupled general circulation model, *Clim. Dyn.*,
29 doi:10.1007/s00382-016-3231-3, 2016.
- 30
31 Waters, J. W., Froidevaux, L., Read, W. G., Manney, G. L., Eslon, L. S., Flower, D. A., Jarnot, R. F., and Harwood, R. S.:
32 Stratospheric ClO and ozone from the Microwave Limb Sounder on the Upper Atmosphere Research Satellite, *Nature*, 362, 597-
33 602, 1993.
- 34
35 Waters, J. W., Froidevaux, L., Harwood, R. S., Jarnot, R. F., Pickett, H. M., Read, W. G.,
36 Siegel, P. H., Coffield, R. E., Filipiak, M. J., Flower, D. A., Holden, J. R., Lau, G. K., Livesey, N. J., Manney, G. L., Pumphrey,
37 H. C., Santee, M. L., Wu, D. L., Cuddy, D. T., Lay, R. R., Loo, M. S., Perun, V. S., Schwartz, M. J., Stek, P. C., Thurstans, R. P.,
38 Boyles, M. A., Chandra, S., Chavez, M. C., Chen, G.-S., Chudasama, B. V., Dodge, R., Fuller, R. A., Girard, M. A., Jiang, J. H.,
39 Jiang, Y., Knosp, B. W., LaBelle, R. C., Lam, J. C., Lee, K. A., Miller, M., Oswald, J. E., Patel, N. C., Pukala, D. M., Quintero,
40 O., Scaff, D. M., Snyder, W. V., Tope, M. C., Wagner, P. A., and Walch, M. J.: The Earth Observing System Microwave Limb
41 Sounder (EOS MLS) on the Aura satellite, *IEEE Trans. Geosci. Remote Sens.*, 44 (5), 1075–1092,
42 doi:10.1109/TGRS.2006.873771, 2006.
- 43
44 Waugh, D. W., Considine, D. B., and Fleming, E. L.: Is upper stratospheric chlorine decreasing as expected?, *Geophys. Res.*
45 *Let.*, 28(7), 1187–1190, doi:10.1029/2000GL011745, 2001.
- 46
47 Waugh, D. W., and Eyring, V.: Quantitative performance metrics for stratospheric-resolving chemistry-climate models, *Atmos.*
48 *Chem. Phys.*, 8, 5699–5713, www.atmos-chem-phys.net/8/5699/2008/, 2008.
- 49
50 Weatherhead, E. C., Reinsel, G. C., Tiao, G. C., Meng, X.-L., Choi, D., Cheang, W.-K., Keller, T., DeLuisi, J., Wuebbles, D. J.,
51 Kerr, J. B., Miller, A. J., Oltmans, S. J., and Frederick, J. E.: Factors affecting the detection of trends: Statistical considerations
52 and applications to environmental data, *J. Geophys. Res.-Atmos.*, 103, 17149–17161, <https://doi.org/10.1029/98JD00995>, 1998.
- 53
54 Wegner, T., Kinnison, D. E., Garcia, R. R., and Solomon, S.: Simulation of polar stratospheric clouds in the specified dynamics
55 version of the whole atmosphere community climate model, *J. Geophys. Res.*, 4991–5002, <https://doi.org/10.1002/jgrd.50415>,
56 2013.
- 57
58 WMO: Scientific Assessment of Ozone Depletion: 2014, Global Ozone Research and Monitoring Project-Report No. 55, WMO
59 (World Meteorological Organization), Geneva, Switzerland, available at: <https://www.esrl.noaa.gov/csd/assessments/ozone>,
60 2014.



1

2



3

4 **Figure 1.** The left panels show percent differences ((model-data) divided by data) for binned climatological average O_3 from
5 2005 through 2014 (see Fig. S1 for these averages), for (top panel) the free-running model (average of 3 realizations) FR-
6 WACCM, (middle panel) the specified dynamics model version SD-WACCM, and (bottom panel) SD-WACCM minus FR-
7 WACCM, also as a percent difference. The two top right panels give ratios of the absolute value of average model (FR-WACCM
8 in top panel, SD-WACCM in middle panel) minus average MLS O_3 to the MLS systematic O_3 errors, based on the climatological
9 fields and estimated MLS errors (2σ estimates, see text and Fig. S2). The bottom right panel gives the ratios of average absolute
10 differences between SD-WACCM and MLS to those differences for FR-WACCM and MLS; e.g., this shows that upper
11 stratospheric tropical SD-WACCM mean O_3 is larger than O_3 from FR-WACCM, which is why SD-WACCM matches MLS
12 better there (see top 2 right panels).

13

14

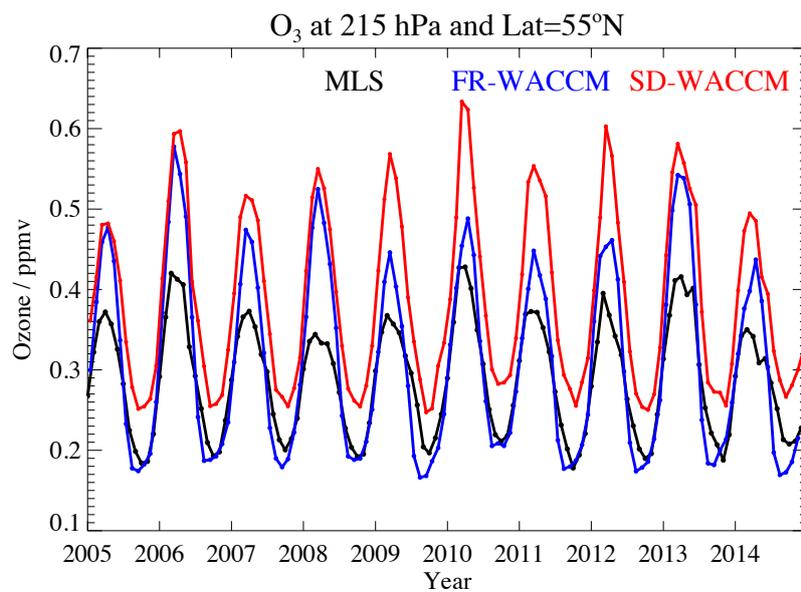
15

16



1

2



3

4

5 **Figure 2.** Monthly mean ozone time series at 215 hPa and the 55°N latitude bin (for averages over 50°N-60°N) from MLS, FR-
6 WACCM, and SD-WACCM (see legend for color coding).

7

8

9

10

11

12

13

14

15

16

17

18

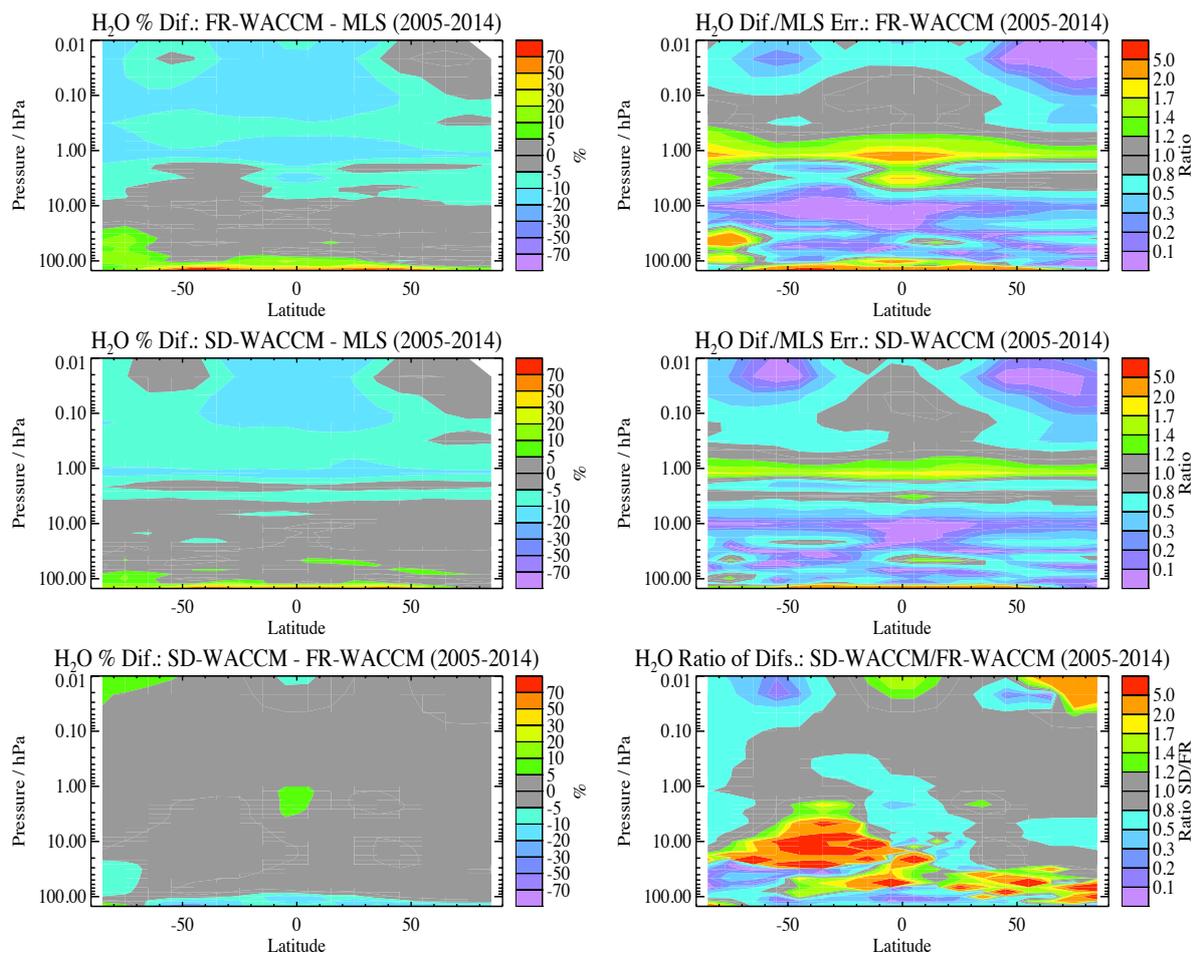
19

20



1

2



3

4

5 **Figure 3.** Same as Fig. 1, but for stratospheric and mesospheric water vapor.

6

7

8

9

10

11

12



1
2

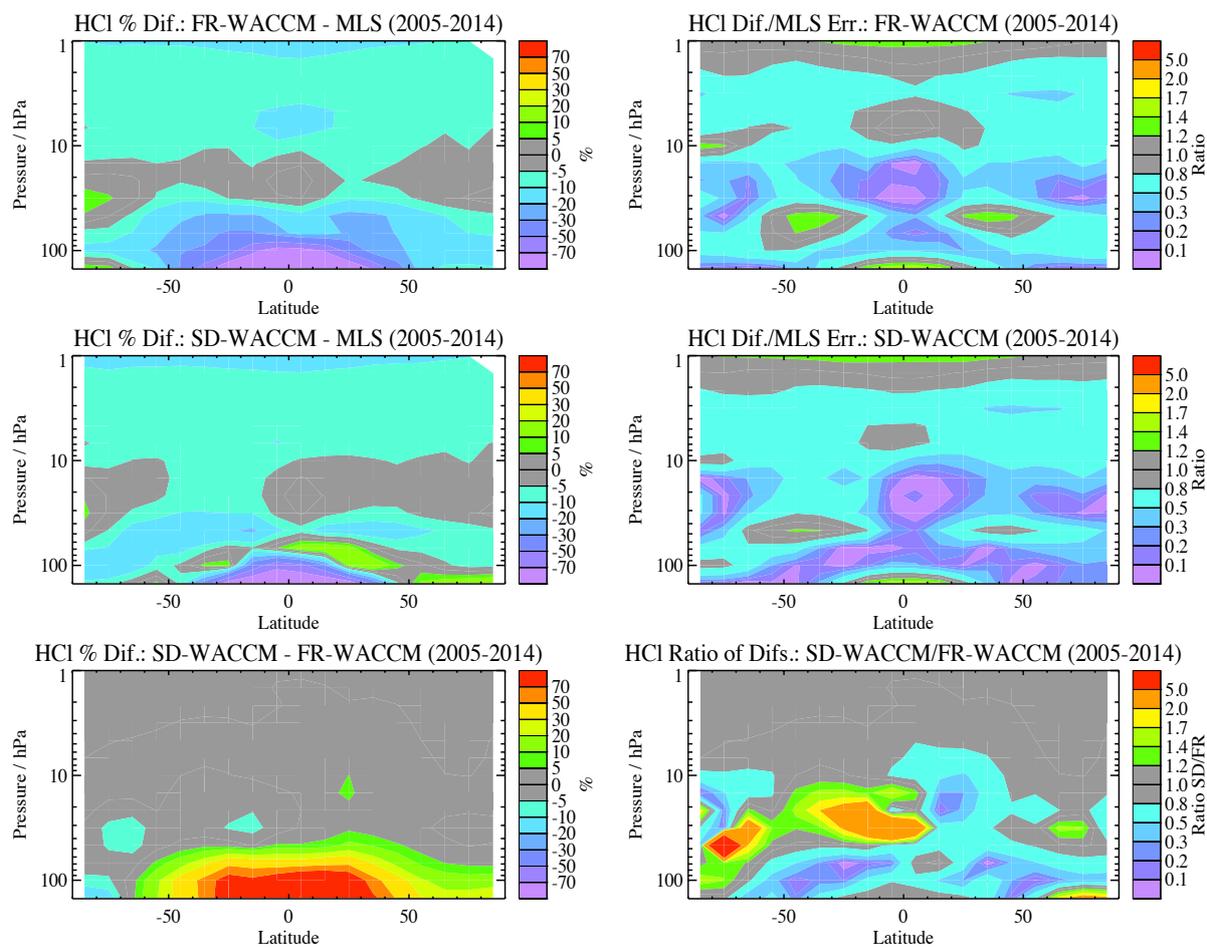


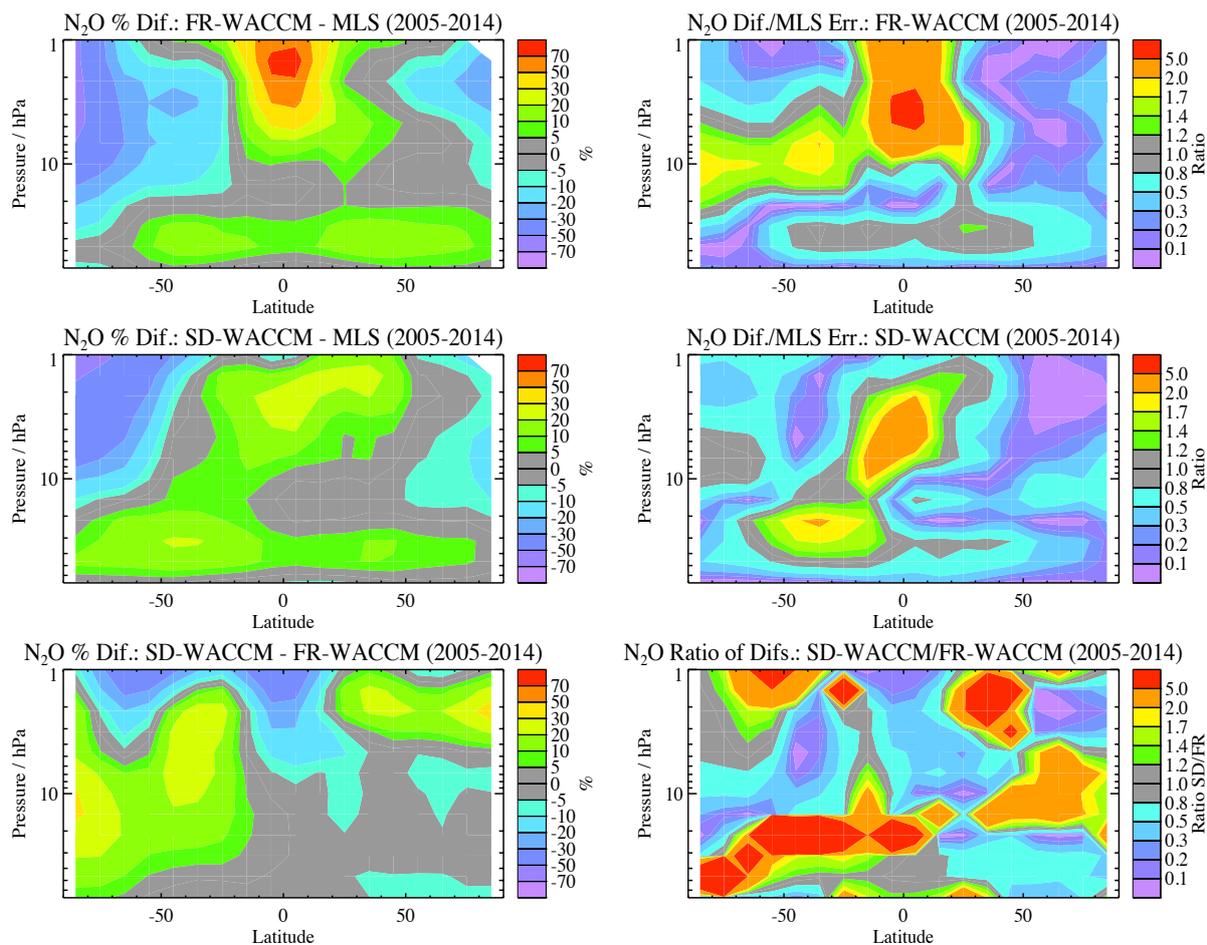
Figure 4. Same as Fig. 1, but for stratospheric HCl.

3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20



1

2



3

4

5 **Figure 5.** Same as Fig. 1, but for stratospheric N_2O .

6

7

8

9

10

11

12

13

14

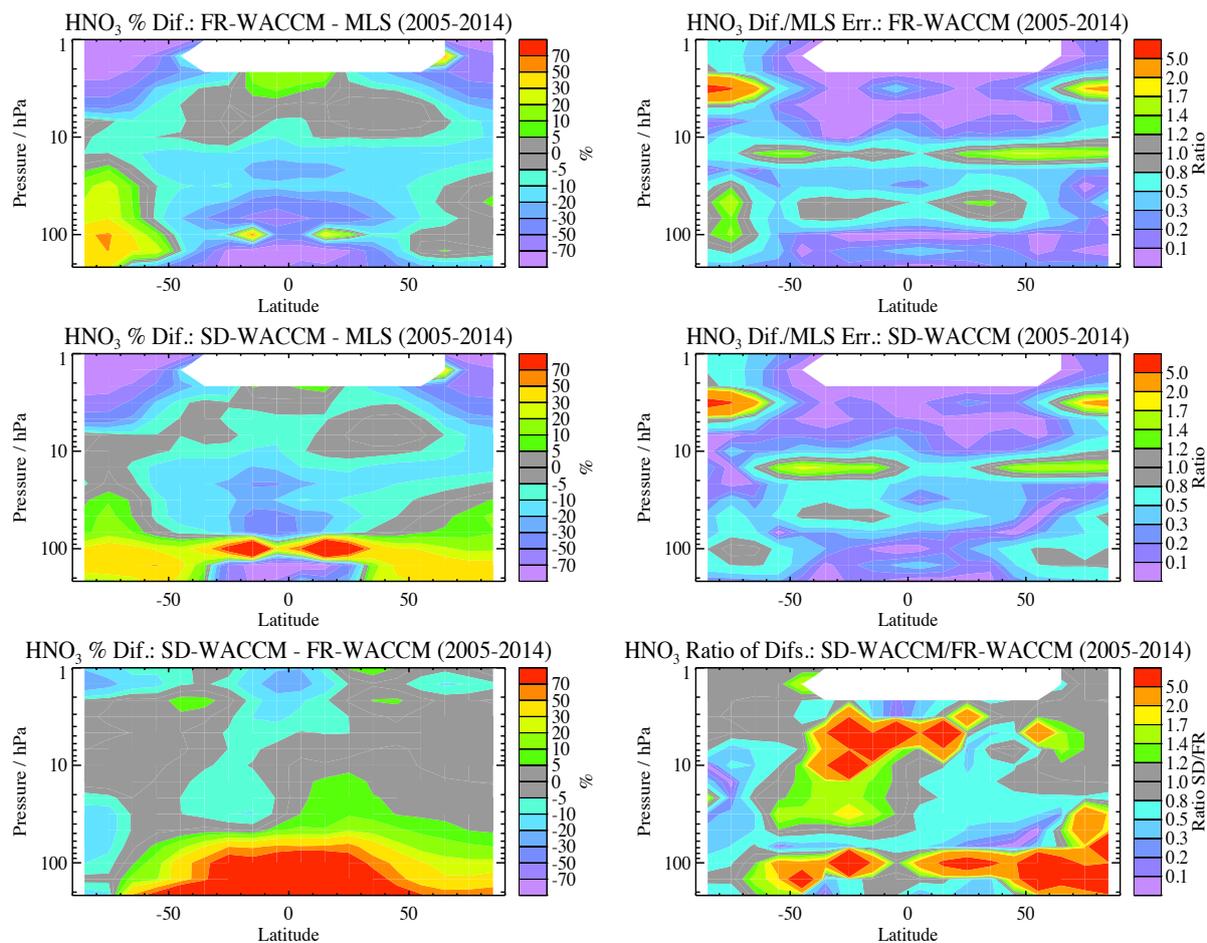
15

16

17



1
2

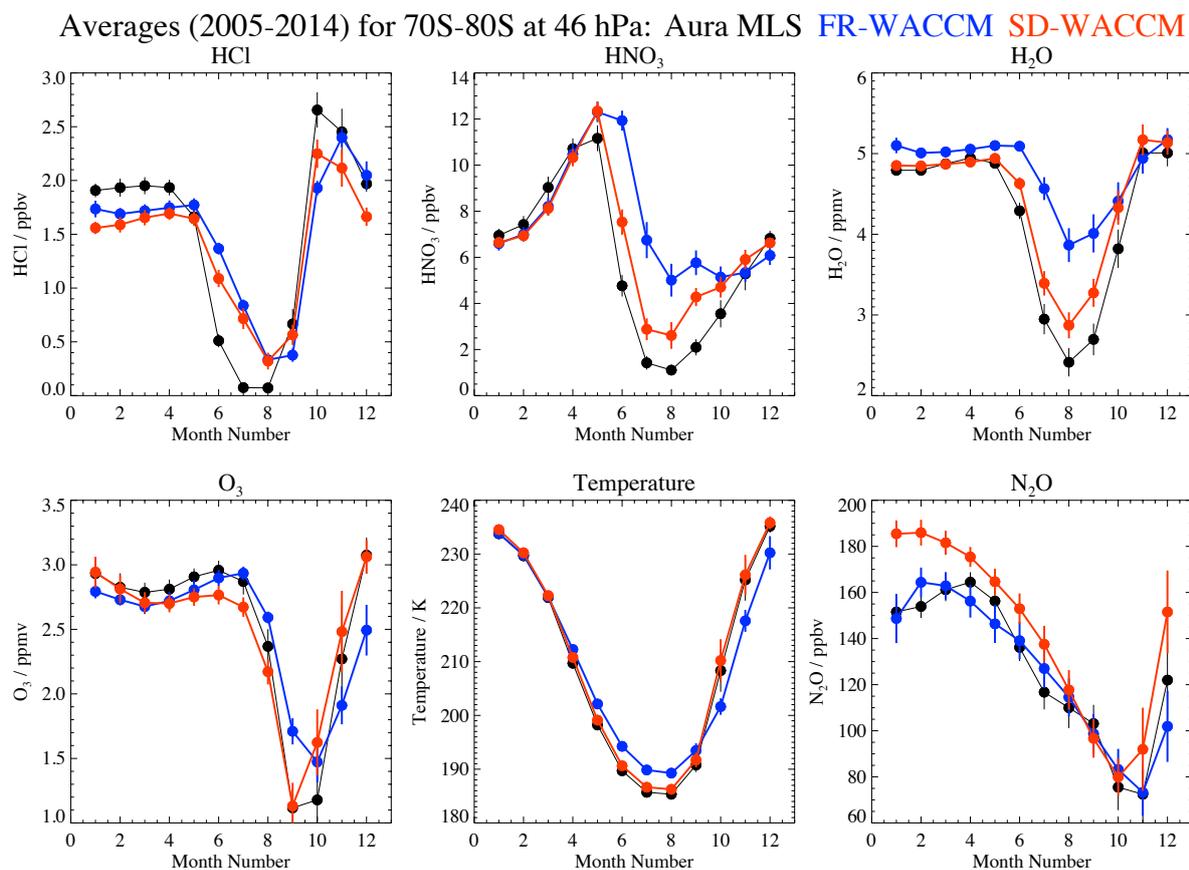


3
4
5
6

Figure 6. Same as Fig. 1, but for stratospheric HNO₃.



1
2

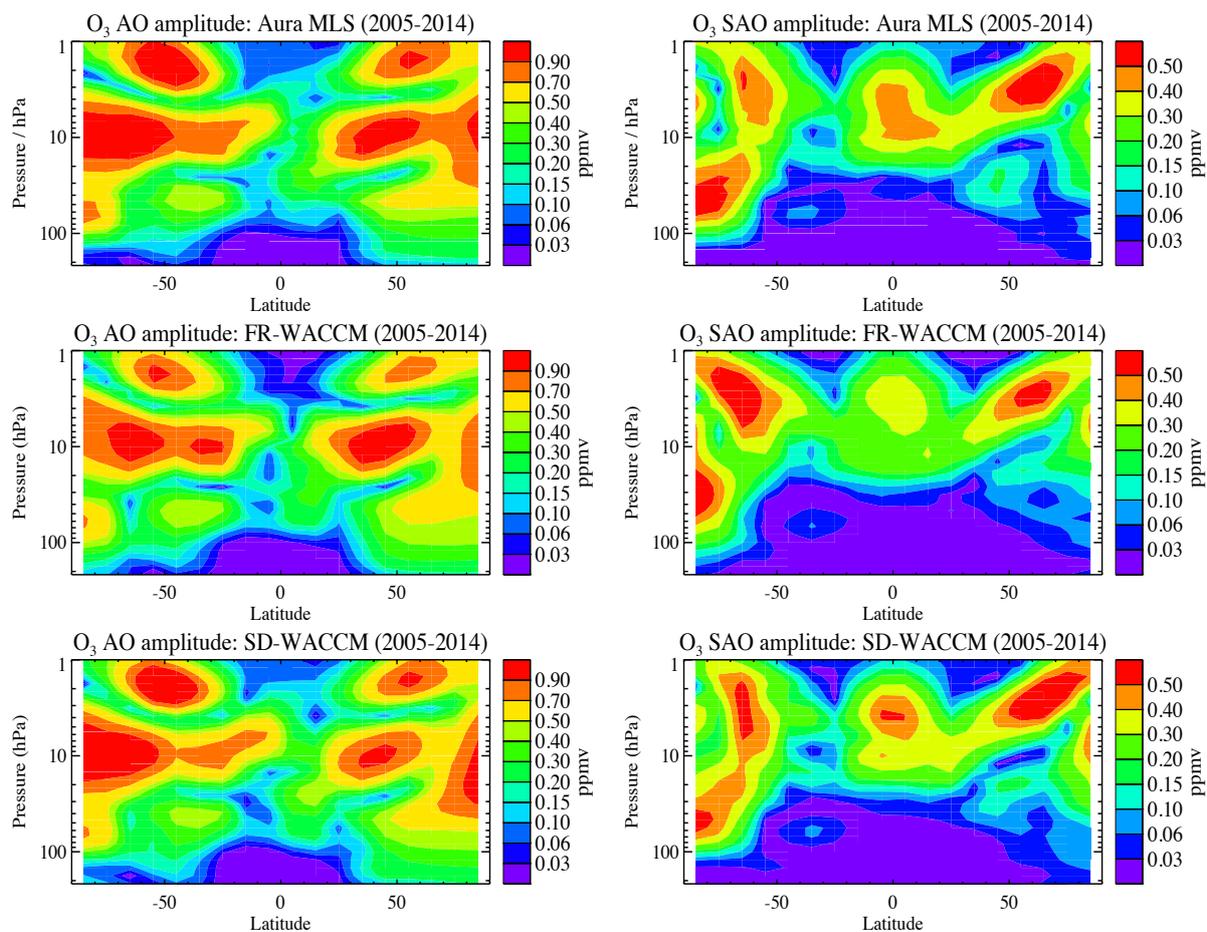


3
4
5
6
7
8
9

Figure 7. Each of the panels shows average seasonal changes from 2005-2014 for the 70°S-80°S region at 46 hPa. Data values (black) are from Aura MLS and model comparisons (FR-WACCM in blue, SD-WACCM in red) are provided for HCl (top left), HNO₃ (top center), H₂O (top right), O₃ (bottom left), temperature (bottom center), and N₂O (bottom right). For each month, the error bars represent twice the standard errors in the means, based on the set of 10 monthly averages (from 2005 through 2014).



1
2



3

4

5 **Figure 8.** Amplitude of the stratospheric ozone annual cycle (left panels) and semi-annual cycle (right panels) for Aura MLS
6 (top), FR-WACCM (middle), and SD-WACCM (bottom), based on fits to time series from 2005 through 2014.

7

8

9

10

11

12

13

14

15

16

17

18

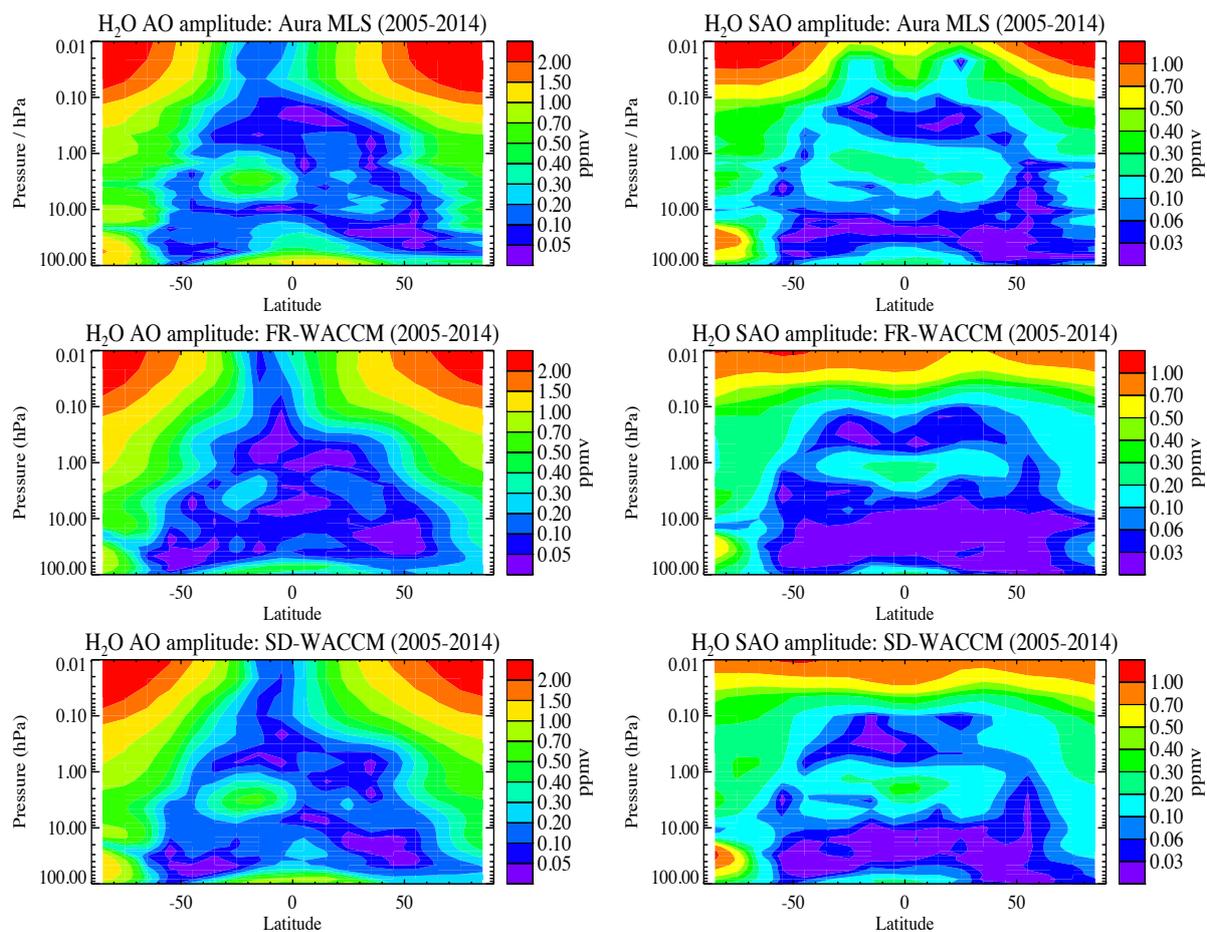
19

20



1

2



3

4

5 **Figure 9.** Same as Fig. 8, but for H₂O annual and semi-annual cycles in the stratosphere and mesosphere.

6

7

8

9

10

11

12

13

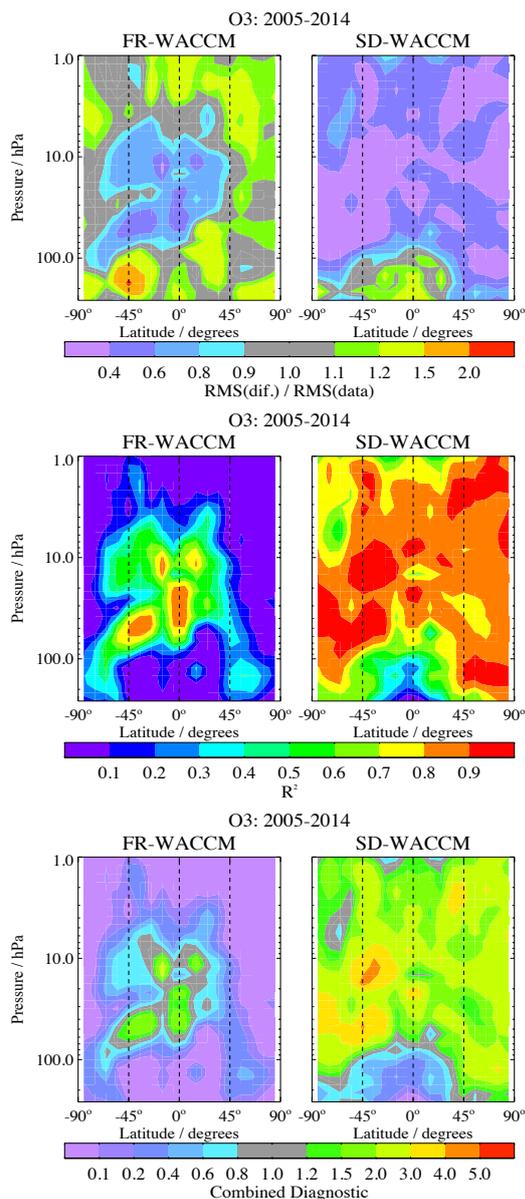
14

15



1

2



3

4

Figure 10. Latitude/pressure contours of diagnostics that show how well the deseasonalized anomalies of model ozone time series (FR-WACCM at left, SD-WACCM, at right) compare to MLS O₃ anomaly series for 2005-2014. Top panels show the RMS difference diagnostic (see text) and middle panels show R² values; small RMS difference values represent a closer fit, while large R² values represent highly correlated results. The bottom panels provide a combined diagnostic, namely the ratio of R² to the RMS difference diagnostic from the top panels; larger values here represent a better result for comparisons to the observed time series.

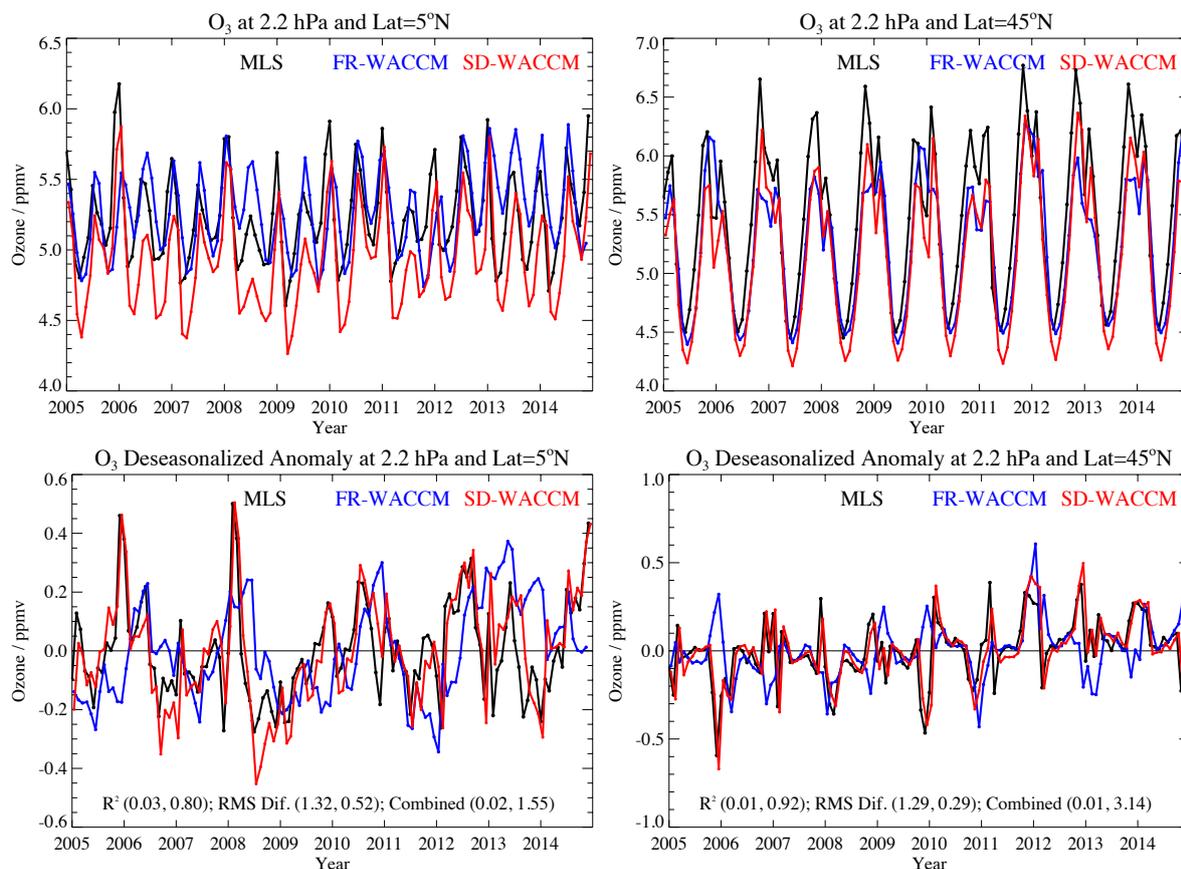
11

12

13



1
2

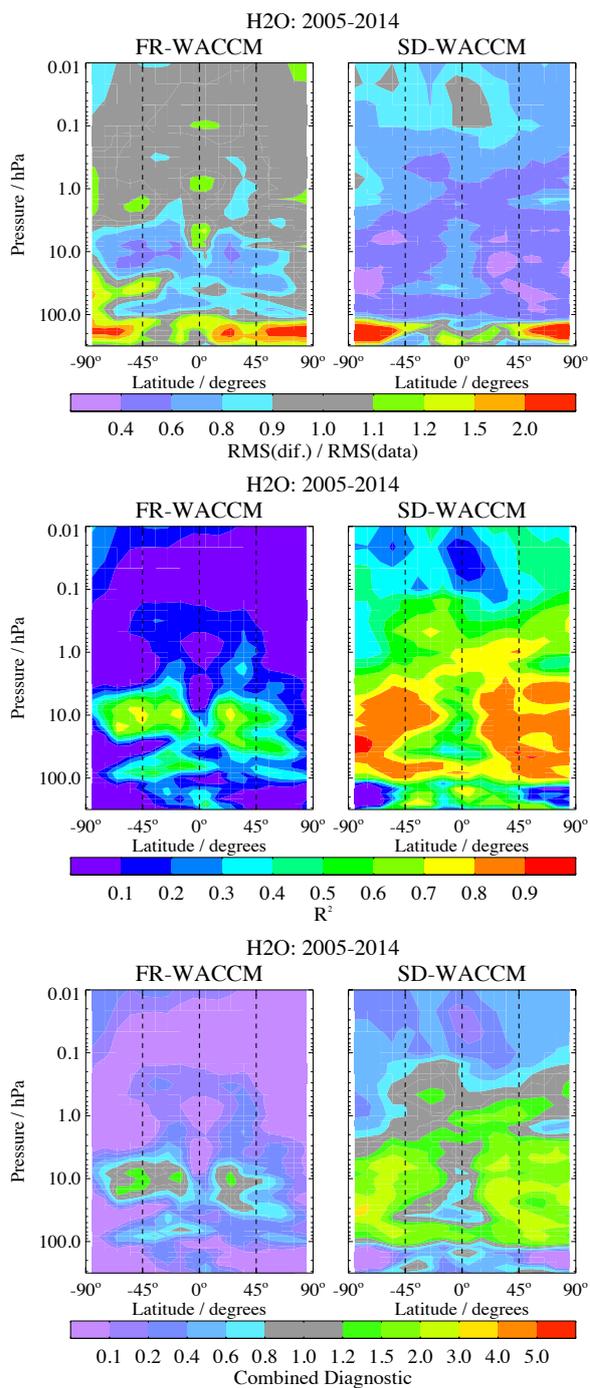


3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24

Figure 11. Time series of monthly zonal mean O₃ mixing ratios at 2.2 hPa (top panels) and deseasonalized anomalies (bottom panels), with the 0-10°N and 40°N-50°N latitude bins on the left and right, respectively. The two model time series (FR-WACCM in blue and SD-WACCM in red) are compared to the MLS series (in black) for 2005-2014. Diagnostic values (see text for a description) are shown in parentheses in the bottom two panels, with the 1st number referring to FR-WACCM and the 2nd number to SD-WACCM.



1
2

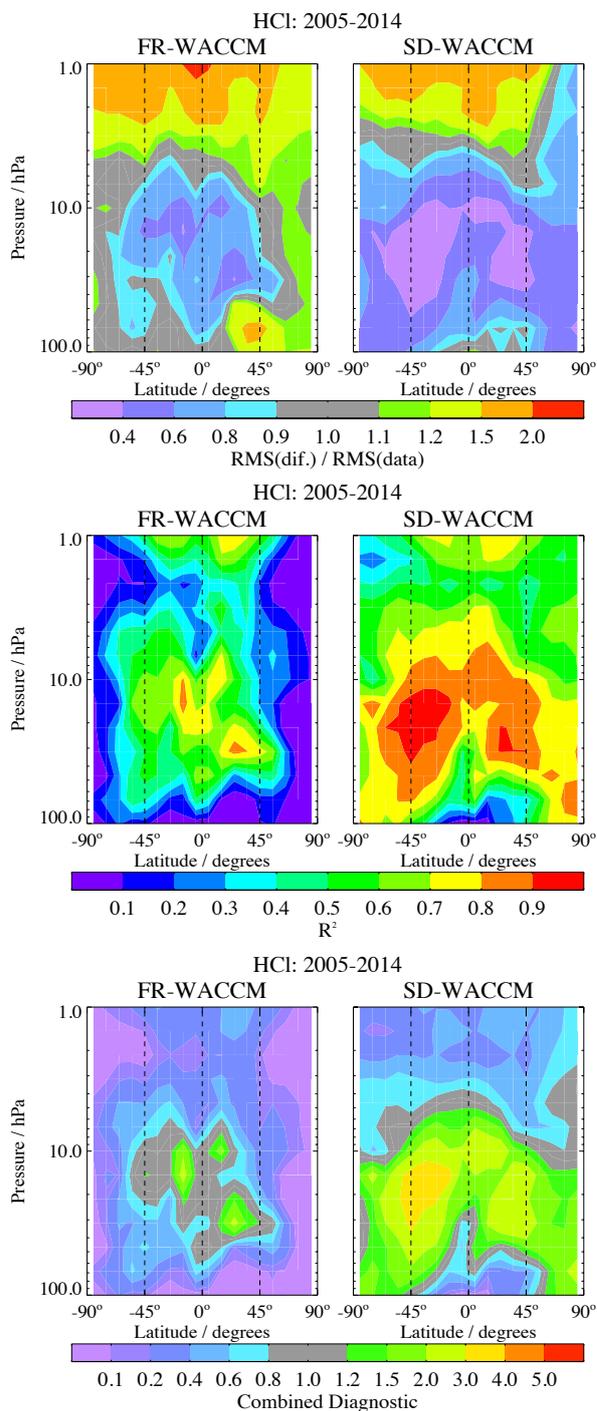


3
4
5
6
7

Figure 12. Same as the Fig. 10 diagnostics, but for H₂O up to 0.01 hPa.



1
2

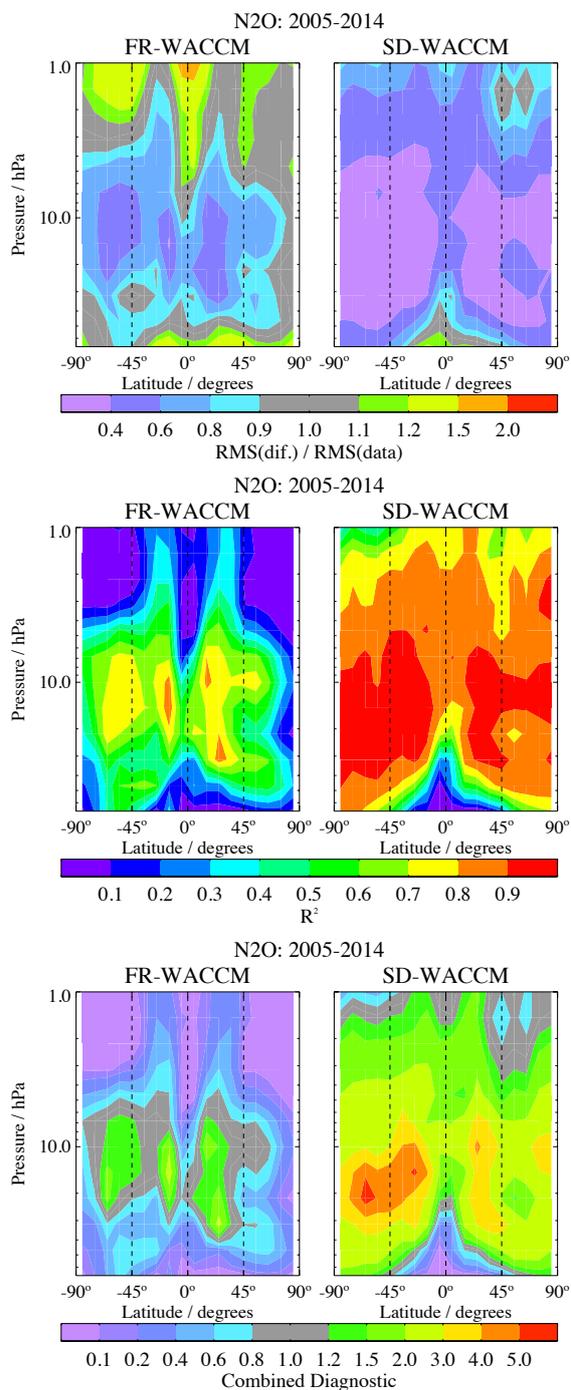


3
4
5
6
7

Figure 13. Same as the Fig. 10 diagnostics, but for stratospheric HCl.



1
2

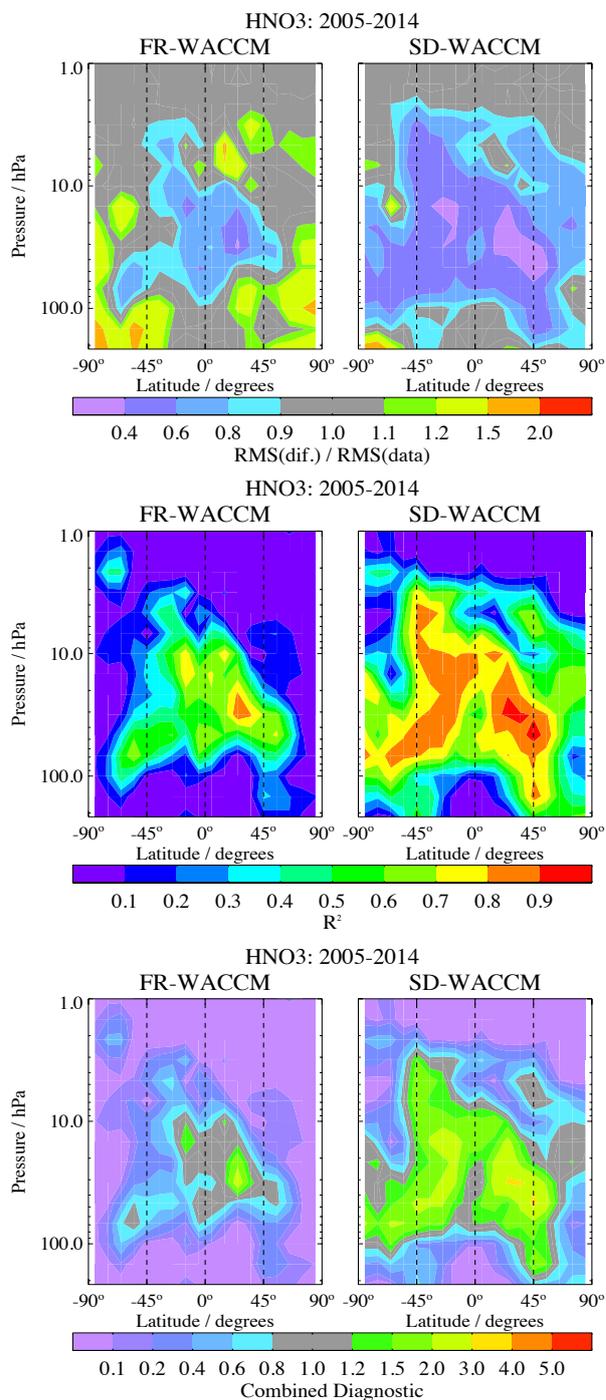


3
4
5
6
7

Figure 14. Same as the Fig. 10 diagnostics, but for N₂O (from 68 to 1 hPa).



1
 2

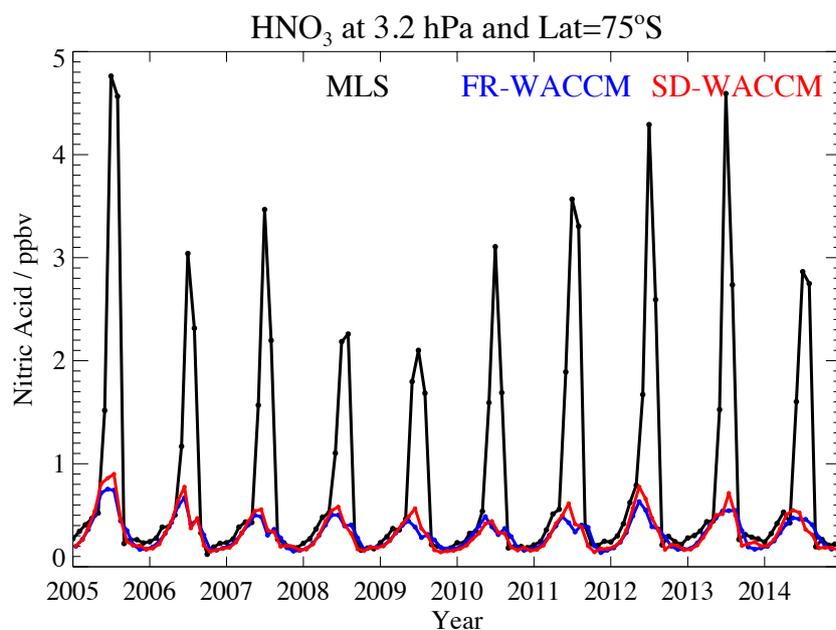


3
 4
 5
 6
 7

Figure 15. Same as the Fig. 10 diagnostics, but for HNO₃ (from 215 to 1 hPa).



1
2

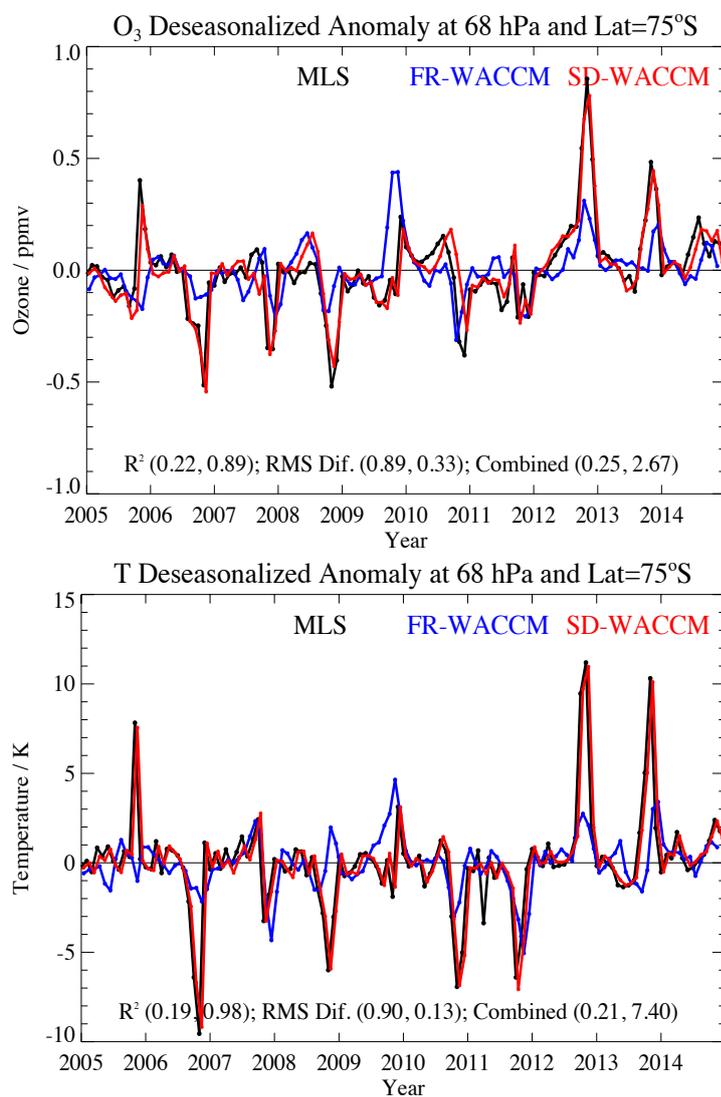


3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31

Figure 16. HNO₃ monthly zonal mean mixing ratio time series (2005 through 2014) from MLS, FR-WACCM, and SD-WACCM for 3.2 hPa and 70°S-80°S.



1
2

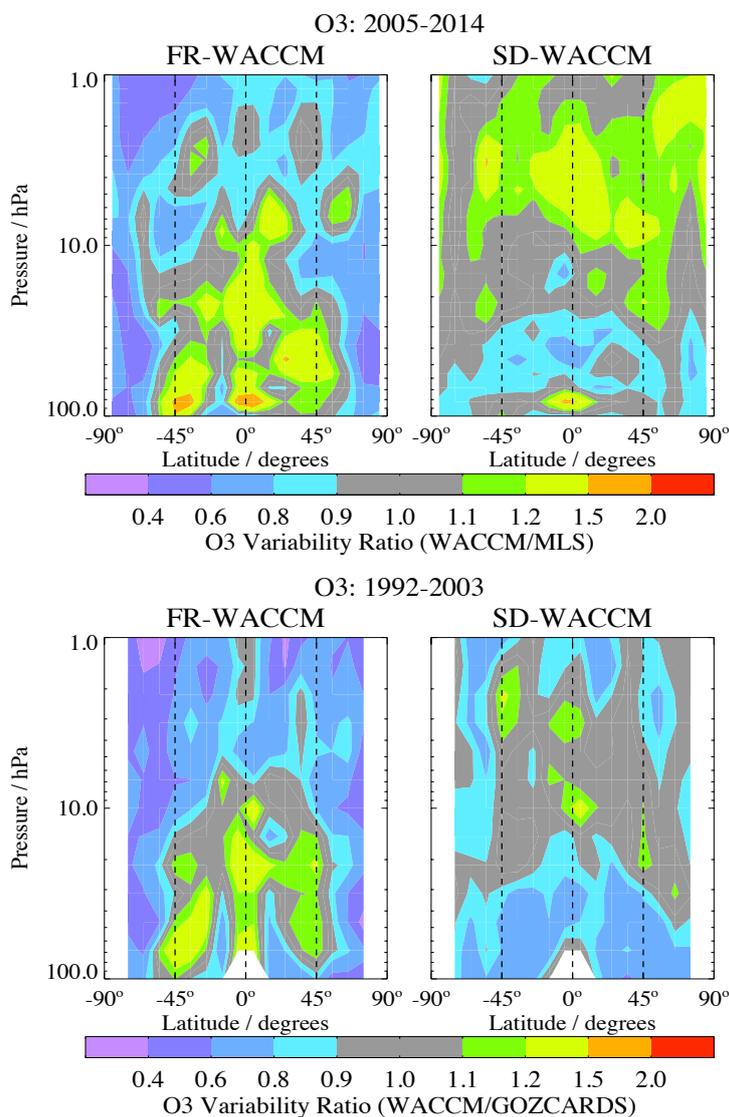


3
4
5
6
7
8
9
10
11
12
13
14
15
16
17

Figure 17. Deseasonalized anomaly time series (2005-2014) at 68 hPa and 70°S-80°S from Aura MLS, FR-WACCM, and SD-WACCM for ozone (top panel) and temperature (bottom panel); calculated values are provided in parentheses for the R², RMS fit, and combined diagnostics used in this work (the 1st number refers to FR-WACCM and the 2nd number to SD-WACCM).



1
2

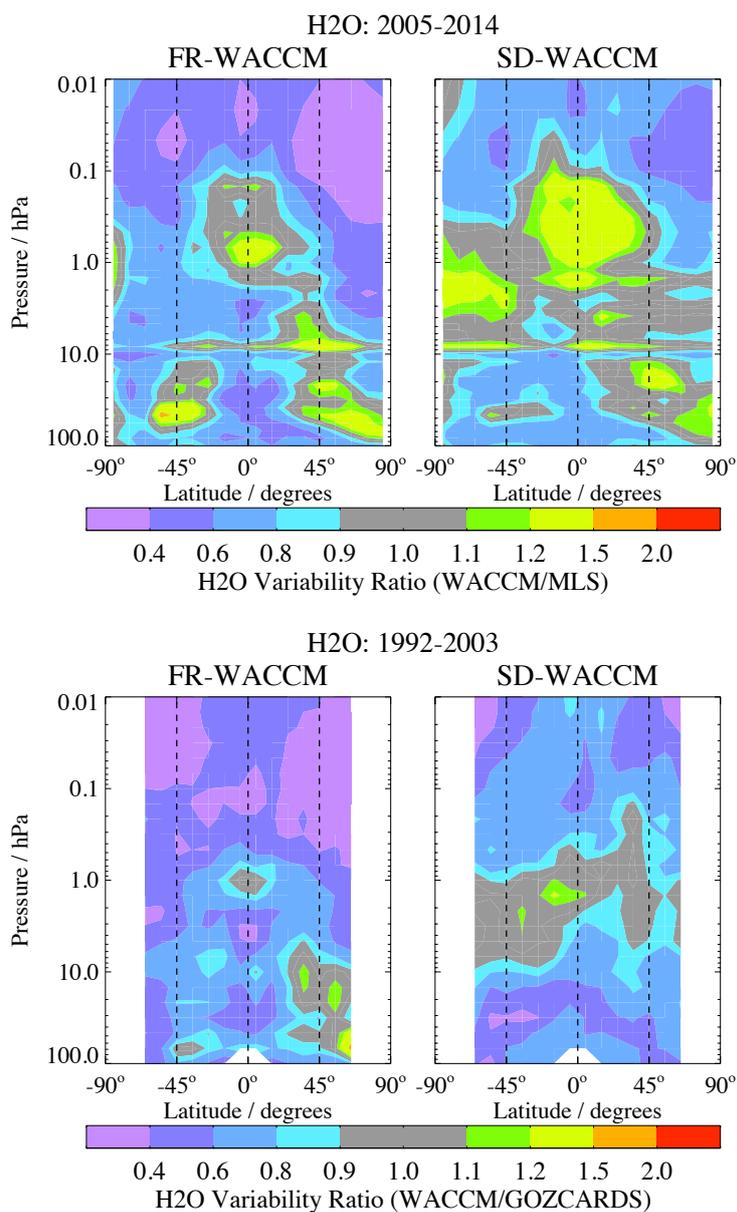


3
4
5
6
7
8
9
10
11
12
13
14
15
16

Figure 18. Variability ratios (model results divided by data results) for stratospheric O₃, with FR-WACCM results on the left, and SD-WACCM on the right. Before calculating the ratios, the variability values are obtained as the root mean square of detrended deseasonalized monthly anomaly time series, and expressed as a percentage of mean (climatological) abundances; the top panels show comparisons to MLS data for 2005-2014, whereas the bottom panels are for 1992-2003 comparisons to GOZCARDS.



1
2

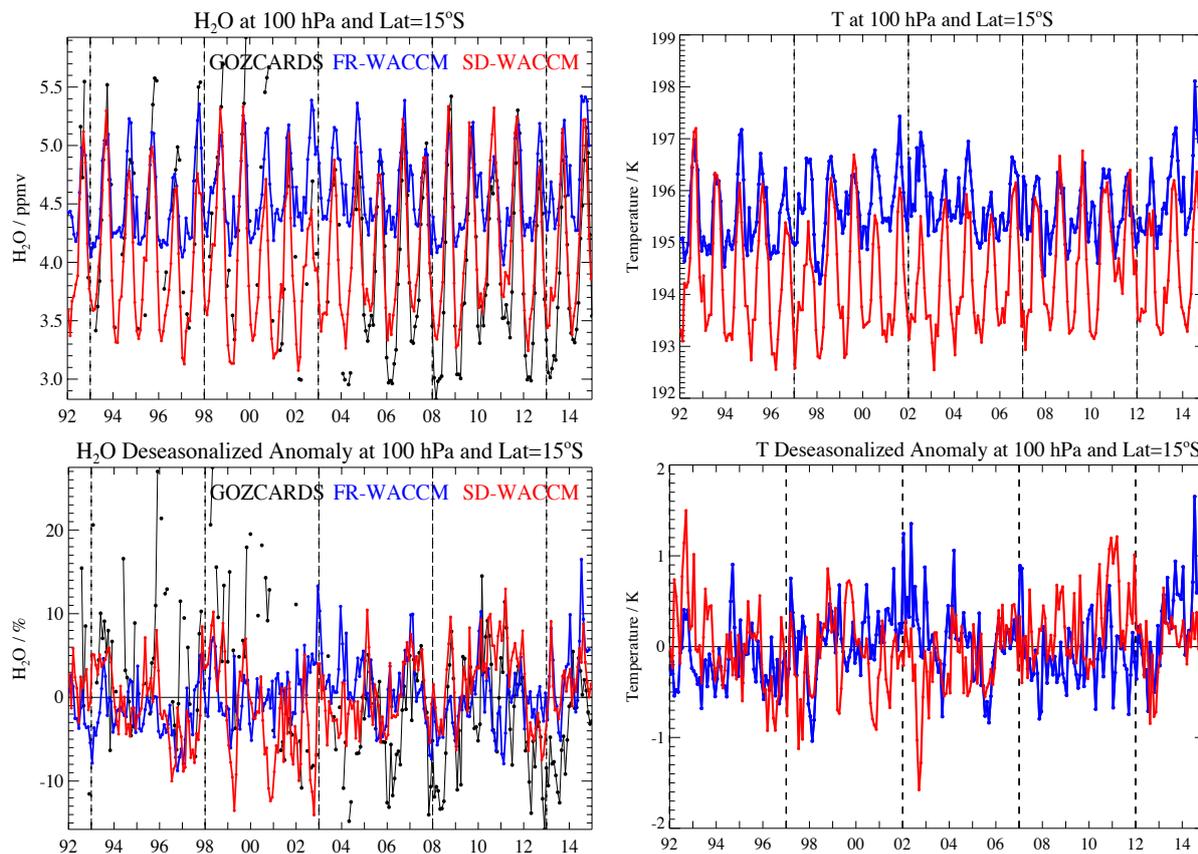


3
4
5
6
7

Figure 19. Same as Fig. 18, but for ratios (model/data) of H₂O stratospheric and mesospheric variability for two different time periods.



1
2

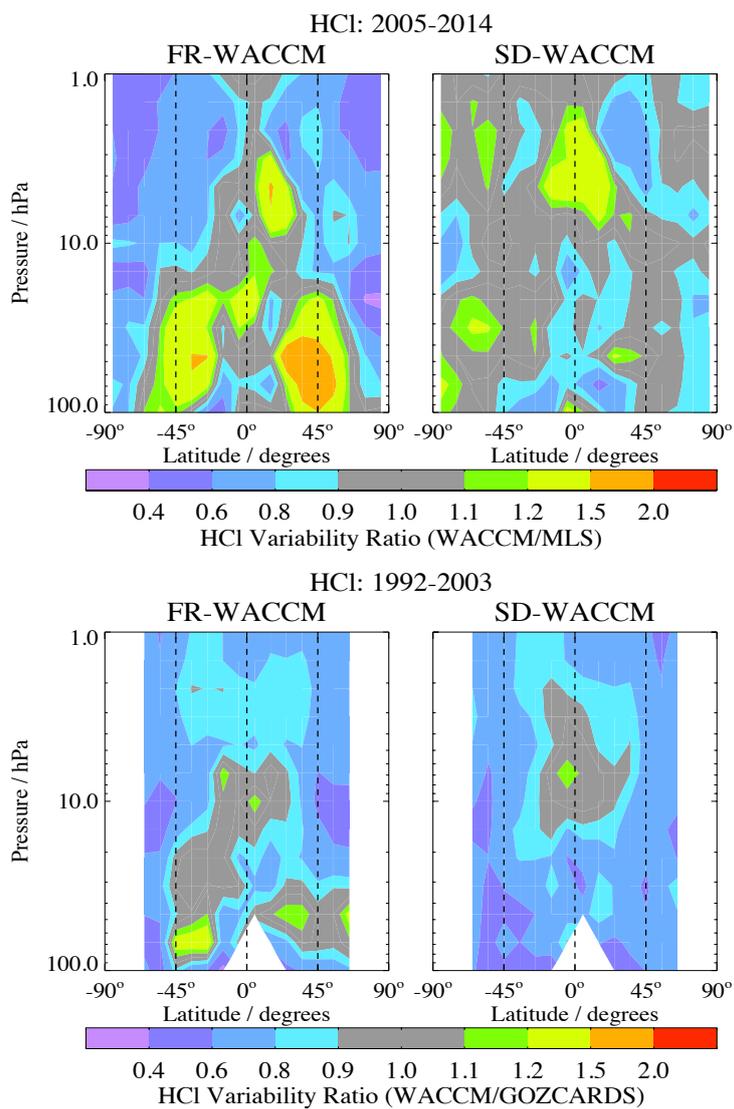


3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22

Figure 20. Time series (1992–2014) at 100 hPa and 10°S–20°S for temperature (right two panels) and H₂O (left two panels), with deseasonalized anomalies shown in the bottom two panels. The temperature plots just show the two models (FR-WACCM in blue, SD-WACCM in red), whereas the H₂O series show the comparisons for the models versus GOZCARDS merged H₂O data (in black).



1
2

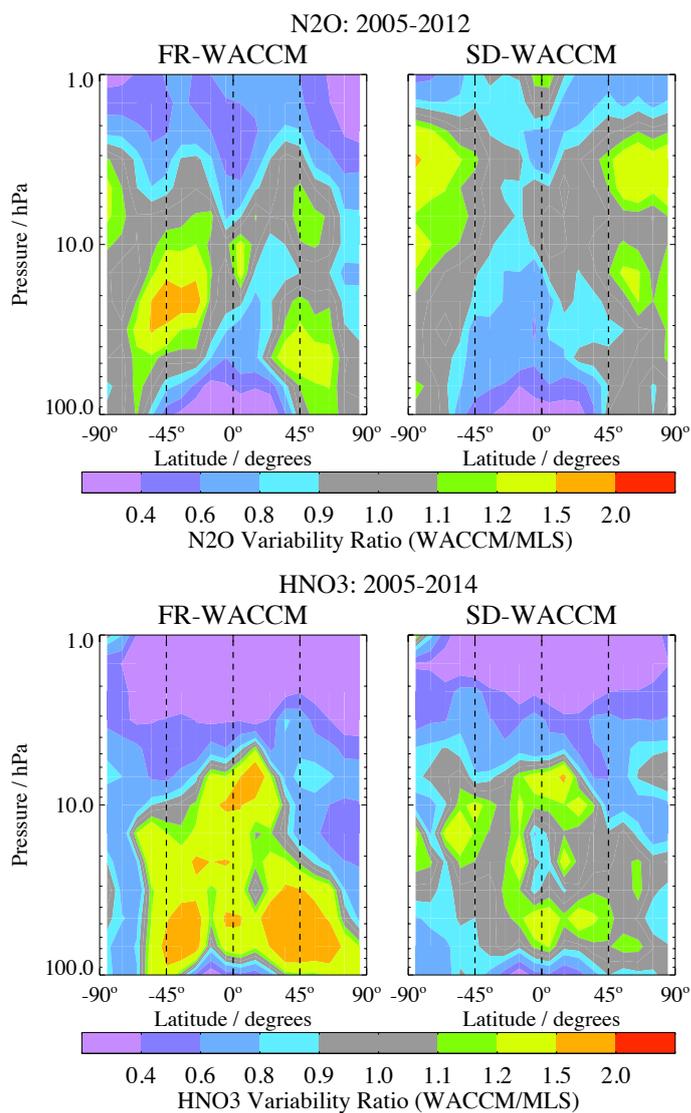


3
4
5
6
7
8
9
10
11
12
13
14
15

Figure 21. Same as Fig. 18, but for ratios (model/data) of HCl stratospheric variability.



1
2

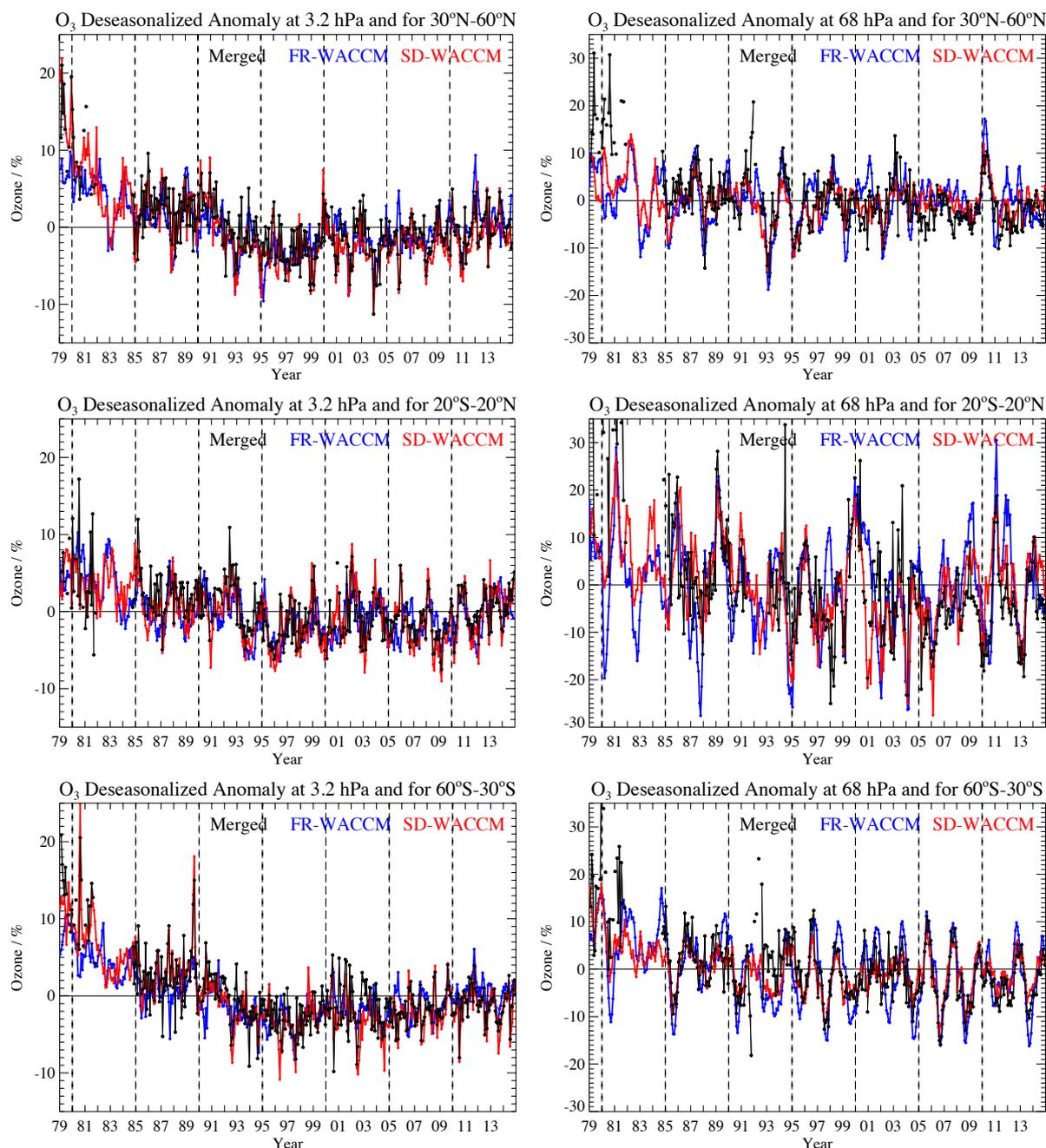


3
4
5
6
7
8
9
10
11
12
13
14
15
16

Figure 22. Same as Fig. 18, but for ratios (model/data) of the variability in stratospheric N₂O for 2005-2012 (top panels) and HNO₃ for 2005-2014 (bottom panels). The MLS N₂O-640 product (see text) is used for the comparisons in the top panels.



1
2

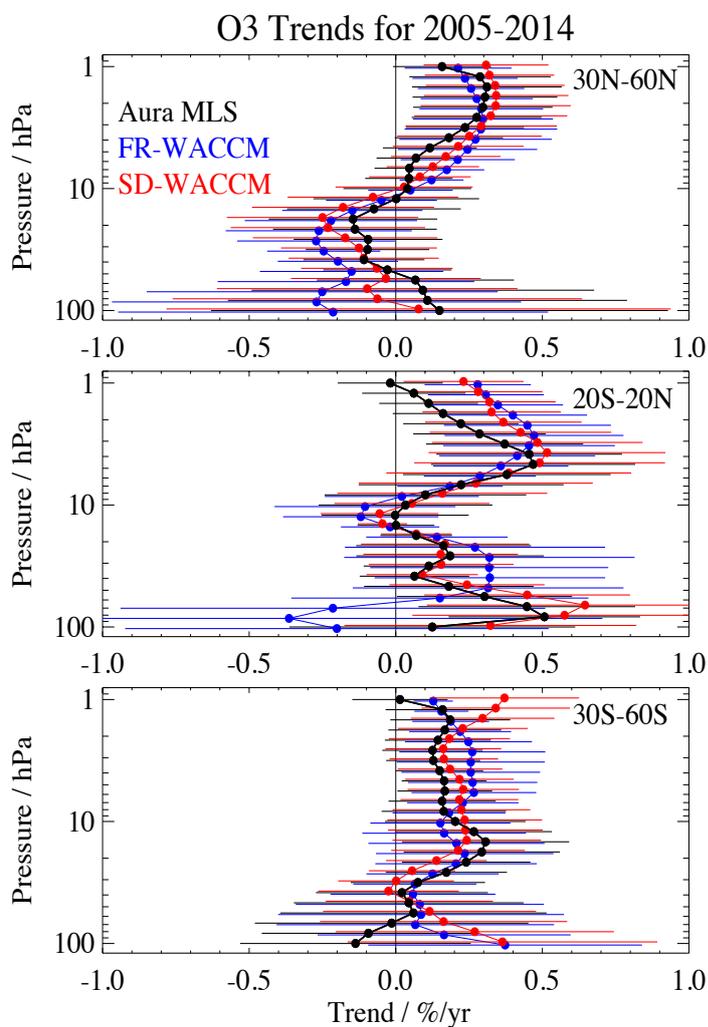


3
4
5
6
7
8
9
10

Figure 23. Sample time series of deseasonalized ozone anomalies (%) from 1979 through 2014 from the GOZCARDS data record (version 2.20) compared to the corresponding model anomalies from FR-WACCM (blue) and SD-WACCM (red). Upper stratospheric series at 3.2 hPa are shown in left panels and lower stratospheric series at 68 hPa are on the right; three latitude bins are displayed (30°N-60°N, top; 20°S-20°N, middle, and 30°S-60°S, bottom).



1
2



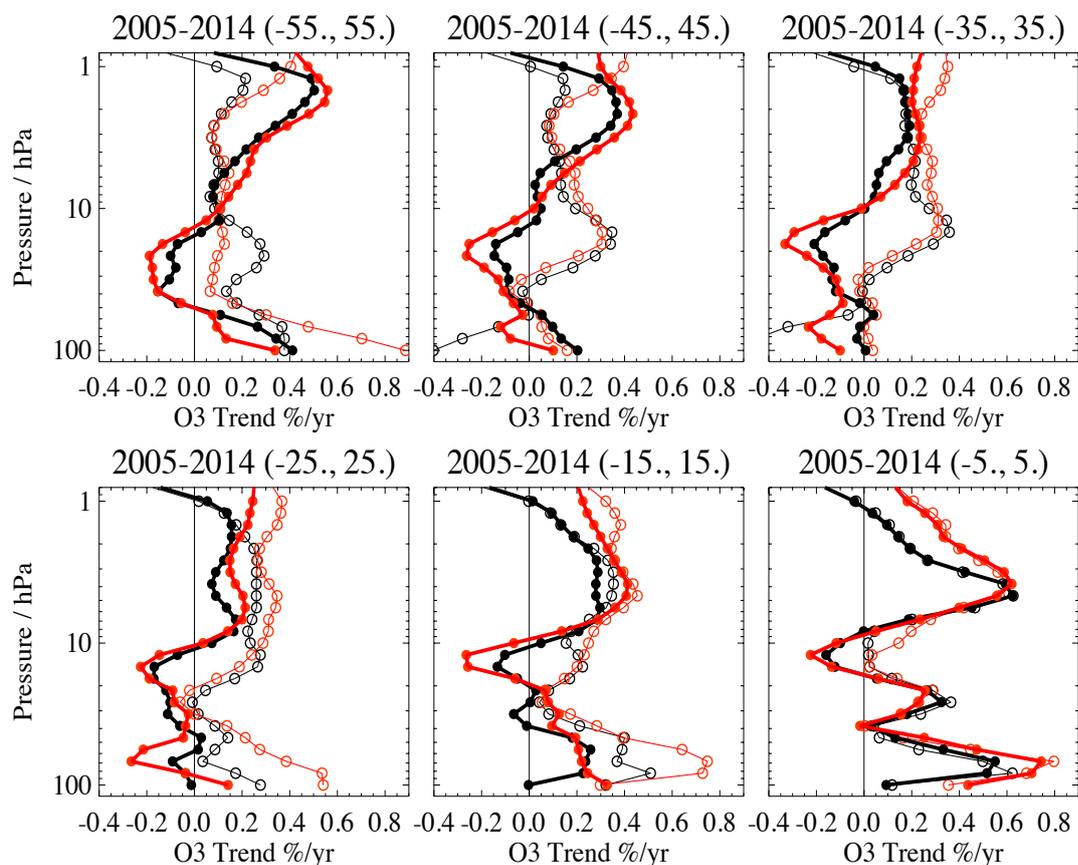
3
4
5
6
7
8
9

Figure 24. Ozone stratospheric trends for 2005 through 2014 obtained from monthly zonal mean data (version 4.2 Aura MLS) and models (FR-WACCM and SD-WACCM), after multiple linear regression analyses of deseasonalized anomaly time series, as described in the text. Each panel refers to results from different latitude band average series (see legend). The error bars are 2σ estimates based on bootstrap resampling results (see text).



1

2



3

4

5

6

7

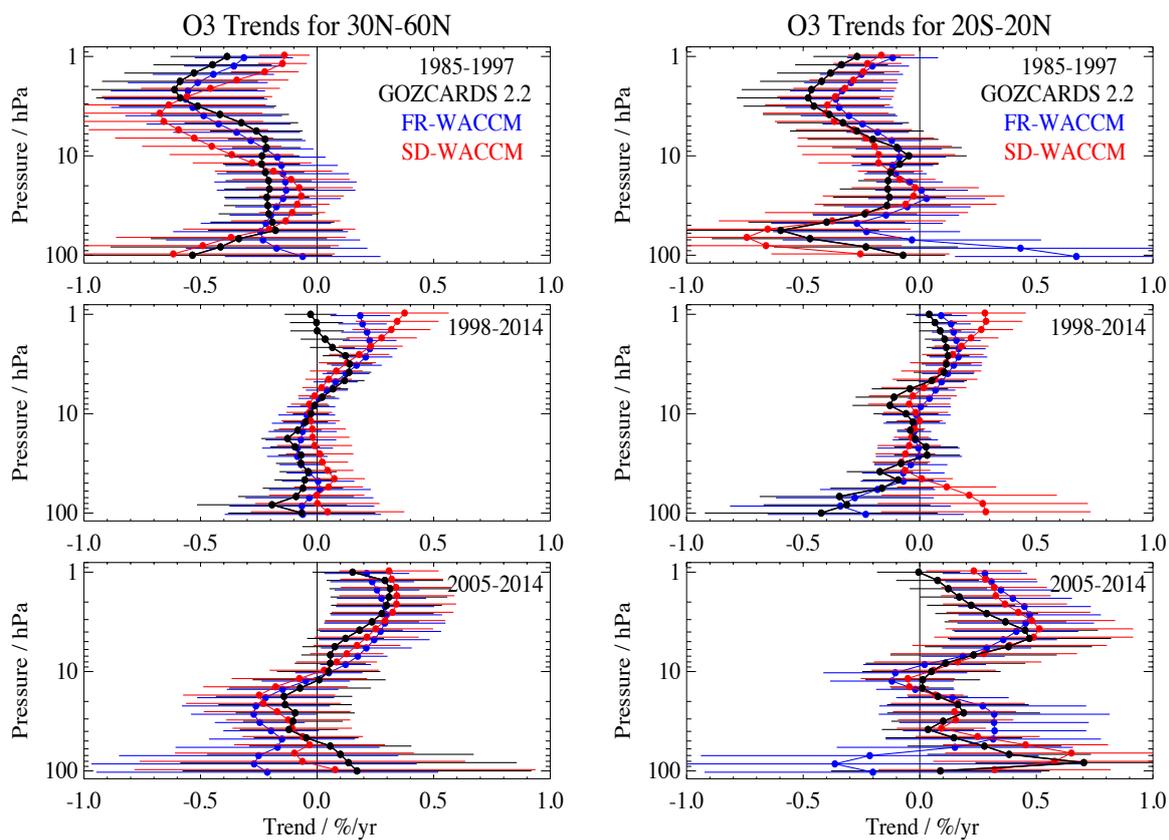
8

9

Figure 25. Ozone trends in different latitude bins for SD-WACCM (red) versus MLS data (black) for 2005-2014. Closed and open circles are for northern and southern latitude bins, respectively. For clarity, error bars are omitted here, as these generally show that model/data trend differences are not significant for this time period.



1
2

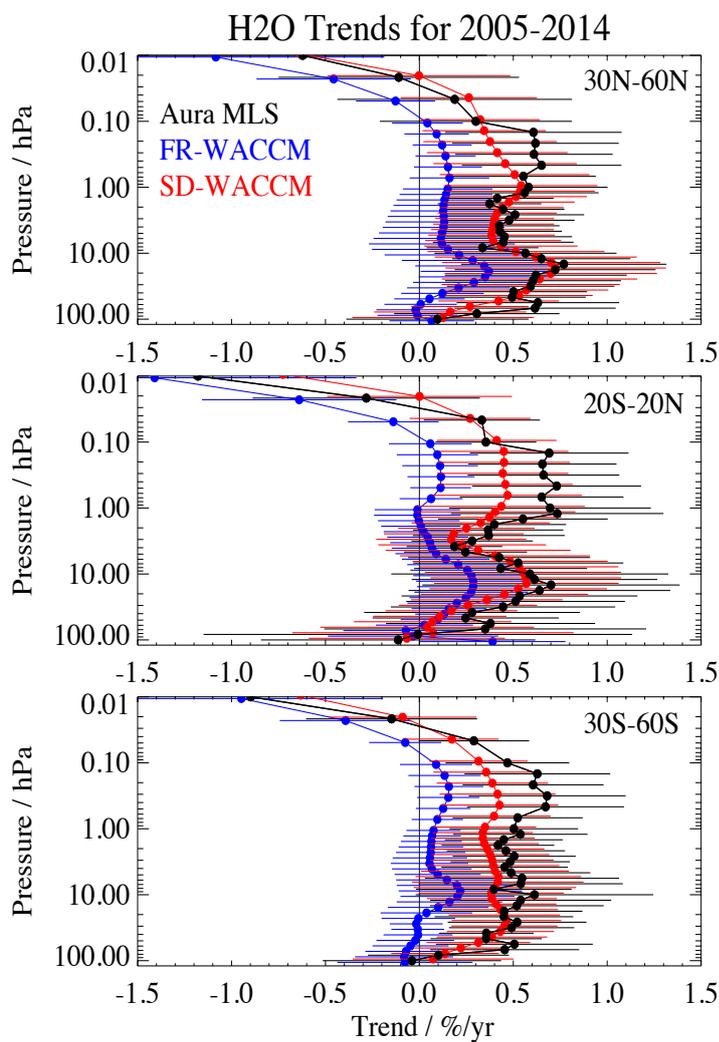


3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22

Figure 26. Same as Fig. 24 for ozone trend comparisons, except for the use of two latitude bands (30°N-60°N on left side and 20°S-20°N on right side) and three different time periods (from top to bottom panels, see legend); the data record here is from GOZCARDS merged ozone version 2.20 (see text).



1
2

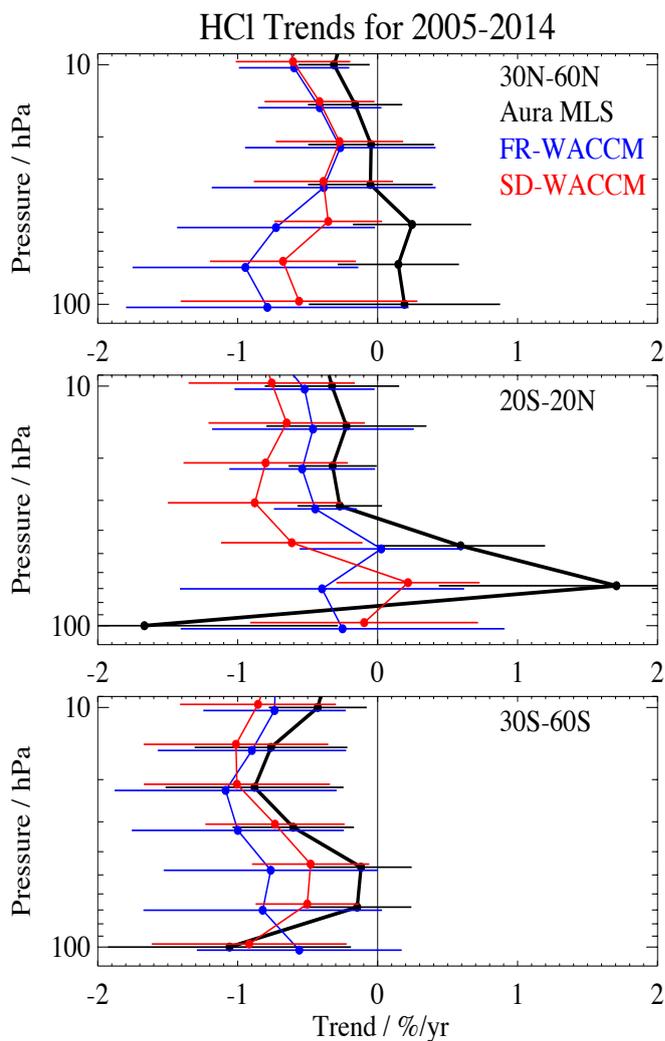


3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18

Figure 27. Trends in three latitude bins for stratospheric and mesospheric H₂O from an analysis of the 2005-2014 MLS data and the two WACCM models over the same time period.



1
2



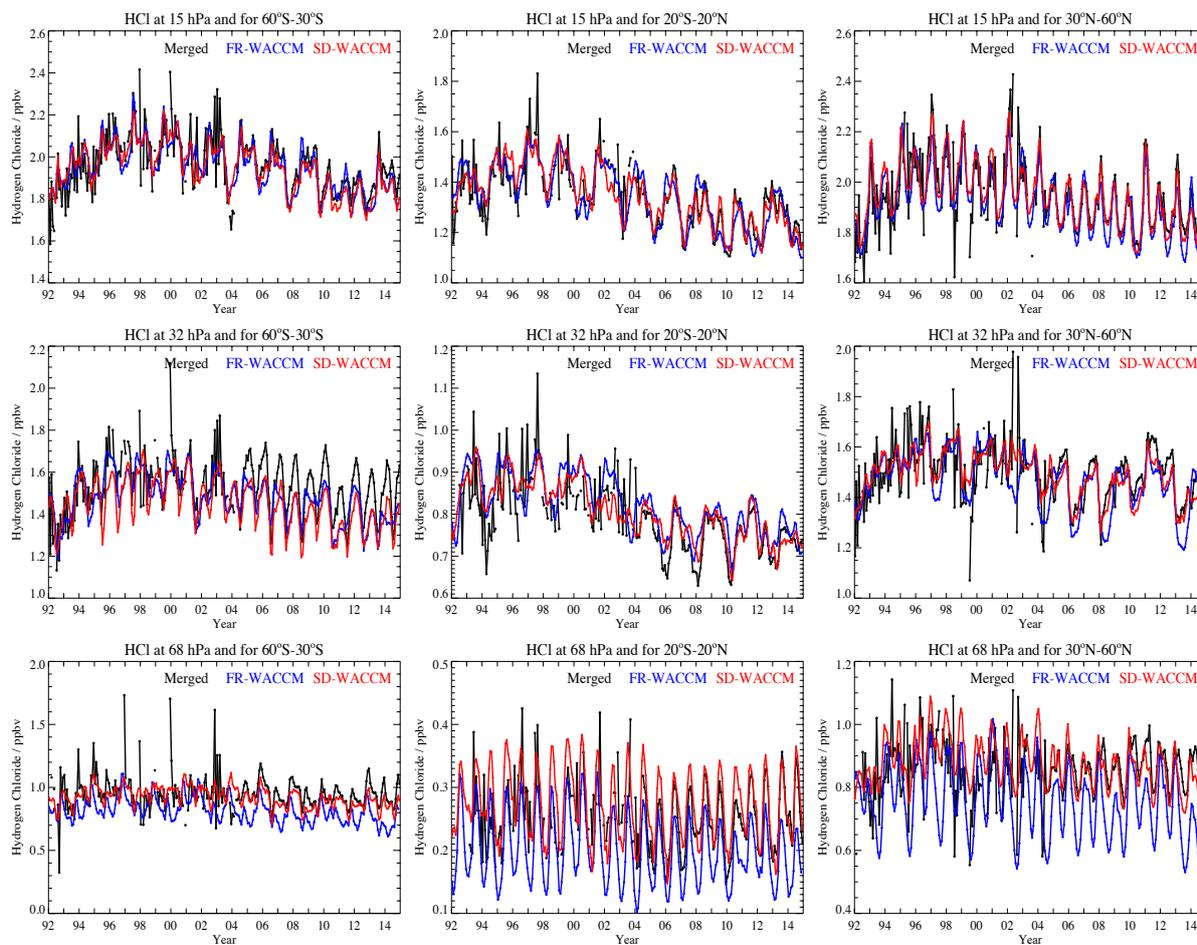
3
4
5
6
7
8
9
10
11
12
13
14
15
16

Figure 28. Same as Fig. 24, but for HCl data and model trends in the lower stratosphere.



1

2



3

4

Figure 29. Time series (1992-2014) of lower stratospheric HCl (ppbv) for the GOZCARDS HCl merged data record (black), as well as models (FR-WACCM in blue and SD-WACCM in red). Each panel is for a different pressure level and latitude bin, as labeled (15 hPa, top; 32 hPa, middle; 68 hPa, bottom); the three latitude bins used in this work are 30°S-60°S (left panels), 20°S-20°N (middle panels), and 30°N-60°N (right panels).

9

10

11

12

13

14

15

16

17

18

19

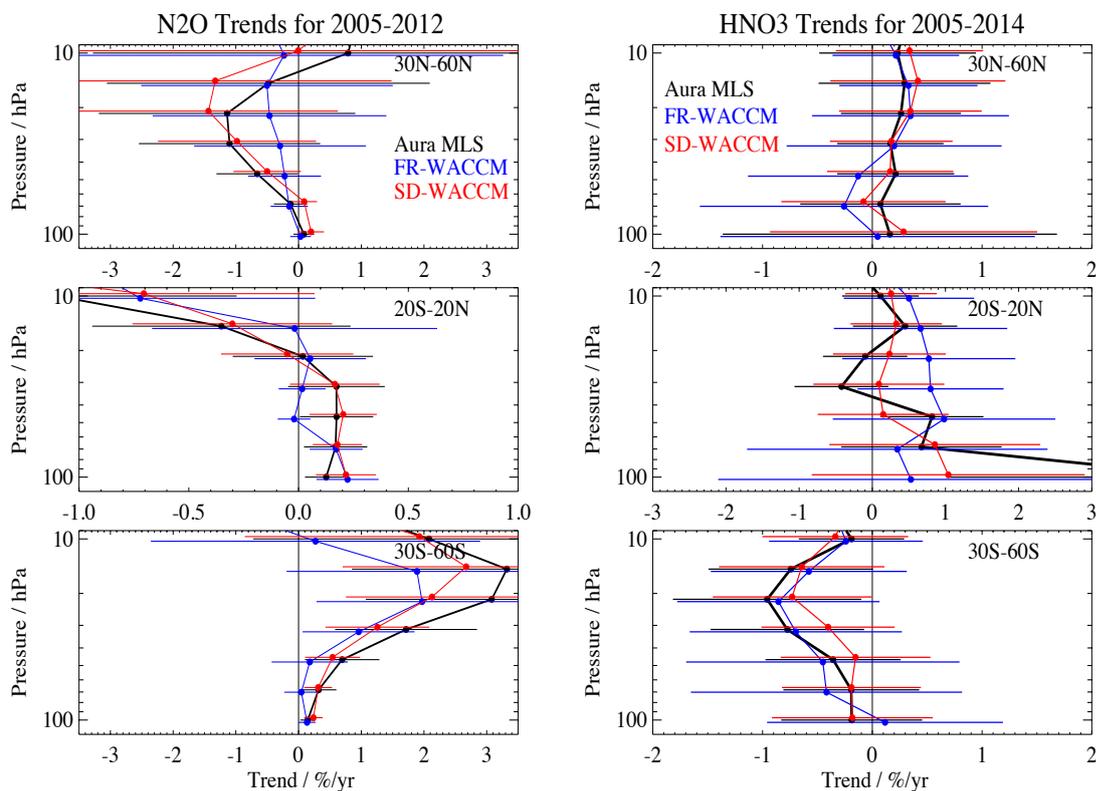
20

21

22



1
2



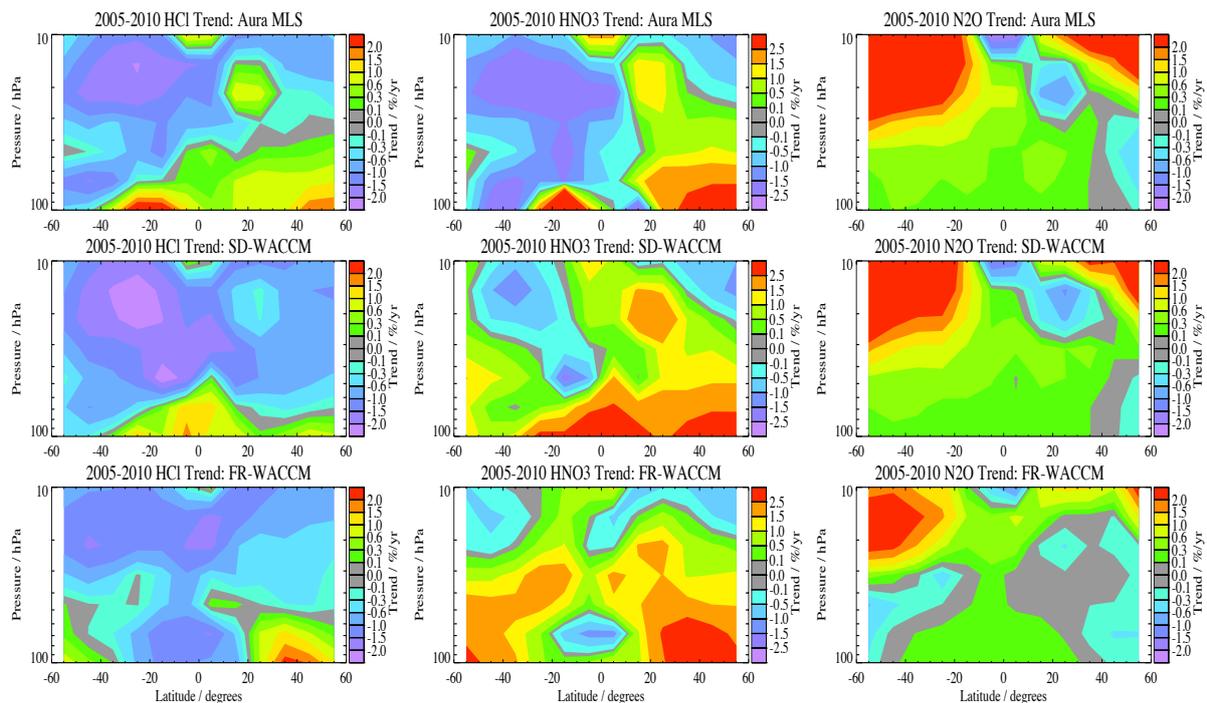
3
4
5
6
7
8
9

Figure 30. Same as Fig. 28, but for N₂O (left 3 panels) and HNO₃ (right 3 panels) data and model trends in the lower stratosphere. The N₂O data results are from the N₂O-640 MLS product (retrieved from the 640 GHz radiometer band data), which was discontinued in 2013 because of an instrument issue affecting this band (see text), and these data and model trends apply to the 2005-2012 period. The HNO₃ trend results (data and models) are for 2005-2014.



1

2



3

4

5

6

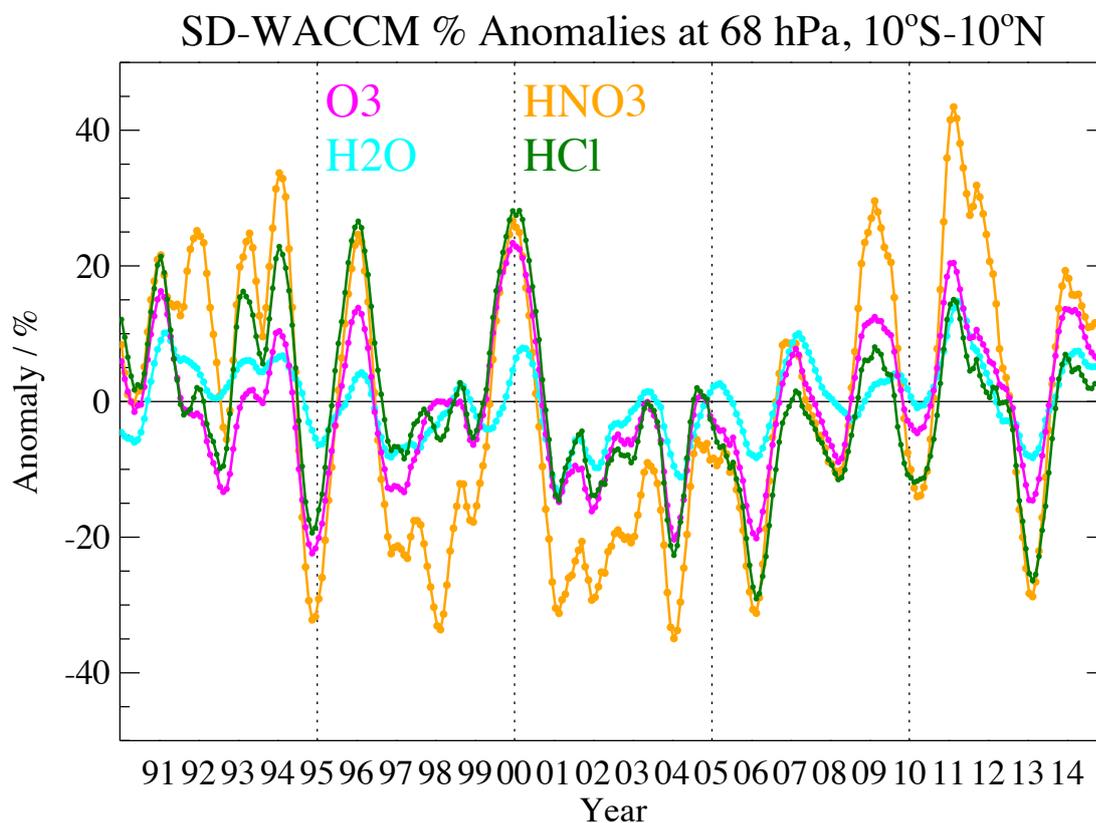
7

8

Figure 31. Pressure/latitude contour comparison of short-term (2005-2010) lower stratospheric trends from Aura MLS data (top panels) versus model trends (from SD-WACCM in middle panels and FR-WACCM in bottom panels). The left column is for HCl, the middle column for HNO₃, and the right column is for N₂O (using data from the MLS 640 GHz radiometer band).



1
2



3
4
5
6
7
8
9
10
11
12

Figure 32. 25-year simulation from SD-WACCM for 1990 through 2014, showing monthly mean deseasonalized anomalies (%) for O₃, H₂O, HNO₃, and HCl at 68 hPa for 10°S-10°N, in a region where dynamically-induced variations, both short- and long-term, appear to affect these species in similar ways. Some smoothing (using a running mean with a ±7-month window) has been applied to highlight the variations longer than a year.