

1 **Uncertainties in recent tropical stratospheric and tropospheric ozone**
2 **changes limit the use of observational constraints for assessing future**
3 **ozone change**

Deleted: restrict our understanding of

Deleted: total column

4 Sean M. Davis¹, William T. Ball^{2,3,†}, Yue Jia⁴, Gabriel Chiodo⁵, Justin A. Alsing⁶, James H. Keeble⁷,
5 Hideharu Akiyoshi⁸, Carlo Arosio⁹, Ewa M. Bednarz^{10,1}, Andreas Chrysanthou^{11,12,13}, Melanie Coldewey-
6 Egbers¹⁴, Robert Damadeo¹⁵, Sandip Dhomse^{11,16}, Mohamadou A. Diallo¹⁷, Simone Dietmuller¹⁸, Roland
7 Eichinger^{18,19}, Stacey Frith^{20,21}, Birgit Hassler¹⁸, Michaela I. Hegglin^{22,23,24}, Daan Hubert²⁵, Patrick
8 Jöckel¹⁸, Béatrice Josse²⁶, Natalya Kramarova²¹, Diego Loyola¹⁴, Eliane Maillard Barras²⁷, Marion
9 Marchand²⁸, Olaf Morgenstern^{29,30,31}, David Plummer³², Robert W. Portmann¹, Karen H. Rosenlof³³,
10 Alexei Rozanov⁹, Viktoria F. Sofieva³⁴, Johannes Staehelin³⁵, Timofei Sukhodolov³⁶, Kleareti Tourpali³⁷,
11 Ronald J. Van der A³, H.J. Ray Wang³⁸, Krzysztof Wargan^{21,39}, Shingo Watanabe^{40,41}, Mark Weber⁹,
12 Jeannette D. Wild⁴², Yousuke Yamashita⁸, Jerry Ziemke^{21,43}

14 ¹NOAA Chemical Sciences Laboratory, Boulder, CO, USA

15 ²Department of Geoscience and Remote Sensing, TU Delft, Delft, The Netherlands

16 ³Royal Netherlands Meteorological Institute (KNMI), De Bilt, The Netherlands

17 ⁴Univ. of Texas at Dallas, Dallas, TX, USA

18 ⁵Instituto de Geociencias, Consejo Superior de Investigaciones Científicas (IGEO-CSIC), Madrid, Spain

19 ⁶Oskar Klein Centre for Cosmoparticle Physics, Department of Physics, Stockholm University, Stockholm, Sweden

20 ⁷Lancaster Environment Centre, Lancaster University, Lancaster, UK

21 ⁸Earth System Division, National Institute for Environmental Studies, Ibaraki, Japan

22 ⁹Institute of Environmental Physics, University of Bremen, Bremen, Germany

23 ¹⁰Cooperative Institute for Research in Environmental Sciences (CIRES), University of Colorado at Boulder, Boulder, CO,
24 USA

25 ¹¹School of Earth and Environment, University of Leeds, Leeds, UK

26 ¹²Laboratory of Atmospheric Physics, Aristotle University of Thessaloniki, Greece

27 ¹³now at: Instituto de Geociencias, Consejo Superior de Investigaciones Científicas (IGEO-CSIC), Madrid, Spain

28 ¹⁴Deutsches Zentrum für Luft- und Raumfahrt (DLR), Institut für Methodik der Fernerkundung, Oberpfaffenhofen, Germany

29 ¹⁵NASA Langley Research Center, Hampton, VA, USA

30 ¹⁶National Centre for Earth Observation (NCEO), University of Leeds, Leeds, UK

31 ¹⁷Institute of Climate and Energy Systems – Stratosphere (ICE-4), Forschungszentrum Jülich, Germany

32 ¹⁸Deutsches Zentrum für Luft- und Raumfahrt (DLR), Institut für Physik der Atmosphäre, Oberpfaffenhofen, Germany

33 ¹⁹Department of Atmospheric Physics, Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic

34 ²⁰Adnet Systems, Inc., Lanham, MD, USA

35 ²¹NASA Goddard Space Flight Center, Greenbelt, MD, USA

36 ²²Institute of Climate and Energy Systems - Stratosphere, Forschungszentrum Jülich, Jülich, Germany

37 ²³Institute for Atmospheric and Environmental Research, University of Wuppertal, Wuppertal, Germany

38 ²⁴Department of Meteorology, University of Reading, Reading, United Kingdom

39 ²⁵Royal Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium

42 ²⁶CNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France
43 ²⁷Federal Office of Meteorology and Climatology, MeteoSwiss, Payerne, Switzerland
44 ²⁸LATMOS, Institut Pierre-Simon Laplace, Sorbonne Université/CNRS/UVSQ, Paris, France
45 ²⁹National Institute of Water and Atmospheric Research, Te Whanganui-a-Tara / Wellington, Aotearoa New Zealand
46 ³⁰School of Physical and Chemical Sciences, Canterbury University, Ōtautahi / Christchurch, Aotearoa New Zealand
47 ³¹Now at: Deutscher Wetterdienst, Offenbach, Germany
48 ³²Climate Research Division, Environment and Climate Change Canada, Montreal, QC, Canada
49 ³³NOAA Chemical Sciences Laboratory, Boulder, CO, USA (now retired)
50 ³⁴Finnish Meteorological Institute, Helsinki, Finland
51 ³⁵Institute for Atmospheric and Climate Science, ETH Zurich, Switzerland
52 ³⁶Physikalisch Meteorologisches Observatorium Davos/World Radiation Center, Davos, Switzerland
53 ³⁷Aristotle University of Thessaloniki, Thessaloniki, Greece
54 ³⁸School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA, USA
55 ³⁹Science Systems and Applications Inc., MD, USA
56 ⁴⁰Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Yokohama, Japan
57 ⁴¹Advanced Institute for Marine Ecosystem Change (WPI-AIMEC), Tohoku University, Sendai, Japan
58 ⁴²ESSIC/UMD & NOAA/NESDIS/STAR, College Park, MD, USA (now retired)
59 ⁴³Morgan State University, Baltimore, MD, USA
60 [†]Deceased

61 Correspondence to: Sean M. Davis (sean.m.davis@noaa.gov)

62 **Abstract.** A variety of chemical and dynamical processes in the troposphere and stratosphere affect tropical total column
63 ozone (TCO), the net effect of which may cause changes in surface UV radiation and impact human and ecosystem health.
64 We use dynamical linear modeling to estimate **changes in** tropical TCO and partial column ozone (PCO) in the troposphere
65 and three stratospheric layers to assess agreement between satellite composites and chemistry-climate model simulations from
66 two multi-model experiments (CCMI-1 and CCMI-2022). While both model experiments show tropical TCO increases over
67 2000–2021, multimodel-mean CCMI-2022 **changes** (+2.5 DU) agree slightly better with observations (**+3.3 DU**) than CCMI-
68 1 (+1.6 DU). However, this overall agreement obscures multiple systematic differences in PCO **changes** between the models
69 and observations across atmospheric layers. For example, since 2000 tropical tropospheric PCO increased significantly in
70 CCMI-2022 (+1.5 DU) but not in CCMI-1 (+0.3 DU), largely explaining the difference in TCO **changes**. Also, despite nearly
71 identical stratospheric PCO **changes**, CCMI-2022 **changes** are slightly more negative **than CCMI-1** in the lower stratosphere
72 (by ~0.5 DU) **and more positive in the middle/upper stratosphere**. Crucially, substantial differences exist **between** observational
73 PCO **changes**, particularly in the troposphere and middle/upper stratosphere, and these disagreements limit the ability to
74 evaluate CCM fidelity **for past changes**. **While early and late century trends are correlated across models, suggesting a potential**
75 **emergent constraint on ozone, the spread in observational trends means they are unable to provide guidance on the credibility**
76 **of model projections for a given emissions scenario.**

Deleted: trends in
Deleted: observational
Deleted: –
Deleted: (CCM)
Deleted: –
Deleted: trends
Deleted: trends
Deleted: –
Deleted: trends
Deleted: trends
Deleted: trends
Deleted: ic
Deleted: ,
Deleted: compensated by
Deleted: trends compared to CCMI-1
Deleted: across
Deleted: trends
Deleted: Furthermore, while the inter-model correlation between late and early 21st century trends is suggestive of a potential emergent constraint on future ozone trends, the spread in observational trends precludes its observational implementation.

98 **1 Introduction**

99 The amount of biologically harmful ultraviolet radiation reaching Earth's surface is strongly affected by the amount of total
100 column ozone (TCO), and this motivates the need to understand both natural variability and anthropogenically-driven changes
101 in TCO across a range of timescales. TCO includes contributions from ozone residing in the stratosphere (~90%, see Figure
102 1) and the troposphere (~10%), and can vary due to a variety of chemical, physical, and radiative processes that are impacted
103 by human activities.

104 In the 1970s and 1980s, reductions in stratospheric ozone from the emission of ozone depleting substances (ODS) became a
105 topic of grave societal and scientific concern, and efforts to address this issue culminated in the adoption of the Montreal
106 Protocol on Substances that Deplete the Ozone Layer and subsequent amendments. Because of the Montreal Protocol, the
107 concentration of ODS has been declining in the atmosphere since around 1997, and ozone concentrations in the upper
108 stratosphere and the Antarctic ozone hole are exhibiting signs of recovery (WMO, 2022). However, during this same period,
109 vertically resolved satellite observations of lower stratosphere ozone generally show declines in the tropics (Ball et al., 2018;
110 Ball et al., 2020; Dietmüller et al., 2021; Godin, Beekmann et al., 2022; Sofieva et al., 2026). The tropical lower stratosphere
111 ozone decrease is broadly consistent with the expected greenhouse gas forced acceleration of the lower branch of the Brewer-
112 Dobson circulation (Hegglin and Shepherd, 2009), although there remains uncertainty around the robustness of both the
113 dynamical and ozone trends (Davis et al., 2023; Sofieva et al., 2026), as well as the importance of other factors such as
114 tropopause changes (Thompson et al., 2021) and very short-lived halogens (Villamayor et al., 2023).

115 In addition to the previous inconsistencies noted in recent tropical lower stratospheric ozone trends, this work is also motivated
116 by a broader need to understand recent and future tropical TCO, as this is a region with significant human population, high
117 biodiversity, and strong influence on the terrestrial carbon sink which may be particularly sensitive to changes in surface ozone
118 and UV radiation (e.g., Young et al., 2021). For example, the region between 30°S and 30°N covers half of Earth's surface
119 and is currently home to ~60% of the human population, with that number growing to two-thirds by the late 21st century under
120 a middle of the road Shared Socioeconomic Pathway (SSP) scenario (SSP2, Jones and O'Neill, 2016).

121 Scientifically, the tropics (defined here as 30°S–30°N) are also of particular interest because they are subject to ozone changes
122 that are potentially opposite in sign in different regions of the atmosphere. Compared to higher latitudes (and notwithstanding
123 the Antarctic ozone hole), tropical TCO values are smaller due to the region having a higher tropopause and upwelling as part
124 of the stratospheric overturning circulation (Dobson, 1956; Dobson et al., 1929). Related to this, tropical TCO may be
125 particularly sensitive to dynamically-induced variations in stratospheric ozone. In the stratosphere, an acceleration of the lower
126 branch of the Brewer-Dobson circulation (BDC) is expected through the 21st century due to rising GHG levels (Butchart,
127 2014), and this is expected to decrease tropical lower stratospheric ozone (e.g., Hegglin and Shepherd, 2009; Lamy et al.,
128 2019). Acting in opposition to this, trends in the upper stratosphere are expected to be positive given ongoing chlorine
129 reductions and greenhouse gas-driven cooling (WMO, 2022). These potentially offsetting processes create some ambiguity

Deleted: (UV-B, 280–315 nm)

Deleted: varies due to several factors including latitude, season, cloud cover (both frequency and thickness), and

Deleted: amount, with TCO filtering ~95% of the UV-B radiation incident at the top of the atmosphere

Deleted: . The

Deleted: .

Deleted: A

Deleted: impact ozone, and both secular changes and variations in TCO related to these processes can modulate surface UV

Deleted: D

Deleted: since ~2000

Deleted: and Northern Hemisphere (NH) midlatitudes

Deleted: Benito-Barca et al., 2025;

Deleted: –

Deleted: –

Deleted: –

Deleted: Although tropical lower stratospheric ozone and upwelling mass flux are highly anti-correlated in models, post-2000 upwelling trends from reanalyses vary widely and in some cases are inconsistent with observed tropical ozone changes (Davis et al, 2023). ...

Deleted: –

Deleted: offsetting in nature

Deleted: –

Deleted: –

Deleted: –

157 regarding future changes in the tropics, in contrast to the midlatitudes, where ozone sensitivity to Cl changes (and hence future
158 ozone recovery) is greater and where the ozone-governing processes act in congruence with one another.

159 To add to this complexity, in the tropical troposphere ozone changes may be occurring due to changes in anthropogenic
160 precursor emissions (Shepherd et al., 2014; Zhang et al., 2016), and these emissions are generally expected to increase and
161 then peak and decrease over the course of the 21st century, in contrast to the expected continual decrease of precursor emissions
162 over the NH midlatitudes. With the ongoing development and increased precursor emissions associated with population and
163 economic growth of the region in the past several decades, it is possible that the tropical tropospheric ozone column has already
164 increased, although observational estimates varied widely in the first Tropospheric Ozone Assessment Report (TOAR, Gaudel
165 et al., 2018). However, more recently phase 2 of the Tropospheric Ozone Assessment Report (TOAR-II, Hubert et al., 2026)
166 assessed fifteen satellite records and found much improved agreement compared to the first TOAR assessment. They show
167 that tropical mean tropospheric ozone columns increased between 2003 and 2012, were roughly constant until 2020, and then
168 decreased after 2020 until the end of their assessment period in 2022.

169 Taken together, these potentially offsetting factors affecting TCO motivate a comprehensive analysis of both stratospheric and
170 tropospheric changes in the tropics. We therefore analyze partial column ozone (PCO) covering distinct vertical regions within
171 the troposphere and stratosphere in order to attribute their contribution to changes in TCO and assess agreement between
172 satellite observations and chemistry-climate models (CCMs). An example climatological ozone profile for the tropics is shown
173 in Fig. 1 to illustrate the relative contributions of ozone in the different atmospheric layers to TCO, along with the four regions
174 defined for the analysis in this paper: troposphere (surface to tropopause, ~13% of TCO), lower stratosphere (tropopause to 30
175 hPa/24 km, ~28%), middle stratosphere (30 hPa/24 km to 10 hPa/31 km, ~41%), and upper stratosphere (10 hPa/31 km to 1
176 hPa/48 km, ~18%). While the choice of these four layers is ultimately somewhat arbitrary, these layers roughly encompass the
177 available observational datasets and represent regions with somewhat distinct processes controlling the distribution of ozone.
178 For example, ozone variability in the lower stratosphere is dominated by transport and the upper stratosphere is dominated by
179 chemistry, with a mixture of chemical and dynamical processes contributing in the middle stratosphere.

Deleted: y

Deleted: and are subject to active ongoing analysis

Deleted: –

Deleted: 7

Deleted: 0

Deleted: 20

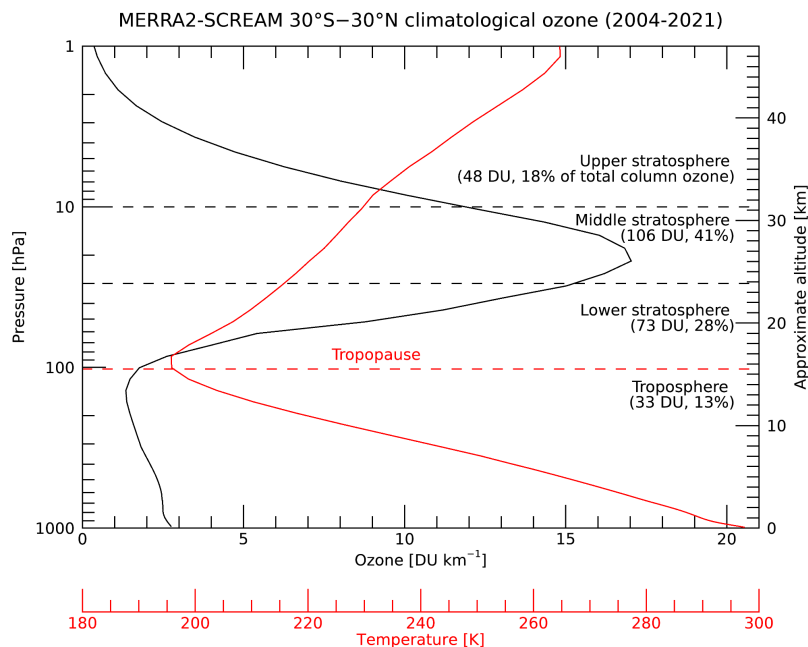


Figure 1: The 2004–2021 climatological tropical average (30°S–30°N) ozone profile from MERRA-2 Stratospheric Composition Reanalysis of Aura MLS (MERRA2-SCREAM, black), along with the corresponding temperature profile (red). The relative contributions to total column ozone from each of the four vertical layers considered in this paper are given on the right hand side.

The first sections (sections 2–4) of the paper detail the observational datasets, CCMs, and statistical methods used. Then, in section 5 we intercompare the recent ozone **changes** simulated by CCMs and derived from observations to illustrate regions of agreement and disagreement. Next (section 6), we analyze the long-term ozone **changes** and demonstrate that in the models, data from the first ~2 decade period of ozone recovery is sufficient to separate the slowly **varying** “recovery” signal from natural variability, and (section 7) that the models’ recovery signal over this period is predictive of the end-of-the-century behavior. While this emergent behavior in the models lends itself in principle to an observational constraint on **the credibility of modeled** future tropical ozone **for a given scenario**, we demonstrate in section 8 that current **range of** observational trend estimates **is too large** to effectively serve as an observational constraint.

Deleted: trends

Deleted: trends

Deleted: –

Deleted: –

Deleted: –

Deleted: span too wide a range

205 **2 Ozone observations and reanalyses**

206 [This section describes the satellite observational and reanalysis datasets that are used for estimating historical ozone changes.](#)
207 [The datasets include vertically-resolved stratospheric data, from which we compute partial column ozone timeseries across](#)
208 [various layers of the stratosphere, as well as total/partial column datasets that provide pre-computed values. The datasets](#)
209 [described herein are summarized in Table 1.](#)

210

211

TABLE 1. Observational datasets

Formatted: Centered

Name	Figure label	Version	Reference	Start Year
Vertically-resolved datasets				
SWOOSH	SWO	2.73	Davis et al. 2016	1984
SAGE-CCI-OMPS+	CCI	v2/fv0008	Sofieva et al. 2023	1984
SAGE-Osiris-OMPS	SOO	n/a	Bourassa et al. 2014	1984
SAGE-Osiris-SAGE	SOS	n/a	Bognar et al., 2022	1984
SAGE-SCIAMACHY-OMPS	SSO	n/a	Arosio et al., 2019	1985
GOZCARDS	GOZ	2.2	Froideveaux et al. 2015	1979
MERRA2-SCREAM	M2S	n/a	Wargan et al., 2023	2004
Tropospheric column				
TOMS-OMI/MLS	TOMS-OMI/MLS	n/a	Ziemke et al., 1998; 2019	1979
TOMS-OMI/CCD	TOMS-OMI/CCD	n/a	Ziemke et al., 1998; 2019	1979
OMI/MLS	OMI/MLS	n/a	Ziemke et al., 2019	2005
MERRA2-SCREAM	M2S	n/a	Wargan et al., 2023	2004
SUNLIT*	SUN	fv001	Sofieva et al. 2022	2004
TOAR II	TOAR II	v1.0	Hubert et al., submitted	2005
TCO datasets				
SBUV-COH	SBUV-COH	v8.6/20250612	Wild et al., 2025	1978
SBUV-MOD	SBUV-MOD	v8.7	Frith et al., 2014; 2017	1979
World Ozone and Ultraviolet Radiation Data Centre	WOUDC	n/a	Fioletov et al., 2002; 2008	1964
Multi-Sensor Reanalysis-2	MSR	fv021-025	van der A et al., 2015	1979
GOME-SCIAMACHY-GOME	GSG	n/a	Weber et al., 2018	1995
GOME-type Total Ozone Essential Climate Variable	GTO	n/a	Coldevey-Egbers et al., 2022	1995
MERRA2-SCREAM	M2S	n/a	Wargan et al., 2023	2004
Bodeker Scientific TCO	BSTCO	3.4.1	Bodeker et al., 2021	1978**

212 [* Dataset also includes stratospheric column ozone](#)

213

** Dataset ends in 2019.

Deleted: ¶

214

2.1 Vertically-resolved datasets

Deleted: –

215

We briefly describe here several merged datasets based on vertically-resolved ozone retrievals from satellite instruments

Deleted: –

216

operating in solar occultation and limb-viewing geometries. These datasets are provided as monthly zonal means with

Deleted: –

217

horizontal grid spacings from 2.5° to 10° in latitude. We note that in general many of these datasets rely on overlapping subsets

218

of instruments, and so are not entirely independent. Of particular note, all of the merged datasets here include data from the

219

Stratospheric Aerosol and Gas Experiment II (SAGE II) instrument (1984–2005), and as such many of the datasets start part

220

way through 1984 with the beginning of the SAGE II record. Also, a number of datasets are subject to large changes in tropical

221

sampling (e.g., with the inclusion of the Aura Microwave Limb Sounder (MLS) in 2004 or the Ozone Profiler and Mapping

222

Suite Limb Profiler (OMPS-LP) in 2012), and these sampling changes have potential impacts on the ability to constrain ozone

223

trends. As an example, whereas the solar occultation instruments collect ~30 vertical ozone profiles per day and may go entire

224

months without sampling the tropics (depending on orbit), the median tropical sampling for Aura MLS and OMPS-LP is 1200

Deleted: –

225

and 800 profiles per day, respectively (SPARC, 2017).

226

Two of the merged datasets used here rely heavily on Aura MLS measurements. The first of these is the Stratospheric Water

227

and OzOne Satellite Homogenized (SWOOSH) dataset (Davis et al., 2016) version 2.7.2, whose merged ozone product includes

Deleted: 2

228

data from SAGE II (v7), the Upper Atmosphere Research Satellite Halogen Occultation Experiment (UARS HALOE v19),

229

UARS MLS (v5), SAGE III/Meteor (v4), Aura MLS (v5.0), Atmospheric Chemistry Experiment Fourier Transform

230

Spectrometer (ACE-FTS, v5.2), and SAGE III/ISS (v5.3). In SWOOSH, the merged product in a given latitude-pressure

Deleted: –

231

month bin is the mean of all available monthly means from each source instrument operating in that month, weighted by the

Deleted: –

232

number of profiles for the given instrument. Because of this weighting, instruments with dense sampling such as Aura MLS

Deleted: –

233

dominate the merged product.

234

The other Aura MLS-based merged dataset considered here is the Global OZone Chemistry And Related trace gas Data records

Deleted: –

235

for the Stratosphere v2.2 (GOZCARDS, Froideveaux et al. 2015). GOZCARDS includes SAGE I (v5.9, 1979–1981), SAGE

236

II (v7), and Aura MLS (v4.2). Although SWOOSH and GOZCARDS are similar, the two datasets use a slightly different

237

approach for homogenization during instrument overlap periods, and GOZCARDS weights each source instrument equally in

238

the combined product, unlike SWOOSH, which weights based on the number of profiles. In practice, the different weightings

239

are most likely to cause differences between the products prior to the end of 2005, when the HALOE and SAGE II records

240

terminated.

241

In addition to the Aura MLS-based datasets, we include in our analysis several datasets that rely on different combinations of

Deleted: W

242

instruments in the 2000’s. First is the SAGE-CCI-OMPS+ (CCI) dataset (v2/fv0008, Sofieva et al., 2023), which combines

Deleted: e also include

243

data from SAGE II (v. 7), three instruments (GOMOS ALGOM2s v.1, MIPAS KIT v8, SCIAMACHY UBr v3.5) aboard the

Deleted: –

244

Environmental Satellite (Envisat, 2002–2012), Optical Spectrograph and InfraRed Imaging System (OSIRIS, v7.2, beginning

Deleted: –

245

2002), ACE-FTS, and two retrievals of SNPP OMPS-LP data (UBr v3.5, Arosio et al., 2018; USask2D v1.1.0, Zawada et al.,

Deleted: includes

Deleted: –

Deleted: –

329 We also use data from the gap-filled Bodeker Scientific TCO dataset (BSTCO v3.4.1, Bodeker et al., 2021). This dataset
330 includes data from 17 different satellite instruments operating since the late 1970's, including multiple SBUV, GOME, and
331 TOMS instruments, as well as SCIAMACHY, OMI, and OMPS Nadir Mapper. Note, that this dataset was not used by Weber
332 et al. (2022).

333 In addition to the merged satellite data records, we also use monthly mean zonal mean ground-based Dobson and Brewer data
334 from the World Ozone and Ultraviolet Radiation Data Centre (WOUDC) dataset (Fioletov et al., 2002; Fioletov et al., 2008).
335 Finally, we also use TCO data from two assimilated products that were not considered by Weber et al. (2022): the Multi-
336 Sensor Reanalysis-2 (MSR-2, hereinafter MSR; van der A et al., 2015) and M2-SCREAM. MSR bias-corrects all available
337 TCO data from polar-orbiting satellites based on ground-based Brewer/Dobson instruments from WOUDC, and then
338 assimilates the debiased observations using the Tracer Model Data Assimilation Model (TMDAM) ozone chemistry and data
339 assimilation model driven by ECMWF analysis fields.

340 **2.3 Tropospheric column data**

341 We consider tropospheric partial column ozone (TrCO) from several datasets here. These satellite-based datasets use a
342 'residual' technique whereby the measured or integrated (e.g., from a retrieved ozone profile) stratospheric partial column
343 ozone value is subtracted from a TCO observation.

344 One residual tropospheric column dataset comes from Aura OMI and Aura MLS (hereinafter OMI/MLS, Ziemke et al., 2019),
345 which uses the OMI TCO and the stratospheric column from the Aura MLS v4.2 data above the WMO lapse rate tropopause
346 provided by the National Centers for Environmental Prediction (NCEP) reanalysis.

347 Another residual dataset is provided by the Synergy of Using Nadir and Limb Instruments for Tropospheric Ozone Monitoring
348 (SUNLIT) project (Sofieva et al. 2022). SUNLIT covers 2004 to present using OMI and TROPOMI TCO, combined with a
349 separate stratospheric column dataset consisting of Aura MLS, OSIRIS, Michelson Interferometer for Passive Atmospheric
350 Sounding (MIPAS), SCanning Imaging Spectrometer for Atmospheric CHartographY (SCIAMACHY), S-NPP OMPS LP,
351 and Global Ozone Monitoring by Occultation of Stars (GOMOS).

352 Two additional TrCO datasets come from the Total Ozone Mapping Spectrometer convective-cloud differential approach
353 (TOMS CCD; Ziemke et al., 1998, 2019) which uses ozone columns retrieved above deep convective clouds as a proxy for
354 the stratospheric column, and subtracts these values from the total column retrieved from clear-sky scenes. This CCD technique
355 is also applied to OMI data to form a TOMS + OMI CCD dataset from 1979 to present. Additionally, we also consider a
356 merged dataset consisting of TOMS CCD before Oct 2004, and OMI/MLS after (hereinafter TOMS+OMI/MLS).

357 In addition to the satellite merged data sets, we also use TrCO from M2-SCREAM. As M2-SCREAM assimilates MLS ozone
358 in the stratosphere and OMI TCO, it is not expected that this product would be substantially different from the OMI/MLS
359 dataset. However, M2-SCREAM also assimilates other ozone-relevant stratospheric species from MLS and represents
360 meteorological variables that could impact ozone, both in the troposphere and stratosphere, so it is not expected to be identical
361 to the OMI/MLS dataset.

Deleted: –

Deleted: ,

Deleted: –

Deleted: o

Deleted: c

Deleted: –

Deleted: –

Deleted: –

Deleted: –

Deleted: –

Deleted: –

Deleted: –

Deleted: –

Deleted:

Deleted: –

Deleted: –

Deleted: –

Deleted: Alternatively

Deleted: Finally

Deleted: –

Deleted: –

Deleted: –

Deleted: –

385 Finally, we consider a multi-dataset merged product produced by the TOAR II activity (Hubert et al., submitted). This TOAR
386 II dataset consists of 15 satellite TrCO data records spanning nadir profiling, tropospheric ozone residual (e.g., like OMI/MLS),
387 and convective-cloud differential techniques. We note that the datasets contributing to the TOAR II multi-dataset product
388 include the OMI/MLS and SUNLIT datasets also analyzed separately in this paper.

389 **2.3 Stratospheric column data**

390 Ozone from the vertically-resolved satellite datasets (SWOOSH, GOZCARDS, SOO, SOS, CCI) and M2-SCREAM is
391 vertically integrated to obtain partial ozone columns in the stratosphere. The whole-stratosphere PCO is computed by
392 integrating from the MERRA-2 zonal mean tropopause to 1 hPa (or 48 km for the altitude-gridded datasets). Prior to this
393 integration, we infill any missing ozone data values by interpolating merged ozone anomalies in the latitude-height plane, as
394 in Davis et al. (2016). In addition to the integration over the entire stratosphere, we also consider PCO in layers covering the
395 lower stratosphere (tropopause to 30 hPa/24 km), middle stratosphere (30 hPa/24 km to 10 hPa/31 km) and upper stratosphere
396 (10 hPa/31 km to 1 hPa/48 km), as illustrated in Fig. 1.

397 In addition to the vertically-resolved datasets that we integrate to get stratospheric PCO, we also consider the stratospheric
398 column data produced as part of the SUNLIT dataset (Sofieva et al., 2022). The SUNLIT stratospheric column represents a
399 vertical integration over the stratosphere from a combination of limb instruments including OSIRIS, GOMOS, SCIAMACHY,
400 MIPAS, OMPS-~~LP~~, and Aura MLS.

401 **3. Chemistry Climate Models**

402 An overview of the Chemistry Climate Models (CCMs) used in this study is provided in Table 2. Phase 1 of the Chemistry
403 Climate Model Initiative (CCMI-1) involved a set of coordinated simulations from a set of 20 CCMs and Chemistry Transport
404 Models (CTMs) that explicitly represent atmospheric chemistry and are suitable for studying a variety of processes including
405 both stratospheric and tropospheric chemistry, as well as global climate change (Morgenstern et al., 2017). Models
406 participating in CCMI-1 were generally of the same generation as the 5th Coupled Model Intercomparison Project (CMIP5)
407 and used forcings similar to that activity, but the reader is referred to Morgenstern et al. (2017) for further details on individual
408 models. A variety of simulations covering the middle 20th century through the end of the 21st century were performed as part
409 of CCMI-1, but only the relevant experiments are briefly described here. The REF-C2 simulations cover 1960 through 2100
410 using ozone depleting substance (ODS) concentrations from the WMO 2010 Scientific Assessment of Ozone Depletion
411 (SAOD) A1 scenario (WMO, 2011), and other forcings including aerosols, greenhouse gases, and tropospheric ozone
412 precursors from Representative Concentration Pathway (RCP) 6.0 (Meinshausen et al., 2011). In REF-C2 simulations, the
413 ocean state is either coupled to the atmosphere model or prescribed based on an offline ocean simulation. A comprehensive
414 description of the REF-C2 experimental design is provided by Eyring et al. (2013). Additionally, we include analysis of the

Deleted: –

Deleted: –

Deleted: –

Deleted: –

Deleted: –

Deleted: –

Deleted: –

Deleted: able 1

Deleted: –

Deleted: –

Deleted: –

Deleted: –

Deleted: –

Deleted: –

Deleted: (2011)

Deleted: –

Deleted: –

432 fixedGHG scenario, which is identical to REF-C2 except with GHGs fixed at 1960 values to isolate the impacts of ODSs
433 separately.

434 In addition to CCMI-1 model output, we also consider simulations performed as part of the follow-up CCMI-2022 project
435 (Plummer et al., 2021). We use the CCMI-2022 future projection simulations (REF-D2, 1960 to 2100), which are a similar
436 configuration to the older CCMI-1 REF-C2 simulations. CCMI-2022 model generations and forcings are generally similar to
437 those from CMIP6. For REF-D2, greenhouse gas and anthropogenic precursor forcings are from the CMIP6 historical database
438 (Meinshausen et al., 2017) up to 2014 and SSP2-4.5 thereafter (Meinshausen et al., 2020; O'Neill et al., 2016). Near-surface
439 ODS mixing ratios are specified based on the WMO (2018) baseline scenario. In comparison to the WMO 2010 scenario, the
440 WMO 2018 scenario has a virtually identical peak value of equivalent effective stratospheric chlorine (EESC), but a slower
441 decline that results in a 5 year delay in the return date to 1980 values (Lickley et al., 2018). A comprehensive description of
442 the CCMI-2022 simulations is provided by Plummer et al. (2021).

443

- Deleted: –
- Deleted: –
- Deleted: –
- Deleted: –
- Deleted: corresponding
- Deleted: –
- Deleted: –
- Deleted: –
- Deleted: –
- Deleted: –

TABLE 2. Chemistry climate models used in this study to simulate ozone from 1960 to 2100

CCMI-1				CCMI-2022			
REF-C2				REF-D2			
model name	# realizations	QBO	Solar	model name	# realizations	QBO	Solar
ACCESS-CCM	2	Y	N	ACCESS-CM2-Chem	1	Y	Y
CCSRNIES-MIROC3.2	2	Y	Y	CCSRNIES-MIROC32	1	Y	Y
CESM1-CAM4Chem	3	Y	Y				
CESM1-WACCM	3	Y	Y	CESM2-WACCM	3	Y	Y
CHASER-MIROC-ESM	1	N	Y%	MIROC-ES2H	3	Y	Y
CMAM	1	N	N	CMAM	3	Y	Y
CNRM-CM5-3	2	N	Y%	CNRM-MOCAGE*	4	Y	N
EMAC	2 (L47), 1 (L90)	Y	Y%	EMAC	3	Y	Y
GEOSCCM	1	Y	N	GEOSCCM	1	Y	Y%
GFDL-CM3	1	Y	N**				
HadGEM3-ES	1	Y	Y%				
IPSL	1	-	-	IPSL-CM6A-ATM-LR-REPROBUS	0	-	-
MRI-ESM1r1	1	Y	Y				
NIWA-UKCA	5	Y	N	NIWA-UKCA2	3	Y	N
SOCOL3	1	Y	Y	SOCOL4	3	Y	Y
ULAQ-CCM	3	Y	Y%				
UMSLIMCAT	1	Y	N				

* Tropospheric ozone not included in analysis, due to unphysical jumps.

** Solar variability in radiative scheme, but not implemented in photolysis code.

% Non-standard solar forcing used.

Deleted: 1

Deleted: *

Deleted: -

465 For REF-C2 simulations, most of the participating models used nudging of their equatorial winds to represent the QBO (Table
466 2, see also Table S8, Morgenstern et al., 2017), and the other models either had no QBO at all or an internally generated one.
467 For the former, a forcing dataset was provided based on observed equatorial winds through 2011 that then repeats past cycles
468 into the future. As this is not explicitly documented, we report based on visual inspection (not shown), that seven REF-C2
469 models are nudged towards this forcing dataset, four models have no apparent QBO wind variability, and the rest exhibit QBO
470 variability in their winds that is not synchronized with observations in the pre-2012 period.
471 Overall, the QBO representation in models improved from the CMIP5-class models used in CCMI-1 to the CMIP6 models
472 used in CCMI-2022 (Richter et al., 2020), and the specification of QBO treatment in REF-D2 resulted in a more consistent
473 treatment across models. In REF-D2, models containing an internally generated QBO were allowed to generate their own
474 QBO, whereas models that do not were directed to nudge towards an extended QBO dataset. In contrast to REF-C2, all REF-
475 D2 models exhibit QBO-like variability in their equatorial stratospheric winds, and only three models appear to have nudged
476 their wind fields towards historical observations.

- Deleted: CCMI-1
- Deleted: -
- Deleted: able 1
- Deleted: -
- Deleted: -
- Deleted: -
- Deleted: -
- Deleted: CCMI-2022
- Deleted: -
- Deleted: -
- Deleted: -

477 **4. Dynamical linear modeling**

478 Dynamical linear modeling (DLM) is a statistical method for time series modelling that is functionally similar to multiple
479 linear regression (MLR), but with several important differences, most notably the use of a flexible non-parametric background
480 term for capturing “trends”, as well as principled treatment of auto-regressive processes and time-varying uncertainties (Laine
481 et al., 2014). The DLM technique has been used in a number of recent studies to quantify ozone depletion and recovery trends
482 (Ball et al., 2017, 2018, 2019; Bognar et al., 2022; Davis et al., 2023; Maillard Barras et al., 2022; Van Malderen et al., 2025).
483 In this study, we use the dlm_vanilla_ar1 model from the dlmmc software package (Alsing, 2019) with its default
484 hyperparameter settings. This model includes dynamic seasonal cycles (2 cos/sin harmonic pairs) and background trend, fixed
485 (non-dynamic) proxy terms (described in the next section), as well as a first-order autoregressive process (AR1). We run the
486 DLM for each observational and model dataset, and for each atmospheric layer described above, using monthly mean tropical
487 averaged (30° S–30° N) PCO/TCO. From the DLM, we infer ozone changes by sampling the posterior distribution of the
488 background trend term from simulations with 7000 iterations (2000 of which are used as warmup and discarded), and report
489 either the percentage of posterior changes of a given sign (e.g., 95 %) or the standard deviation of the posterior distribution as
490 measures of trend significance and uncertainty, respectively. In some cases, we also show kernel density estimates (KDEs) of
491 changes in the DLM background ozone values. KDEs are a non-parametric method used to estimate the probability density
492 function (PDF) of a variable by applying a smoothing kernel, and by construction are normalized to unity. For each
493 observational dataset, KDEs are computed from the 5000 iterations of the DLM model. Multi-model KDEs are constructed by
494 sampling 5000 DLM iterations across all available realizations of each CCM participating in the given multi-model
495 experiment. For models with only one realization, all 5000 iterations of the DLM are used. Similarly, we construct the multi-
496 observational KDE by sampling across the 5000 DLM iterations available from each observational data set. We note that the

- Deleted: -
- Deleted: -
- Deleted: -
- Deleted: -
- Deleted: trends
- Deleted: trends
- Deleted: T

choice of 7000 iterations is likely far more than necessary, but represents a balance between sample size, reasonable computing time, and stability of the hyperparameters. Our testing has indicated virtually identical results for an order of magnitude fewer iterations.

As with other studies, for the DLM simulations on CCM output we do not input uncertainty values to the DLM model. When running the DLM on the observational composites, we use a 1% uncertainty. While this number may seem small given typical reported precision uncertainties of ~10% from satellite limb instrument ozone retrievals, we note that these random errors are substantially reduced by both vertical averaging and the averaging of the many profiles within the tropics for a given month. The 1% number used here is a reasonable choice based on testing we did (not shown) with the SWOOSH data using propagation of errors including covariance estimates (i.e., not assuming independent measurements), and is also justified based on the clear geophysical variability present in the PCO/TCO timeseries that is on the order of 1% (see, e.g., example in section 4.2). We note that, while the DLM posterior uncertainties are influenced by data uncertainties input to the DLM, the output uncertainties are non-zero even for zero input uncertainty, and are not strongly sensitive to the input uncertainties (e.g., doubling the input uncertainty from 1% to 2% does not double the posterior uncertainty).

4.1 Regressor proxies

Here we briefly describe the regressor proxies used in the DLM runs of PCO from models and observations. As our goal is to isolate the slowly-varying background ozone trend term consistently from models and observations, we use proxies that are appropriate for each dataset.

For applying to observations, we use as proxies the detrended Oceanic Niño 3.4 Index from the NOAA Climate Prediction Center to represent the El Niño Southern Oscillation (ENSO), the first two principal components (PCs) from the MERRA-2 reanalysis equatorial stratospheric (100–10 hPa) winds to represent the QBO, the 30 cm radio flux to represent the 11-year solar cycle, and the tropical (30°S–30°N) stratospheric aerosol optical depth (SAOD) from the most recent (v2.21) GloSSAC dataset (Kovilakam et al., 2020).

For models, we compute the ENSO index by averaging and detrending the surface temperatures from each simulation over the Niño 3.4 region (170° W–120° W, 5° S–5° N). For representing the QBO we use the first two PC timeseries computed for each model using the same methodology described above for MERRA-2.

For representing the solar cycle in REF-D2, we use the recommended CMIP6 solar forcing as described by the SOLARIS-HEPPA working group (<https://solarisheppa.geomar.de/cmip6>, last access: 20 May 2024). We identify which models represent solar ozone variability (see Table 2) by identifying those with a significant correlation between the recommended solar forcing timeseries and a smoothed version of the tropical upper stratospheric partial ozone column (i.e., the residual after fitting using the dlm_noregs_ar1 model), similar to the approach taken by Funk and Matthes (2020, see their Fig. 19). We find that in REF-D2, models either used the recommended solar forcing or don't exhibit solar cycle ozone variability.

In contrast, using the same type of analysis with REF-C2 output, we find that a number of models, while including solar cycle variability, did not implement the recommended forcing (Table 2; forcing details at <https://solarisheppa.geomar.de/ccmi>, last

Deleted: –

Deleted: –

Deleted: –

Deleted: CCMI

Deleted: –

Deleted: –

Deleted: –

Deleted: CCMI

Deleted: and detrending

Deleted: –

Deleted: –

Deleted: –

Deleted: able 1

Deleted: –

Deleted: –

Deleted: CCMI-1

Deleted: able 1

568 access: 20 May 2024). While all models used a solar forcing consistent with observations prior to 2011, after this time many
569 of the models repeated a different solar cycle (or set of solar cycles) than the recommendation, which was to repeat a sequence
570 of the last four solar cycles (20–23). As a result, the correlation coefficients between the recommended forcing and residual
571 ozone are much smaller after 2011 in some models, as the phasing between the two drifts apart. For these models, we used the
572 appropriate solar forcing provided by the modelling groups directly, or in one case (CHASER, MIROC, ESM) inferred the
573 solar input based on mesospheric O¹D photolysis rate output at 0.1 hPa.

574 **4.2 DLM example**

575 To illustrate the DLM fitting process and give a qualitative sense of what the DLM fits look like, Figure 2 shows an example
576 fit to one observational dataset (SWOOSH) and one REF-C2 model (CESM1, WACCM) for the tropical stratospheric PCO.
577 Although DLM fits are performed on monthly data, the figure shows annual mean data to better illustrate interannual
578 variability, such as that due to ENSO, QBO, and solar cycle variability. This figure highlights the large interannual variability
579 over the SWOOSH observational period in comparison to the relatively small DLM background trend changes (Fig. 2a), which
580 ultimately contribute to the uncertainties in the inferred background change. The solar component of the fits demonstrate the
581 prominent 11-year solar cycle in both timeseries to illustrate why it is important to include this regressor, in particular when
582 attempting to extract a background trend over shorter time periods such as the SWOOSH observational record.

Deleted: –

Deleted: –

Deleted: CCMI-1

Deleted: –

Deleted: T

Deleted: –

Deleted: stratospheric PCO and

Deleted: s

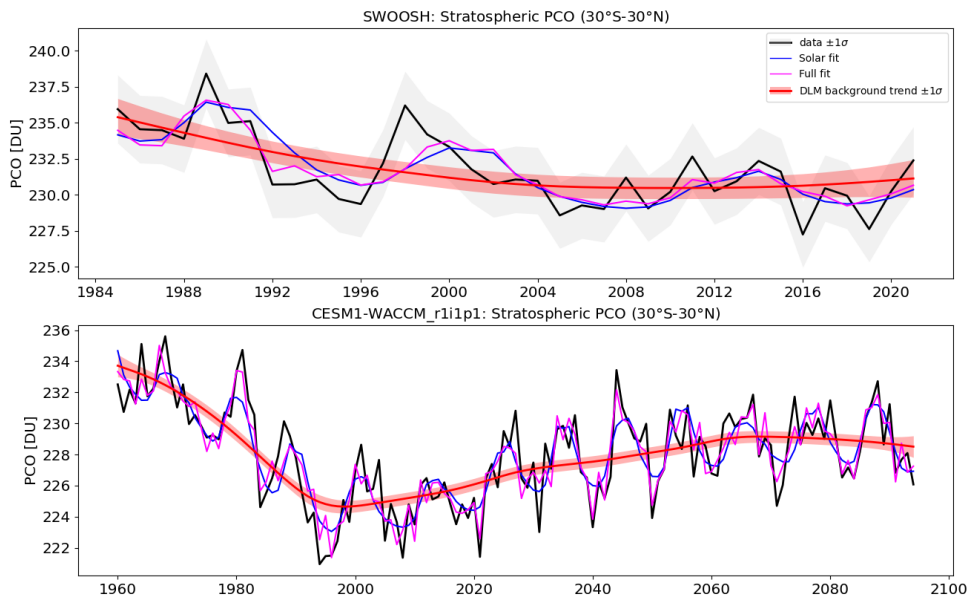


Figure 2: Example DLM fits to the annual and tropical (30° S–30° N) mean partial column ozone (PCO) time series for the whole stratosphere (tropopause to 1 hPa, black lines) from (a) the SWOOSH observational dataset (with $\pm 1\sigma$ uncertainty shading in gray) and (b) a CESM1-WACCM simulation from REF-C2. Shown are the DLM background trend terms (red line is the mean trend with $\pm 1\sigma$ uncertainty shading), which are the primary focus of the analysis presented here. Also shown are the 11-year solar cycle component of the fit (blue) along with the full DLM fit (i.e., trend + all proxies including QBO, ENSO, Solar, and SAOD).

5. Tropical ozone changes: satellite observational period

In this section, we consider tropical ozone changes over the observational time period, starting in 1979 for total column and tropospheric partial column data, and 1985 for stratospheric partial column data. These choices coincide with the first full year of the available datasets, which are constrained by the starting dates of the TOMS (Nov. 1978) and SAGE II datasets (Oct. 1984). For the end date, we use 2021, as it occurs prior to the eruption of the Hunga volcano, which was shown to have impacted stratospheric composition following the eruption, with potential implications for ozone (APARC, 2025; Santee et al. 2023). Any ozone impacts from this eruption (e.g., from stratospheric aerosol and water vapor perturbations) are not included in either of the REF-C2 or REF-D2 simulations, but are potentially present in the observational datasets.

Deleted: A

Deleted: –

Deleted: with $\pm 1\sigma$ uncertainty shading

Deleted: –

Deleted: the CCMI-1

Deleted: –

Deleted: scenario

Deleted: –

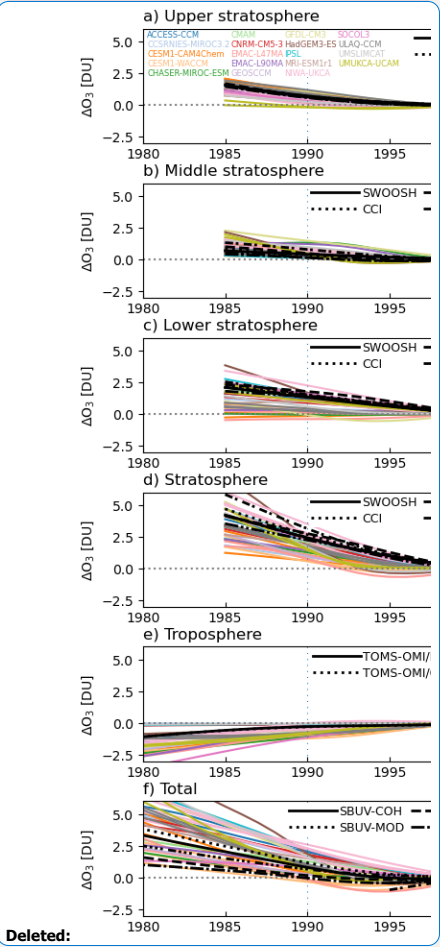
Deleted: O

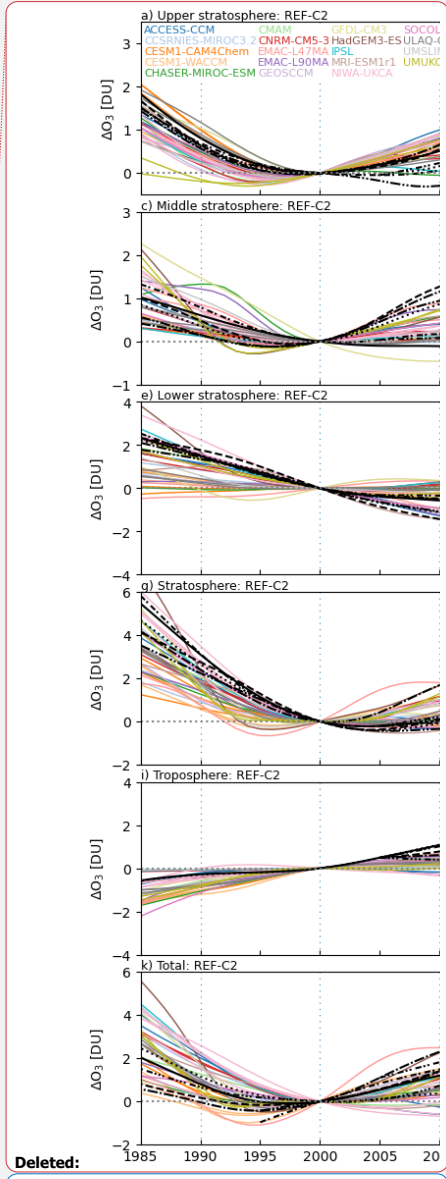
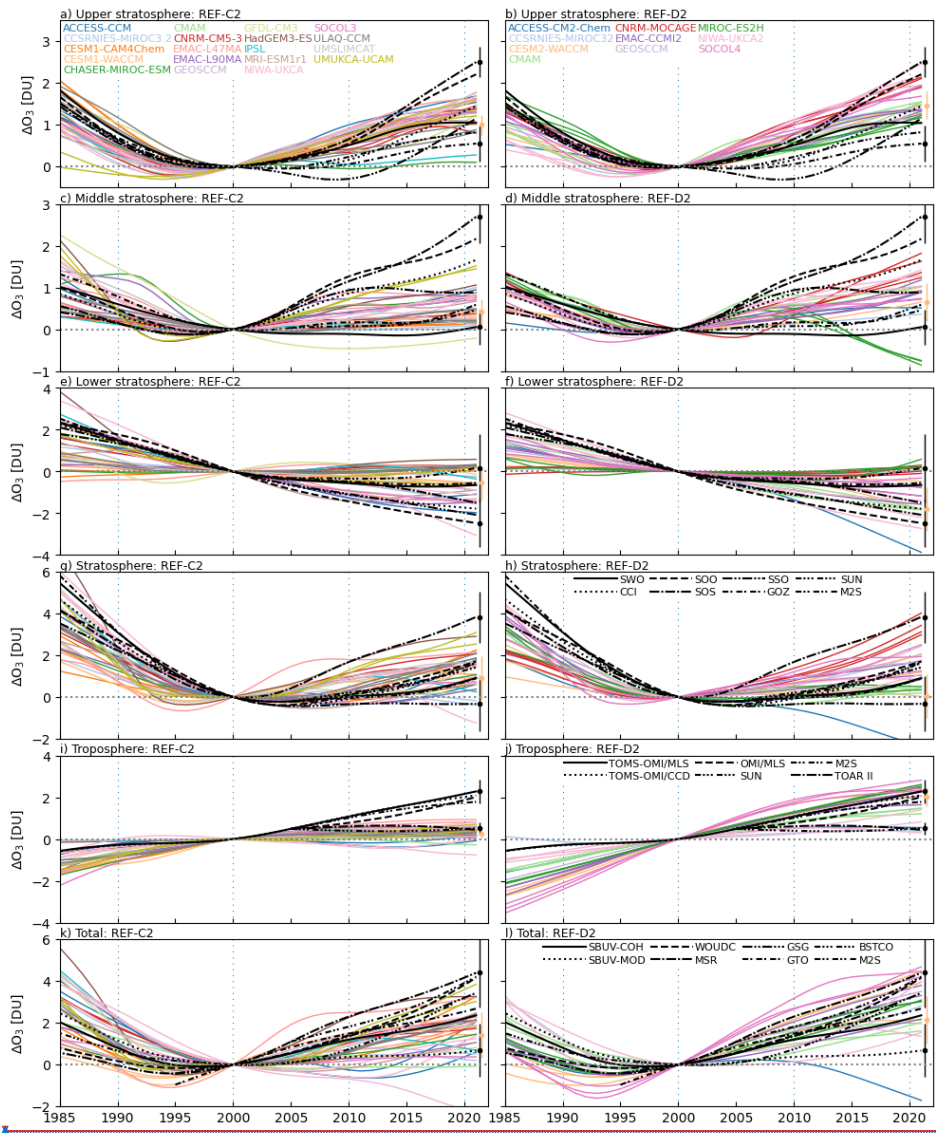
Deleted: –

Deleted: –

616 Figure 3 shows the DLM background trend term from the fit to the tropical average (30°S–30°N) PCO in three stratospheric
617 layers, the whole stratosphere, and troposphere, as well as TCO from both the observations and REF_C2 model simulations.
618 A parallel version of this figure for REF_D2 is shown in Figure 4. All DLM changes are shown relative to the year 2000 to
619 allow for easy comparison between changes since that time.
620

Deleted: CCMI-1
Deleted: –
Deleted: CCMI-2022
Deleted: –
Deleted: trends
Deleted: trends



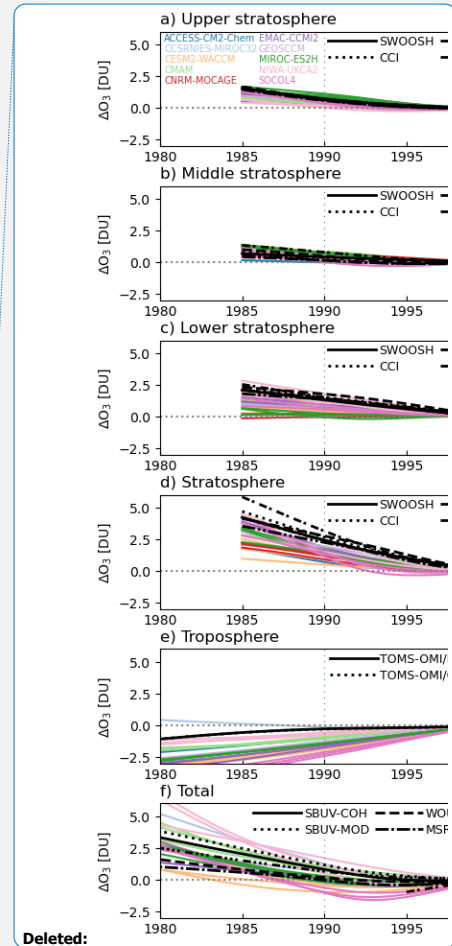
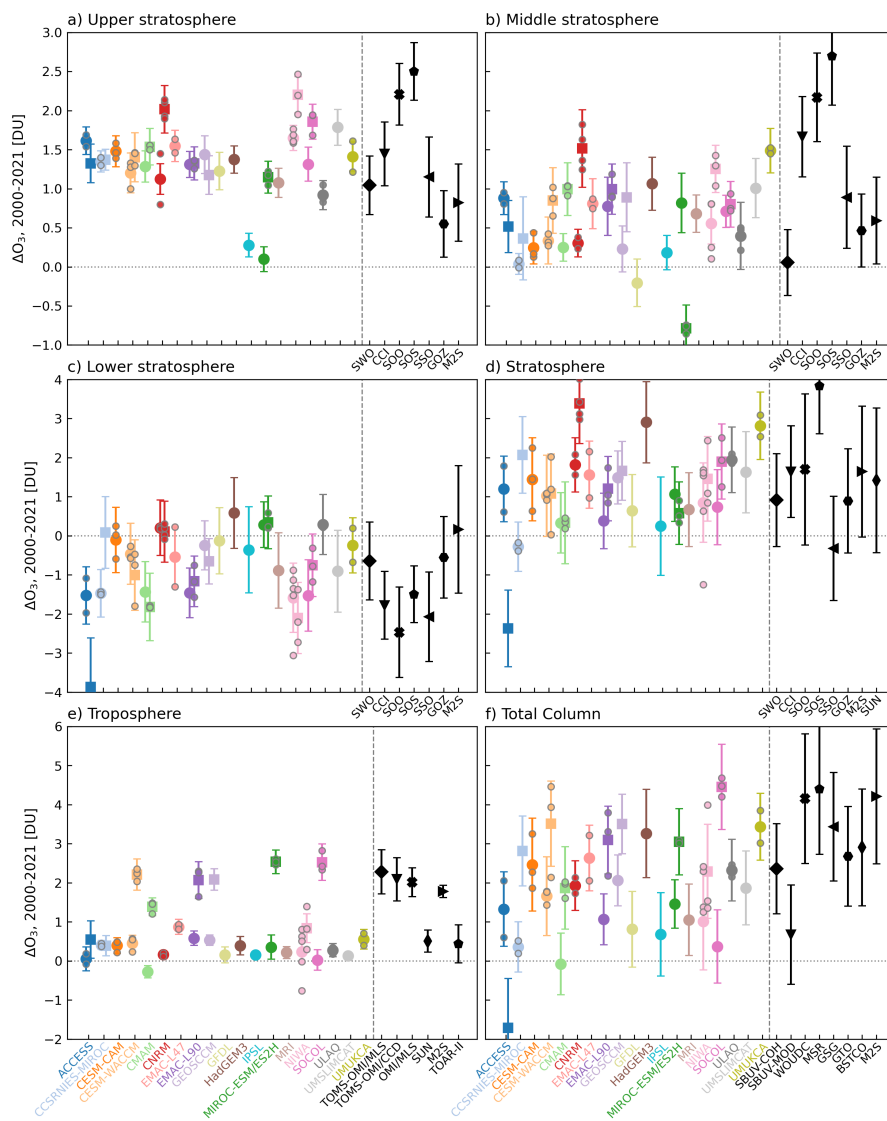


Deleted:

Formatted: Font: 9 pt, Bold

631 Figure 3: DLM background trend timeseries over the “observational period” (1985–2021) for REF-C2 simulations (colors, left
632 column), REF-D2 (colors, right column) and observations (black) relative to 2000 for (a) the upper stratospheric PCO, (b) middle
633 stratospheric PCO, (c) lower stratospheric PCO, (d) whole stratospheric PCO, (e) tropospheric PCO, and (f) the TCO. At the end
634 of the timeseries, a few selected examples are shown of the DLM background trend ozone change over 2000–2021, with error bars
635 denoting ± 1 standard deviation of the DLM changes (see text for discussion). For tropospheric PCO datasets starting in 2005 (i.e.,
636 SUNLIT, OMI/MLS, and M2S), the timeseries have been adjusted to match the TOMS-OMI/MLS value in 2005 for easy comparison
637 across datasets, and the changes have been scaled to represent the equivalent change over 2000–2021. Note, CNRM-MOCAGE data
638 are excluded from panels j and l due to an unphysical jump in tropospheric ozone near the end of simulation. To facilitate easy
639 comparison, the y-axes span 8 DU for all layers except for the middle and upper stratosphere, which span 4 DU.

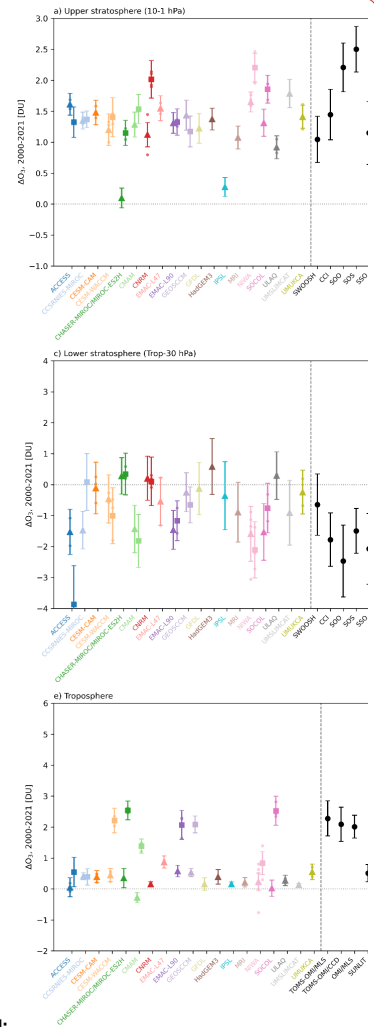
- Deleted: 1979/
- Deleted: CCMI-1
- Deleted: –
- Deleted: The right column shows
- Deleted: trends
- Deleted: trends



Deleted:

Figure 4: DLM background trends over the early recovery period (2000–2021) for each model REF-C2 simulations (triangles), REF-D2 (squares) and observations (black) relative to 2000 for (a) the upper stratospheric PCO, (b) middle stratospheric PCO, (c) lower stratospheric PCO, (d) whole stratospheric PCO, (e) tropospheric PCO, and (f) the TCO. At the end of the timeseries, a few selected examples are shown of The right column shows the DLM background trend ozone change over 2000–2021, with error bars denoting ± 1 standard deviation of the DLM trendschanges (see text for discussion). For tropospheric PCO datasets starting in 2005 (i.e., SUNLIT, OMI/MLS, and M2S), the timeseries have been adjusted to match the TOMS-OMI/MLS value in 2005 for easy comparison across datasets, and the changes have been scaled to represent the equivalent change over 2000–2021. Note, CNRM-MOCAGE data are excluded from panels j and l due to an unphysical jump in tropospheric ozone near the end of simulation. To facilitate easy comparison, the y-axes span 8 DU for all layers except for the middle and upper stratosphere, which span 4 DU.

To complement these figures, Figure 5 shows kernel density estimates (KDEs) of changes in the DLM background ozone trend since 2000, from both multi-model experiments and observations. To quantify the degree of agreement between the trend estimates, we compute the overlap coefficients between pairs of trend distributions (Inman and Bradley, 1989). The overlap coefficient is defined as the area intersected by two probability density functions, and is computed here by numerically integrating the minimum of the two KDEs. The overlap of the two KDEs is then expressed as a percentage, with 0% indicating distributions with no overlap and 100% indicating identical distributions. Overlap coefficients are discussed below in the context of specific comparisons between models and observations. A comprehensive set of overlap coefficients for all combinations of models, observations, and atmospheric layers are provided in the Supplement (Figs. S1–S6), alongside a table of the mean DLM changes (Table S1). Together, these figures demonstrate a varied level of agreement amongst the models, amongst the observations, and between models and observations across the different vertical layers of the atmosphere. Below, we summarize some of the key the similarities and differences between the models and observations, and highlight the notable differences between the REF-C2 and REF-D2 experiments.



Deleted:

Formatted: Font: 9 pt, Bold

Deleted: trends

Formatted: Font: Not Bold

Deleted: Same as Fig. 3, but for CCMI-2022 REF-D2 simulations. Note, CNRM-MOCAGE data are excluded from panels e–f due to an unphysical jump in tropospheric ozone near the end of simulation.

Deleted: –

Deleted:

Deleted: KDEs are a non-parametric method used to estimate the probability density function (PDF) of a random variable by applying a smoothing kernel, and by construction are normalized to unity. For each observational dataset, KDEs are computed from the 5000 iterations of the DLM model. Multi-model KDEs from CCMI-1 and CCMI-2022 are constructed by sampling 5000 DLM iterations across all available realizations of each CCM participating in ... [1]

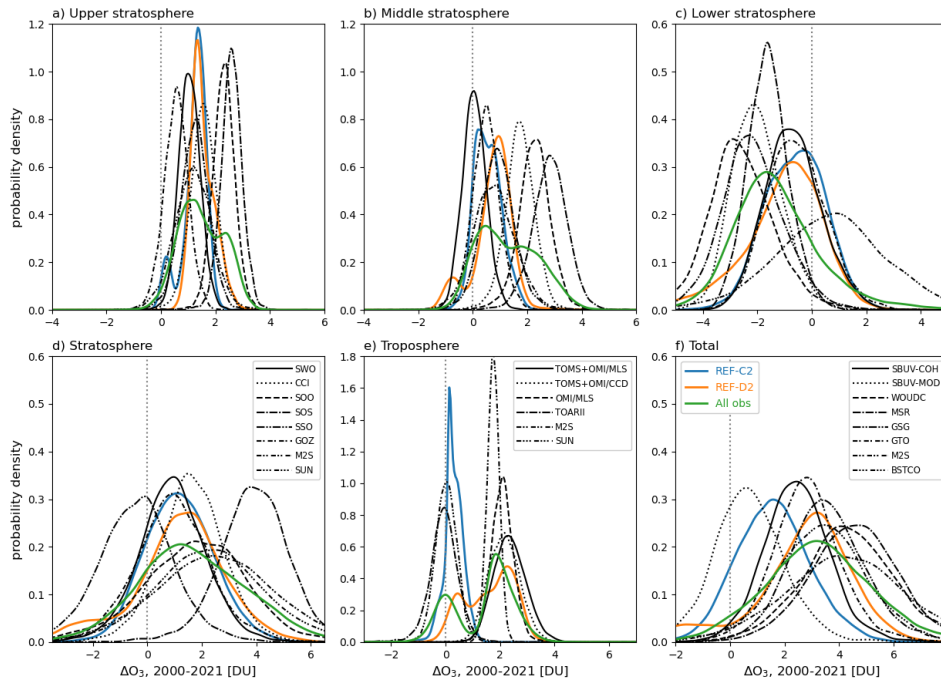
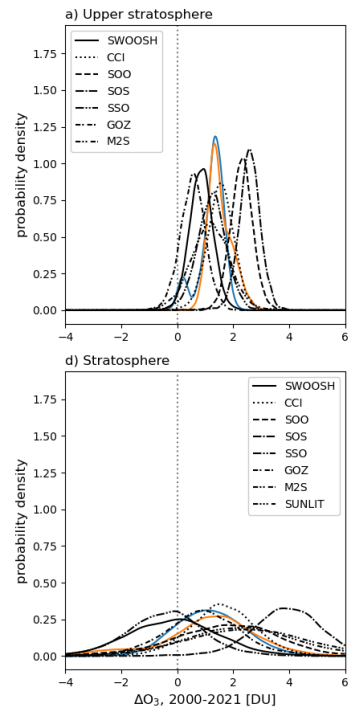


Figure 5: Kernel density estimates (KDEs) of DLM background trend change (2000–2021) in tropical ozone columns from observations (black lines for individual data sets, green lines for the multi-observational distributions), REF-C2 (blue), and REF-D2 (orange). Results from observational datasets starting in 2005 (i.e., OMI/MLS, SUNLIT, M2S) have been scaled to represent changes over the 2000–2021 period instead of the 2005–2021 period.

5.1 Upper stratosphere

In the upper stratosphere (10–1 hPa), the spread of CCM responses is fairly narrow and there is a large degree of overlap between REF-C2 (+1.2 DU over the 2000–2021 period, Fig 5a) and REF-D2 experiments (+1.5 DU). However, there is a wide spread amongst the observational datasets since 2000, with some datasets showing quite strong recovery values that are generally above the spread of REF-C2 models (SOS, mean=+2.6 DU; SOO, mean=+2.3 DU), some near the middle of the model spread (M2-SCREAM, mean=+1.2 DU; SSO, mean=+1.2 DU; CCI, mean=+1.5 DU), and the MLS-based datasets on the lower end (SWOOSH, mean=+1.0 DU; GOZCARDS, mean=+0.6 DU).



Deleted:

Deleted: CCMI-1

Deleted: Together, these figures demonstrate a varied level of agreement amongst the models, amongst the observations, and between models and observations across the different vertical layers of the atmosphere. Below, we summarize the similarities and differences between the models and observations, and highlight the notable differences between the CCMI-1 and CCMI-2022 experiments. ¶

Deleted: CCMI

763 For the tropical upper stratosphere changes over 2000–2021, the observational datasets with the least (GOZCARDS, mean
764 change of +0.55 DU) and most (SOS, mean change of +2.6 DU) upper stratospheric ozone increase have very little overlap
765 with one another (2%). In the middle lies the CCI dataset, which overlaps somewhat with each of these observational datasets
766 (28% for GOZCARDS and 21% for SOS). In contrast to the spread amongst observational estimates, the REF-C2 and REF-
767 D2 upper stratospheric changes since 2000 are in good agreement with one another, showing an increase of 1.3 and 1.5 DU,
768 respectively, with an overlap of 85%. Although there is a large spread of observational estimates, the CCI data shows good
769 agreement with the models (75% and 79% overlap for REF-C2 and REF-D2, respectively), as does M2-SCREAM although to
770 a lesser extent (67% and 59% overlap).

771 **5.2 Middle stratosphere**

772 In the middle stratosphere (30–10 hPa), the model experiments are also in good agreement, with REF-C2 (REF-D2) showing
773 an increase of 0.7 DU (0.5 DU) since 2000 (81% overlap). Similarly to the upper stratosphere, the observations show a wide
774 range, with SWOOSH showing a small positive change (mean of +0.1 DU), SOS showing a strongly positive change (mean
775 increase of 2.5 DU), and a negligible 2% overlap between the two. Even neglecting these two outlier datasets, the range of
776 observational estimates in the middle stratosphere is large. For example, the next most negative dataset (GOZCARDS) has a
777 +0.5 DU increase, whereas the next most positive dataset (SOO) contains a +2.3 DU increase, with only 10% overlap between
778 the two. As for the agreement between models and observations, in this region both M2-SCREAM and GOZCARDS have the
779 best overlap with the models (78%/79% between REF-C2/REF-D2 and M2-SCREAM; 86%/72% for GOZCARDS).
780 Given that GOZCARDS, M2-SCREAM, and SWOOSH are all primarily reflecting Aura MLS data after 2004, it is somewhat
781 surprising that they give different results. A sensitivity analysis where the DLM was provided SWOOSH and GOZCARDS
782 data starting in 2005 shows nearly identical 2005–2021 changes between the two datasets (95% overlap), in contrast to changes
783 over the same time period (2005–2021) but with DLM input data starting in 1985 as in Fig. 5 (71% overlap). Also using 2005–
784 2021 data as input to the DLM, we find that SWOOSH and GOZCARDS show good agreement with M2-SCREAM (75% and
785 78% overlap, respectively). Together, these results indicate that trend differences in Fig. 5b over 2000–2021 are not a result
786 of using different versions of Aura MLS data (SWOOSH uses v5, whereas GOZCARDS and M2-SCREAM use v4.2). An
787 additional test using DLM input data starting in 2000 produces a 78% overlap between SWOOSH and GOZCARDS for 2000–
788 2021 changes, as compared to the 63% overlap in Fig. 5b (based on input data starting in 1985). This suggests that the trend
789 differences between SWOOSH and GOZCARDS in Fig. 5b are due to both underlying data differences during the 2000–2004
790 pre-Aura MLS time period, as well as to differences in pre-2000 data impacting the background trend value in the year 2000.
791 As SWOOSH and GOZCARDS use the same satellite data and data versions in the pre-Aura period, the exact cause of these
792 differences is presumably due to different processing/filtering of the input profiles or merging methodology, and will be the
793 subject of future investigation.

Deleted: To quantify the degree of agreement between the trend estimates, we compute the overlap coefficients between pairs (... [2])

Deleted: CCMI-1

Deleted: -

Deleted: CCMI-2022

Deleted: -

Deleted:

Deleted: trends

Deleted: CCMI

Deleted: -

Deleted: 1

Deleted: CCMI

Deleted: -202

Deleted: 2

Deleted: -

Deleted: CCMI-1

Deleted: -

Deleted: CCMI-2022

Deleted: -

Deleted: negative

Deleted: -0.5

Deleted: -

Deleted: CCMI

Deleted: -

Deleted: 1

Deleted: CCMI

Deleted: -

Deleted: 202

Deleted: -

Deleted: -

Deleted: trends

Deleted: trends

Deleted: 44

Deleted: -

Deleted: ERRA

Deleted: -

Deleted: trends

Deleted: 32

Deleted: -

Deleted: -

842 5.3 Lower stratosphere

843 In the lower stratosphere (tropopause to 30 hPa), the CCMs show negative mean **changes** over 2000–2021 (-0.6 DU for **REF-C2** and -1.1 DU for **REF-D2**) and significant overlap between the two distributions (88%). The CCM distributions are 69%
844 and 77% negative for REF-C2 and REF-D2, respectively. All of the observations except for M2-SCREAM (+0.7 DU) show
845 negative mean **changes** over this time period, with values ranging from -2.6 DU (SOO) to -0.6 DU (GOZCARDS). Although
846 the models and observations mostly show robust negative **changes**, we note that the percentage of posteriors with negative
847 values ranges from insignificant (71% for GOZCARDS) to highly significant (>95% for SSO, SOO, SOS, and CCI). Finally,
848 as with the other regions of the stratosphere, in the lower stratosphere the degree of overlap between the observations and
849 models varies widely, from about 40% (SOO and **REF-C2**) to 95% (GOZCARDS and **REF-C2**).
850

851 5.4 Whole stratosphere

852 Considering the stratosphere as a whole, both observational estimates and CCMs predominately show increases over the 2000–
853 2021 time period. As with the other regions of the stratosphere, the CCM ensembles show a large degree of overlap (90%),
854 even though their mean increases are somewhat different (1.4 DU for **REF-C2** versus 1.0 DU for **REF-D2**). It is worth noting
855 that the ACCESS-CM2-Chem is an outlier in the stratospheric column timeseries (Fig. 4d). Excluding this model, the mean
856 **REF-D2** change is +1.5 DU, and the two distributions have 98% overlap. Observations also span a wide range, varying from
857 -0.4 DU for SSO to +3.9 DU for SOS with 11% overlap between the two. The observational estimates closest to the CCM
858 means (i.e., GOZ, +0.9 DU; SOO, +1.7 DU) show quite strong overlap with one another (72%). The overlap between CCMs
859 and observations ranges from ~30–50% for the SOS and SSO datasets, respectively, to ~70-90% for the other observational
860 datasets (Fig. S4).

861 5.5 Troposphere

862 In the troposphere, there are stark differences between the observational datasets and also between the two CCM experiments.
863 For example, **REF-C2** shows a relatively small change in tropical tropospheric ozone over the 2000–2021 period (+0.4 DU),
864 whereas the overall **REF-D2** distribution shows substantial increases (+1.4 DU). With only 40% overlap between the two CCM
865 distributions, the troposphere is the region of greatest (and arguably, only substantial) difference between the two modeling
866 experiments. Although we do not attempt to quantitatively attribute the differences in tropospheric ozone **changes** between
867 REF-C2 and REF-D2, we do note that the tropospheric precursor emissions are very different between the two experiments,
868 and that this seems to be the most likely contributor to differences among the datasets (Acquah et al., 2025). For example,
869 from 2000 to 2020 the REF-C2 tropical-average NO_x emissions increased from ~30 Tg year⁻¹ to ~40 Tg year⁻¹, whereas in
870 REF-D2 the emissions rose from ~50 Tg year⁻¹ to ~70 Tg year⁻¹, almost a factor of two greater increase. Our analysis cannot
871 rule out, however, that the differences in tropical tropospheric ozone evolution between REF-C2 and REF-D2 also have a
872 contribution due to changes in the complexity of tropospheric ozone chemistry in the newer generation of models, with a larger

Deleted: trends

Deleted: CCMI

Deleted: –

Deleted: 1

Deleted: CCMI

Deleted: –

Deleted: 202

Deleted: trends

Deleted: trends

Deleted: CCMI

Deleted: 1

Deleted: CCMI

Deleted: 1

Deleted: CCMI

Deleted: 1

Deleted: CCMI

Deleted: 202

Deleted: CCMI

Deleted: 202

Deleted: CCMI

Deleted: 1

Deleted: CCMI

Deleted: 202

Deleted: trends

897 proportion of the earlier REF-C2 models (e.g. UМУKCA-UCAM, HadGEM3-ES, CMAM) using chemistry schemes tailored
898 primarily for stratospheric ozone studies while employing a reduced-complexity approach to tropospheric chemistry.
899 Similarly large discrepancies exist between the observational estimates of tropospheric change. For example, the SUNLIT
900 dataset shows a 0.0 DU mean change, whereas the TOMS+OMI/MLS dataset shows an increase of 2.4 DU (with 1% overlap
901 between the two). We note that the SUNLIT dataset shows strong overlap with the REF-C2 distribution (64%), whereas the
902 TOMS+OMI/MLS dataset overlaps strongly with the REF-D2 distribution (65%).

903 **5.6 TCO**

904 Finally, for TCO we find relatively minor differences between the CCM distributions but a wide range of observational changes
905 over the 2000–2021 period. The CCM TCO changes are +1.7 DU and +2.2 DU for REF-C2 and REF-D2, respectively, with
906 72% overlap. It is worth noting that the ~0.5 DU difference between the two CCM experiments increases to over 1.5 DU if the
907 outlier model ACCESS-CM2-Chem is excluded (see Fig 4f), and this difference is virtually identical to the difference between
908 the tropospheric column estimates from the two CCM experiments with ACCESS-CM2-Chem excluded. This suggests that
909 the differences in the tropical TCO responses between the two generations of CCM experiments are largely due to differences
910 in the tropospheric ozone column changes occurring in the last ~20 years.

911 **6. 21st century tropical ozone changes**

912 In this section, we consider changes in tropical ozone from DLM fits over the entirety of the REF-C2/REF-D2 simulations
913 covering 1960–2100, but with a focus on 21st century behavior, the beginning of which roughly corresponds to the start of
914 ODS-driven chemical ozone recovery. We note that while for the observational period (Section 5) the two scenarios should be
915 very close, it is in the second part of the 21st century where the differences should become more apparent given the differing
916 GHG emission scenarios used by the two experiments. The differences amount to an approximately 25% reduction in 21st
917 century CO₂ emissions in REF-D2 compared to REF-C2.
918 Figure 6 shows the DLM background trend timeseries in the various vertical layers in the stratosphere and troposphere, and
919 Figure 7 shows kernel density estimates of the DLM background trend change over the 21st century (2000–2094 for REF-C2
920 to cover the time period common to all models), similar to Fig. 5. These plots reveal generally consistent patterns of ozone
921 changes across the two experiments and across models, but with a few notable exceptions discussed below.

Deleted: CCM1

Deleted: 1

Deleted: CCM1

Deleted: 1

Deleted: CCM1

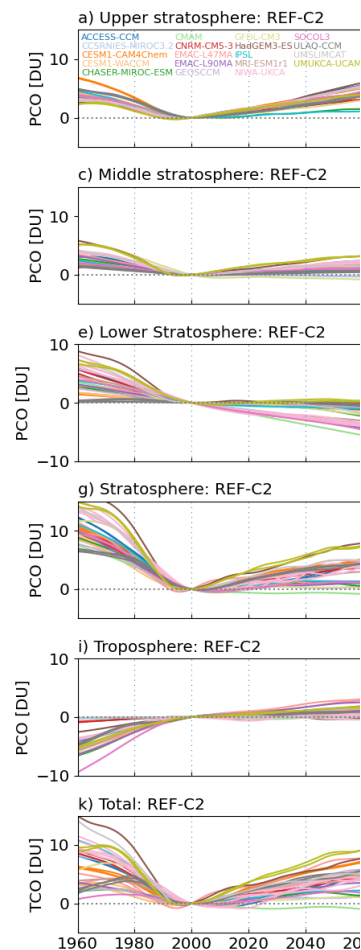
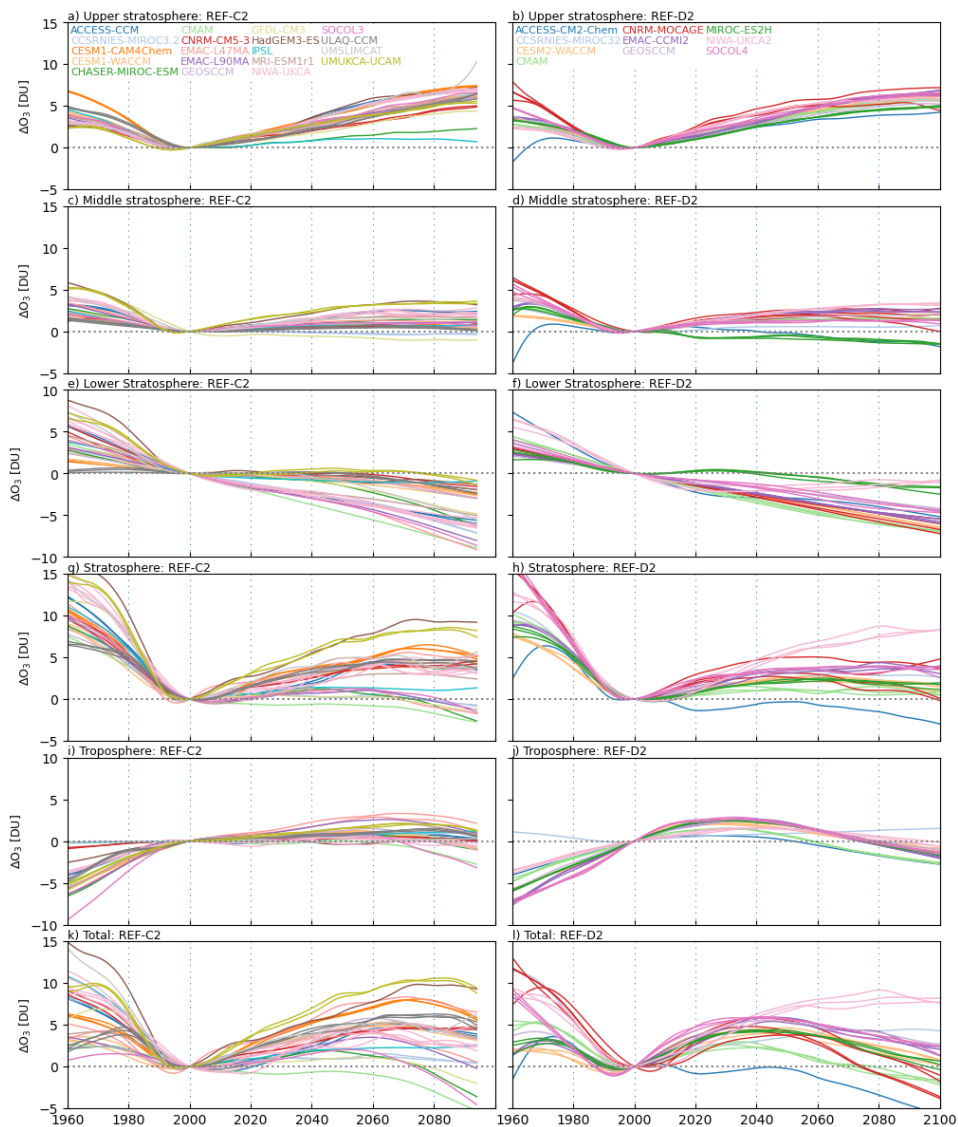
Deleted: 202

Deleted: CCM1

Deleted: 1

Deleted: CCM1

Deleted: 202



Deleted:

Figure 6: DLM background values over the “full period” (1960–2100) for REF-C2 simulations (left column) and REF-D2 (right column), relative to 2000 for (a) the upper stratospheric PCO, (b) middle stratospheric PCO, (c) lower stratospheric PCO, (d) whole stratospheric PCO, (e) tropospheric PCO, and (f) the TCO.

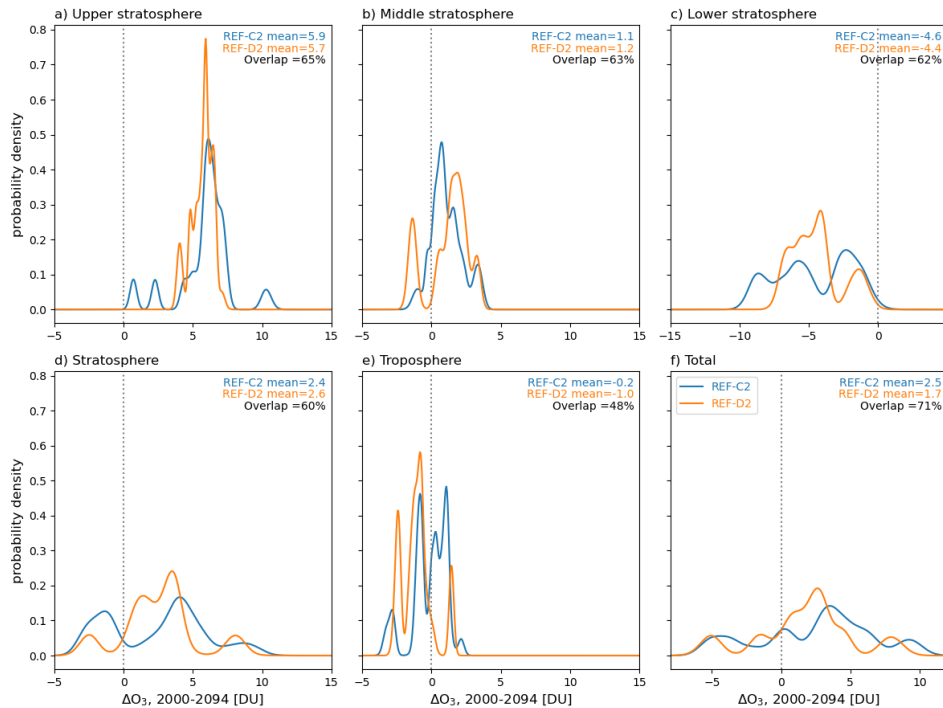


Figure 7: Kernel density estimates (KDEs) of DLM-based tropical ozone column change over 2000–2094 from REF-C2 (blue) and REF-D2 (orange) model simulations. The multi-model mean changes and overlap between the two experiments are indicated in each panel.

In the upper and middle stratosphere, ozone recovery is quite similar between the two experiments over the 21st century, with the CCMI experiments showing 63% and 58% overlap, respectively, and the bulk of ozone recovery (~ 6 DU) coming from the upper stratospheric layer, as expected. In the tropical lower stratosphere, models robustly show an ozone decline throughout the 21st century, with a mean decrease of ~5 DU and virtually no models showing positive changes. There is good overlap

Deleted: trends
Deleted: CCMI-1
Deleted: –
Deleted: CCMI-2022

Deleted: CCMI-1
Deleted: –
Deleted: CCMI-2022
Deleted: –
Deleted: CCMI

Deleted: trends

957 between the two CCMI experiments (59%), but with a substantial spread among models within each experiment ranging from
958 almost -10 DU to nearly zero.

959 Considering the stratosphere as a whole, the mean 21st century tropical ozone change is positive for both CCMI experiments
960 (~2–3 DU), but with substantial spread within and among the experiments. Notably, 6 REF-C2 models (CCSRNIES-
961 MIROC3.2, CHASER-MIROC-ESM, CMAM, EMAC-L47MA, EMAC-L90MA, SOCOL3) and one REF-D2 model
962 (ACCESS-CM2-Chem) have mean declines in stratospheric PCO over the 21st century. Of the REF-C2 models with negative
963 changes, 5 of them comprise the most negative lower stratospheric PCO changes and range from -9 to -7 DU. The sixth model
964 (CHASER-MIROC-ESM) has a lower stratospheric trend (-5.4 DU) close to the REF-C2 multi-model mean response (-4.9
965 DU), but has a significantly less positive upper stratospheric recovery (+2.2 DU, compared to the CCMI-1 mean of +6.3 DU)
966 that contributes to its negative PCO trend for the whole stratosphere.

967 In contrast to REF-C2, where the models with negative stratospheric changes are largely those with the most negative lower
968 stratospheric changes, the outlying negative stratospheric trend in the REF-D2 ACCESS-CM2-Chem simulation (Fig. 6h) is
969 primarily due to its outlying negative trend in the middle stratosphere (Fig. 6d) which is not fully offset by its relatively weak
970 positive trend in the upper stratosphere (Fig. 6b). Specifically, its combined middle/upper stratosphere trend is +2.6 DU,
971 compared with the REF-D2 multi-model mean of +7.2 DU. In contrast, the ACCESS lower stratosphere trend is -4.9 DU,
972 which is near the REF-D2 multi-model mean of -4.6 DU.

973 Over the 21st century, the multi-model mean tropospheric column ozone changes are different between the two model
974 experiments, with a slightly negative trend in REF-D2 (-1.1 DU) and almost no decrease in REF-C2 (-0.2 DU). These results
975 are opposite of the response over the first two decades of the 21st century, where REF-D2 shows an increase (+1.6 DU) and
976 REF-C2 shows relatively little change (+0.3 DU).

Deleted: CCMI

Deleted: CCMI

Deleted: trends

Formatted: Superscript

Deleted: CCMI

Formatted: Superscript

977 7. Can observational period changes predict future changes?

978 In this section and the next one, we examine the degree to which tropical stratospheric ozone trends vary systematically across
979 different time periods in models and observations, the extent to which these relationships might serve as an emergent constraint
980 on model projections, and the feasibility and limitations of using existing observations to provide guidance on the credibility
981 of model projections for a given emissions scenario. The motivating hypothesis for analyzing model ozone trends across time
982 periods is that biases in ozone-controlling processes may lead some models to have stronger (or weaker) ozone trends that end
983 up being correlated over different time periods. For example, a model that accelerates the BDC too strongly in response to
984 GHG increases may show a too-negative lower stratospheric ozone trend during both the 20th and 21st centuries. Or in the
985 upper stratosphere, a model that is overly sensitive to chlorine changes might show both a too-rapid depletion of ozone in the
986 late 20th century and a too-rapid recovery in the 21st century.

987 For this analysis, we consider three time periods: the depletion period (before the turnaround, which occurs around 2000), the
988 early 21st century recovery period (turnaround–2021), and the remainder of the 21st century (2022 and beyond). The key

Deleted: test

Formatted: Superscript

Deleted: over

question, we seek to answer in this section is whether or not the simulated later 21st century trends for a given scenario are correlated with modeled trends over the earlier periods encompassing available observations (~1980–2021). Before doing that, we first test whether trends computed from input over short (one to two decade) time periods can accurately quantify the “true” model trends (i.e., those based on using the entire model time series as input to the DLM), as this is a necessary precondition for having confidence in any comparison between observations and models. Then, we explore inter-model correlations using different pairs of time periods to establish 1) how ozone depletion and early recovery trends covary across models, and 2) how late 21st century model trends covary with early 21st century recovery. The results from the latter of these tests establish the extent to which DLM trends from observations might provide some constraint on future model ozone projections. Then, in the following Section 8, we use the observational trends from Section 5 to assess the current ability to provide an observational constraint on 21st century tropical ozone trends.

7.1 Diagnosing trends over the observational period

In this section, we consider how well trends can be characterized over shorter time periods, such as the first few decades of the 21st century (comprising the early recovery period, i.e. turnaround to 2021). We address this issue by comparing 2000–2021 DLM trends from runs where the DLM is fed CCM output only from the observational time period (1985–2021) with 2000–2021 DLM trends from runs where the DLM uses data from the full time period of the model simulations (1960–2100). The latter of these is assumed to represent the “true” model trend that is unaffected by the limited sampling and endpoint issues that could potentially impact a shorter observational data record of only a few decades. Figure 8 shows the correlation between trends computed over the early 21st century recovery time period, with the only difference being the length of the data used as input to the DLM. As can be seen from this figure, the correlation between the two versions of the DLM are high for all regions of the stratosphere and for both multi-model experiments, and the models lie along the 1:1 line as expected. We note that excluding outliers (i.e., CHASER-MIROC32 in the middle stratosphere, ACCESS-CM2-Chem in the lower stratosphere and whole stratosphere) does not significantly alter the correlation coefficients. These results suggest that the true long-term trends over relatively short periods such as two decades can be accurately quantified from data records of limited length, and therefore lends confidence to the comparison of trends between observational datasets (whose data length is necessarily limited) and those from models.

Deleted: s

Deleted: here are: 1)

Deleted: “predictable” from

Deleted: observational

Deleted: , and 2) whether this potential model “predictability” of 21st century trends can be constrained by existing observations

Deleted: Below

Deleted: Finally

Deleted: considering

Formatted: Superscript

Deleted: sensitivity simulations

Deleted: model

Deleted: over

Deleted: , as opposed

Deleted: to

Deleted: over

Deleted: very

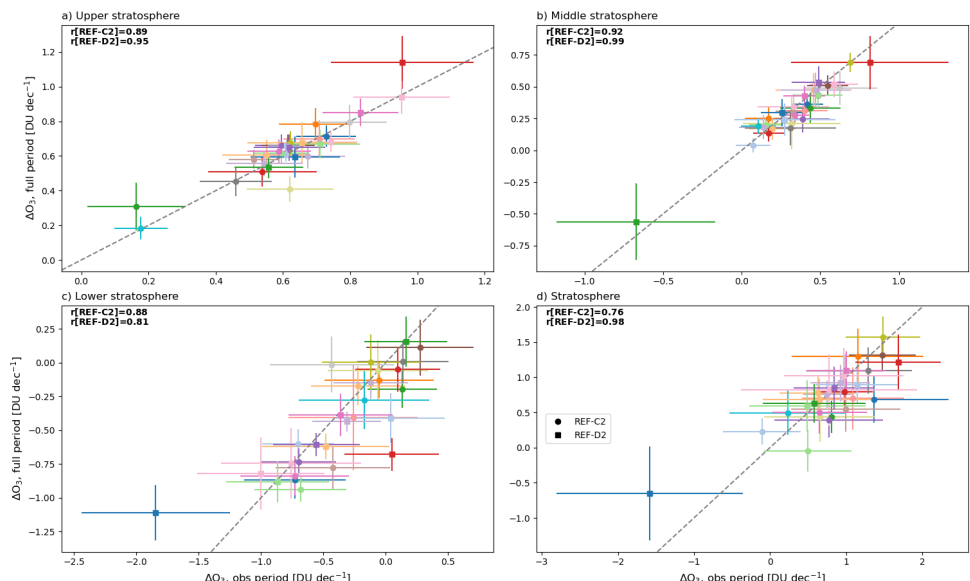


Figure 8. Scatter plots of model tropical partial column ozone trends (2000–2021 for lower stratosphere, turnaround–2021 for all others) for the DLM run with data from the full time period (1960–2100, y-axes), versus for the DLM run with data only over the observational period (1985–2021). Each marker shows the DLM-inferred trend from a given model, with different symbols for different model experiments (see legend) and different colors for each model (see legends in Figs. 4 and 6). Horizontal and vertical error bars represent the ± 1 standard deviation about the mean DLM change, as in previous figures. The inter-model correlation coefficients (all of which are significant) for each multi-model experiment are shown in each panel, as are the 1:1 lines (gray, dashed).

7.2 Depletion vs. recovery period changes

We next consider the correlation between ozone trends over the depletion and recovery periods across models, separated by different vertical regions of the stratosphere. For regions of the stratosphere with a well defined turnaround (minimum) in their PCO time series (i.e., upper stratosphere, middle stratosphere, and whole stratosphere), we define the time boundary between the depletion and recovery periods as the date of the minimum in the DLM background trend occurring between 1990 and 2015. This turnaround date is identified separately from each of the 5000 DLM samples, for each timeseries (i.e., each combination of model realization and vertical region). Since there is in general no turnaround in ozone trend in the lower stratosphere, for this region the year 2000 is used. The turnaround date can vary significantly across each model realization and vertical region of the stratosphere, as well as to some degree across each of the 5000 DLM samples. As an example, the

Deleted: (significant correlations are in bold)

Deleted:

1097 REF-C2 multi-model mean turnaround time is 1997 for the upper stratosphere (with range 1994–2010), but is delayed to 2002
1098 for the whole stratosphere (with range 1996–2029).
1099 Figure 9 shows the inter-model correlation of trends over the depletion (1980–turnaround) and early 21st century recovery
1100 periods (turnaround–2021). This plot shows significant inter-model correlations between the strength of depletion and recovery
1101 for the whole stratosphere across all model experiments. Perhaps surprisingly, in the upper stratosphere the correlation across
1102 the models between depletion and early recovery trends is not significant. One might expect that models with shorter CFC
1103 lifetimes have both a faster inorganic chlorine (Cl_y) increase before, and a faster Cl_y decrease after, the peak Cl_y in the late
1104 1990’s, and that this behavior might lead to those same models having both faster ozone depletion and faster ozone recovery.
1105 To test this hypothesis, we consider the Cl_y trend ratio, which we define as the ratio of the tropical mean (30° S–30° N) Cl_y
1106 recovery trend (from peak Cl_y year to 2021) to the tropical mean Cl_y depletion trend (1985 to peak Cl_y year) calculated using
1107 Cl_y partial column amounts over the vertical region of interest (e.g., upper stratosphere) from each REF-C2 experiment.
1108 Compared to the REF simulations, ozone trends in the fGHG simulation (triangles, Fig. 9a) are generally closest to the multi-
1109 model mean Cl_y trend ratio (dotted line, Fig. 9a). This general agreement between the fGHG ozone trend ratios and the Cl_y
1110 trend ratios in the upper stratosphere makes sense given the “purely chemical” perturbations in the fGHG simulations, and the
1111 absence of complicating factors such as GHG-induced cooling and its associated knock-on effects on ozone. Indeed, almost
1112 all REF-C2 and REF-D2 models show stronger ozone recovery than that suggested from the Cl_y ratio (i.e., they are above the
1113 line in Fig. 9a), qualitatively in line with the fact that in the recovery period GHG and ODS both act to increase ozone, whereas
1114 in the depletion period the GHG forcing acts to offset the ODS forcing in the upper stratosphere.

Deleted: ,

Deleted: ,

Formatted: Subscript

Formatted: Subscript

Formatted: Subscript

Formatted: Subscript

Deleted: recovery/depletion trend ratio for

Deleted: ,

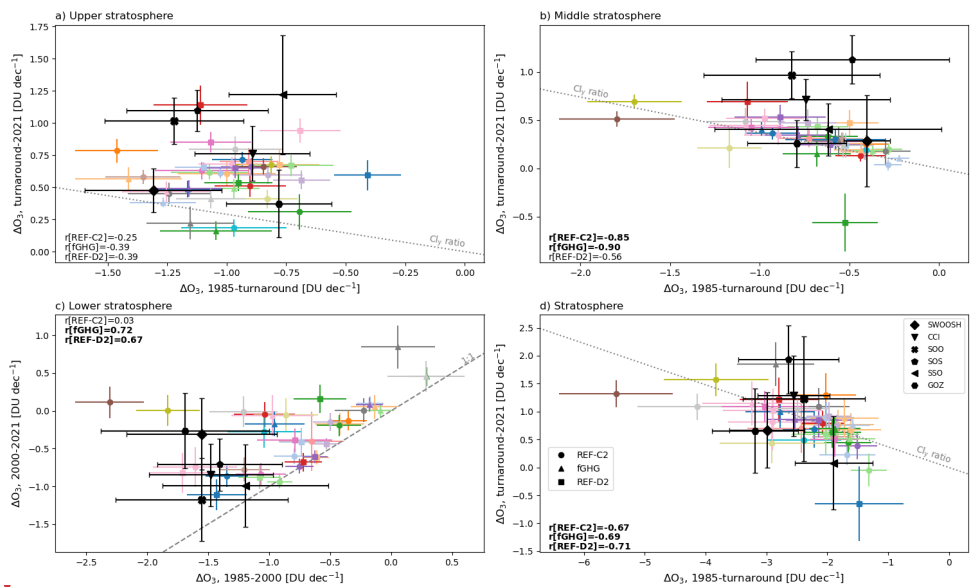
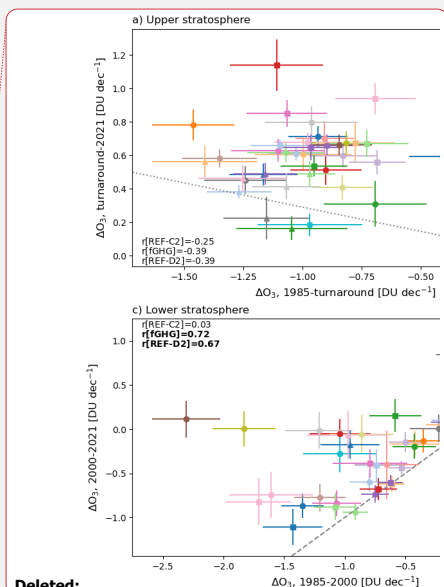


Figure 9. Scatter plots of the tropical partial column ozone depletion (horizontal axes) and early recovery (vertical axes) trends from models and observations. Each model marker shows the DLM-inferred trend from a single model realization, with different symbols for different model experiments (see legend) and different colors for each model (see legends in Figs. 4 and 6). Horizontal and vertical error bars indicate ± 1 standard deviation about the mean DLM change, as in previous figures. The inter-model correlation coefficients for each multi-model experiment are shown in each panel (significant correlations are in bold). The dotted gray lines in (a,b,d) show the multi-model mean value of the Cl_y recovery/depletion trend ratio calculated over the same latitude (30° S– 30° N) and vertical region from the REF-C2 experiment. The dashed gray line in (c) is the 1:1 line, indicating values for which the LS PCO trends are the same both before and after 2000.

In the middle stratosphere (Fig. 9b), there is a remarkably good correlation between ozone depletion and recovery rates for the two CCM1-1 experiments (i.e. REF-C2 and fGHG), whereby models with more rapid depletion also show more rapid recovery. The ozone recovery/depletion trend ratios in this region are in general agreement with the multi-model mean Cl_y ratio (gray dotted line, Fig. 9b), suggesting a rather straightforward connection between Cl_y changes and ozone changes in this region, and also potentially allowing for the use of an observational constraint on model projections. Although the REF-D2 correlation in Fig. 9b is not significant, we note that if the outlier model is removed (MIROC-ES2H), the correlation becomes significant for that experiment as well ($r=-0.70$). The MIROC-ES2H model exhibits a dry bias in stratospheric water vapor due to a cold bias in the tropical upper troposphere and lower stratosphere. Excessive NO_x and associated catalytic reactions in the middle



Deleted:

Deleted: the

1139 stratosphere lead to underestimated ozone, with differing ozone trends possibly linked to biases in nitrogen partitioning. It is
1140 possible that these biases may lead to the outlier behavior of this model in Fig. 9b.
1141 In the lower stratosphere (Fig. 9c), there is a significant positive correlation of ozone trends between the pre- and post-2000
1142 periods in the fGHG and REF-D2 experiments, in contrast to the negative correlations for other regions of the stratosphere.
1143 Given the predominately dynamical control over ozone in the lower stratosphere (with ozone changes driven predominantly
1144 by the GHG-induced acceleration of the BDC), this positive correlation is unsurprising and simply shows that, to some degree,
1145 models with stronger ozone trends in the first half of the observational time period tend to have stronger trends in the latter
1146 half (and vice versa for models with weak trends). This spread of model responses illustrates the variety of model BDC
1147 sensitivity to combined ODS/GHG forcing. Indeed, many of the model trends in this region fall near the 1:1 line (i.e., their
1148 trends are identical during each period; dashed line, Fig. 9c), although they are generally above the line, indicating relatively
1149 less negative trends during the early 21st century when ODS decreases tend to oppose BDC changes.
1150 Taken together, the full stratospheric PCO correlations are significantly negative for all experiments (Fig. 9d), as expected
1151 based on the dominant contribution from the middle stratosphere. This behavior suggests that overall, models that show
1152 stronger ozone declines during the depletion period tend to show stronger ozone increases in the few decades during the early
1153 recovery period.

1154 7.3 Observational period vs. late 21st century changes

1155 We next consider the inter-model correlation between trends for the beginning of ozone recovery (turnaround–2021) and the
1156 rest of the 21st century time period (2022–2094), based on DLM fits using data covering the whole 1960–2100 period as in the
1157 previous section (Fig. 10). At all levels, there is a significant inter-model correlation for REF-C2, whereby models with faster
1158 recovery in the first two decades also tend to show more positive trends over the rest of the 21st century. For the stratosphere
1159 as a whole (Fig. 10d), such correlation is not found for the other model experiments, a point which will be further discussed
1160 below. We also note that most models fall below the 1:1 line (dashed gray line, Fig. 10d), indicating that ozone trends in the
1161 later 21st century are mostly weaker than in the first two decades.

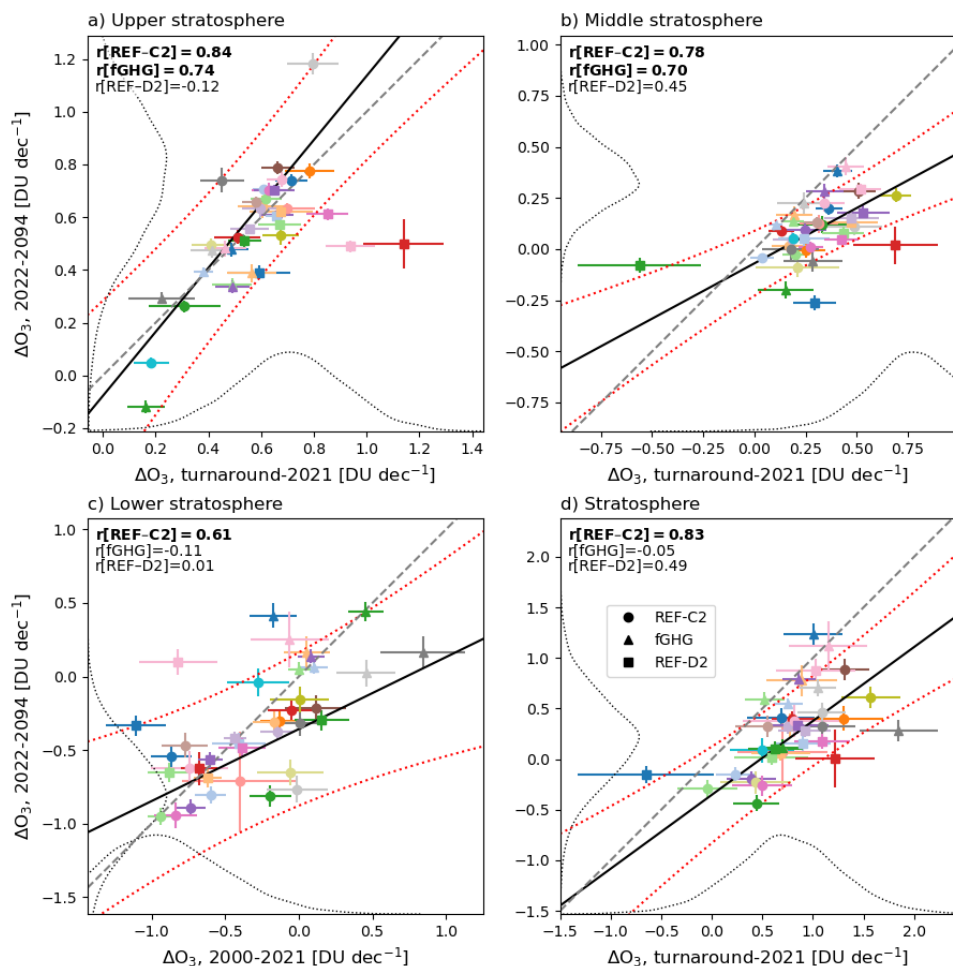


Figure 10. Scatter plots of the tropical partial column ozone early recovery (horizontal axes, 2000–2021) and later century (vertical axes, 2022–2094) trends. Each marker shows the DLM-inferred trend from a given model, with different symbols for different model experiments (see legend) and different colors for each model (see legends in Figs. 4 and 6). Horizontal and vertical error bars indicate the ± 1 standard deviation about the mean DLM change, as in previous figures. The intermodel correlation coefficients for each multi-model experiment are shown in each panel (significant correlations are in bold), as are the 1:1 lines (gray, dashed), the linear

1168 fit to the REF-C2 models (black lines), and their 95% prediction interval (red dotted lines). Dotted black lines show the CCI dataset's
1169 DLM posterior trend distribution (along the horizontal axes) and the CCI-predicted late 21st century trends (along the vertical axes).

1170

1171 There are notable differences between the two generations of CCMI REF experiments throughout the stratosphere (Fig. 10a-
1172 c), with REF-C2 showing significant correlations at all levels, in contrast to the insignificant correlations in REF-D2. While
1173 diagnosing the exact mechanisms driving the differing correlations is beyond the scope of this analysis, we found that the
1174 differences between REF-C2 and REF-D2 are still present even when considering only the 10 models that have output from
1175 both experiments. Thus, it seems most likely that the difference in behavior is due to differing forcings (e.g., GHG forcing is
1176 weaker in REF-D2), or different representations of processes that might change through the 21st century (e.g., the QBO, Richter
1177 et al., 2020), or as-of-yet unidentified model biases.

1178 In the middle and upper stratosphere, both ~~REF-C2~~ experiments produce a significant positive correlation between early and
1179 late 21st century ozone trends, with REF-D2 showing borderline significant correlation in the middle stratosphere but very
1180 little correlation in the upper stratosphere. We note that unlike in Fig. 9b, removal of the outlier model in Fig. 10b does not
1181 change the correlation coefficient significantly. In the lower stratosphere, there is also a large difference in the correlations
1182 between REF-C2 ($r=0.61$) and REF-D2 ($r=0.01$). These results are somewhat surprising, given the expected GHG-induced
1183 strengthening of tropical upwelling in chemistry climate models and its impact on lower stratospheric ozone, in contrast to the
1184 multiple dynamical, chemical, and radiative processes that might influence middle stratospheric ozone.

1185 In summary, in REF-C2 there is a strong positive correlation between model trends in the first two decades of the 21st century
1186 and the trends over the rest of the century in all layers of the stratosphere, whereas in REF-D2 there is a borderline significant
1187 correlation in the middle and whole stratosphere but insignificant correlations in the other layers. Except in the upper
1188 stratosphere, the model responses mostly lie below the 1:1 line, indicating that the trends over the later 21st century are generally
1189 smaller in magnitude than the first two decades of the century. At least for REF-C2, the early 21st century model ozone trends
1190 in the middle stratosphere appear to provide some “predictive” power for trends later in the century in that models with weak
1191 early-century trends also have weak late-century trends, and vice versa for models with strong trends.

1192 8 Testing an observational constraint on future changes

1193 Although a clear “emergent constraint” on end-of-century ozone trends is not robust in ~~REF-D2~~ simulations, the robustness of
1194 the correlation between early 21st century and end-of-century ozone trends in ~~REF-C2~~ models (particularly in the middle
1195 stratosphere) motivates further investigation of whether this relationship might be used in conjunction with early 21st century
1196 observations to provide an “observational constraint” on end-of-century values.

1197 Figure 10 illustrates the application of an observational constraint for one dataset. The probability distributions along the
1198 horizontal axes of that figure show the DLM-inferred posterior trends from the CCI dataset (as in Fig. 5). This distribution is
1199 then convolved with the best-fit line (black line in Fig. 10) and its prediction interval (red dotted lines, Fig. 10) using a Monte

Deleted: CCMI-1

Deleted: CCMI

Deleted: 202

Deleted: CCMI

Deleted: 1

1205 Carlo sampling technique (see, e.g., Fig. 2, Nowack et al., 2023) to produce a CCI-informed late century trend on the vertical
1206 axis. The CCI-informed trend distribution can then be compared with the simulated end-of-century trends. In principle, if one
1207 believed the observations as “truth” and had strong correlations between early and late century model trends, one could
1208 interpret the difference between the CCI-informed late century trend and the multi-model trend as an indication of the multi-
1209 model bias for a given emissions pathway. However, the uncertainty of the observationally-informed late century trends (i.e.,
1210 width of the distributions along the y axes in Fig. 10), which include contributions from both observational trend uncertainty
1211 (i.e., width of the distributions along the x axes in Fig. 10) and imperfect inter-model correlations (i.e., through the prediction
1212 interval), are in some cases (e.g., the upper stratosphere) too large to make such a straightforward assessment, even for a single
1213 observational data set. In this particular example, the CCI trend distribution agrees rather well with the whole-stratosphere
1214 trends from the models, but this agreement appears to be due to the models showing less positive trends than CCI in the
1215 middle/upper stratosphere (e.g., compare pdf along horizontal axis with model symbols in Fig. 10 a-b) that are compensated
1216 for by more positive trends in the lower stratosphere.

1217 To evaluate how the available observations might agree (or not) with end-of-century values, we consider the distribution of
1218 observation-informed end-of-century values from a range of different data sources, using the method outlined above. Figure
1219 11 shows the distribution of end-of-century trends based on the observations and the REF-C2 relationship, as well as the
1220 simulated trends. As expected, the spread amongst observational estimates (e.g., as seen in Fig. 5) is preserved in the (linear)
1221 transfer from early- to late-21st century trends. In the upper stratosphere, the widths of the observational distributions are
1222 generally greater than those of the models, whereas in the middle stratosphere, observations and model distribution widths are
1223 similar, suggesting that observationally-informed end-of-century trends are sufficiently precise to provide an observational
1224 constraint on the modeled trends in this region. However, in both regions systematic differences between the observational
1225 pdfs span a range that is larger than the width of the model distributions, and the conclusions one would draw about the
1226 accuracy of the model trends depends strongly on which observational estimate one regards as truth. For example, based on
1227 the SWOOSH data, one would conclude that the models systematically overestimate ozone trends in the middle and upper
1228 stratosphere, whereas one would conclude the opposite based on the SOO and SOS trends.

Deleted: the CCI-predicted end-of-century trends

Deleted: predicted

Deleted: Qualitatively

Deleted: predicted

Deleted: predicted values

Deleted: I

Deleted: and upper

Deleted: the

Deleted: s

Deleted: show generally

Deleted: widths

Deleted: predicted

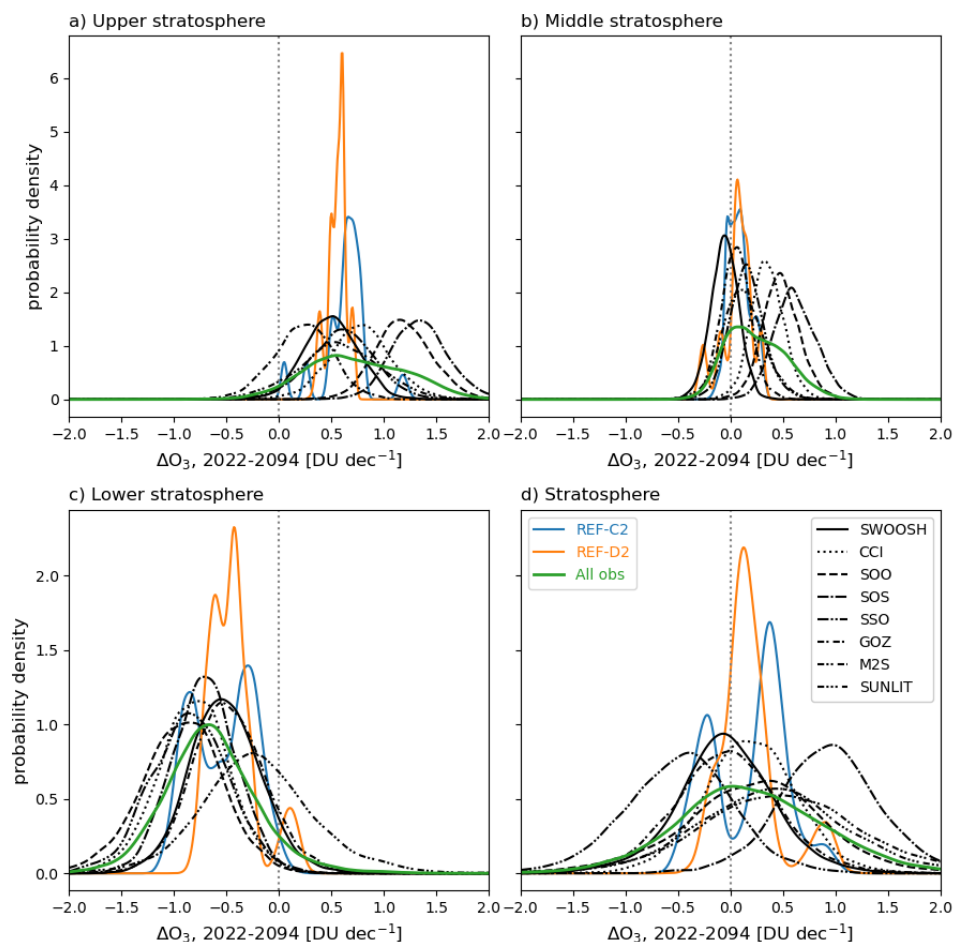
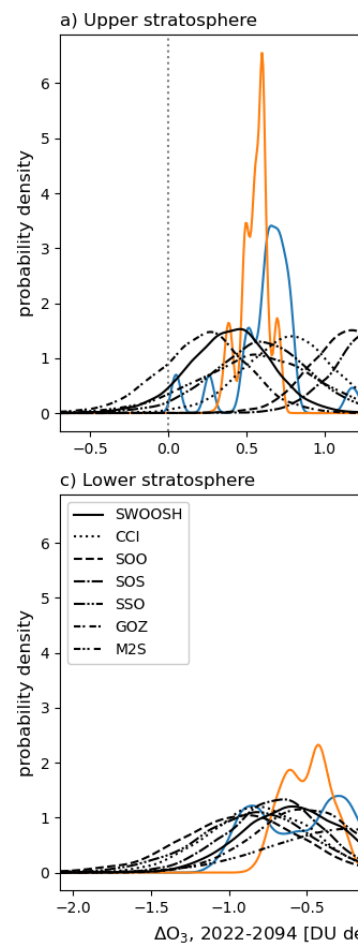


Figure 11: Kernel density estimates (KDEs) of tropical ozone column trends for 2022-2094 from REF-C2 (blue) and REF-D2 (orange) model simulations, as well as “observationally-informed” end-of-century trends constructed by combining observational trends over the early 21st century (2000/turnaround-2021, see Fig. 5) with the linear relationship between later and early century trends from REF-C2 in Fig. 10.



Deleted:

Deleted: CCMI-1

1256 In contrast to the middle and upper stratosphere, for the lower (and whole) stratosphere the observationally-based distributions
1257 largely overlap one another but have widths that are significantly wider than the model distributions (Fig. 11 c-d). The causes
1258 of the broader width of the observational distributions in the lower/whole stratosphere (Fig. 11 c-d) compared to the
1259 middle/upper stratosphere (Fig. 11 a-b) are both 1) wider prediction intervals (see Fig. 10) and wider observation-period trend
1260 distributions (see Fig. 5). Due to these wider distributions, observationally ~~informed~~ end-of-century trends for the lower/whole
1261 stratosphere are limited in their ability to provide an observational constraint on the simulated trends, although most of the
1262 observational datasets would suggest that the models underestimate end-of-century trends ~~in the lower stratosphere~~.
1263 In the end, it is not possible to confidently apply an observational constraint on end-of-century ozone trends, although this
1264 analysis sheds some light on the potential for and current limitations of such an approach. First among these limitations is the
1265 wide spread amongst the observational estimates, which at some levels poorly overlap with one another and should motivate
1266 future studies aimed at reconciling these differences. Additionally, although there are plausible physical reasons for
1267 relationships between early and late century trends, inconsistencies between the results from different model generations
1268 reduce confidence in the robustness of these relationships. Recent work has documented significant differences in stratospheric
1269 transport among the two CCMi model generations discussed here (Abalos et al., 2026), although the specific impacts of these
1270 differences on emergent ozone behavior is not currently known.

1271 9 Conclusions

1272 In this study, we analyzed tropical total and partial column ozone trends in layers spanning the troposphere and stratosphere
1273 in multiple merged satellite datasets and chemistry-climate simulations from the ~~REF-C2~~ and ~~REF-D2~~ multi-model
1274 experiments. By applying dynamical linear modeling (DLM) to isolate background trends from natural variability, we assessed
1275 the consistency between simulated and observed ozone evolution over the observational time period (~1980 to present),
1276 quantified 21st century model trends, and assessed the potential for the use of an observational constraint on future model
1277 trends.

1278 Our results indicate that while both generations of CCMs simulate tropical TCO increases over the 2000–2021 period, the
1279 ~~REF-D2~~ ensemble mean trend is more positive and in better quantitative agreement with observations than ~~REF-C2~~. We find
1280 that the shift towards more positive TCO trends in ~~REF-D2~~ is primarily due to a much stronger increase in tropical tropospheric
1281 ozone in ~~REF-D2~~ (~1.5 DU) compared to ~~REF-C2~~ (~0.3 DU) since 2000, likely driven by stronger increases in precursor
1282 emissions in ~~REF-D2~~. For the stratosphere as a whole, PCO changes are nearly identical between the model experiments,
1283 although ~~REF-D2~~ shows slightly more negative trends (by about 0.5 DU) in the lower stratosphere, which are compensated by
1284 more positive trends in the middle/upper stratosphere compared to ~~REF-C2~~.

1285 A major finding of this work is that substantial disagreement exists among observed tropical ozone trends over the past few
1286 decades, particularly in troposphere and middle/upper stratosphere, and that these observational uncertainties critically limit

Deleted: predicted

Deleted: CCMi

Deleted: 1

Deleted: CCMi

Deleted: 202

Deleted: CCMi

Deleted: 202

Deleted: CCMi

Deleted: 1

Deleted: CCMi

Deleted: 202

Deleted: CCMi

Deleted: -202

Deleted: CCMi

Deleted: -

Deleted: 1

Deleted: CCMi-2022

Deleted: CCMi-2022

Deleted: CCMi-1

our ability to evaluate the fidelity of CCMs in reproducing past changes. For example, in the upper stratosphere, observational estimates for the 2000–2021 period range from +0.6 DU to +2.6 DU, with minimal overlap between many of the datasets. Similarly, estimates of tropospheric ozone changes range from negligible (SUNLIT) and in agreement with REF-C2, to strongly positive (TOMS+OMI/MLS) and in agreement with REF-D2, further complicating the validation of simulated trends. Ongoing work as part of the TOAR II activity is undertaking a more thorough intercomparison of satellite tropospheric column datasets, and their combined record generally supports the smaller end of the observational estimates presented here. Despite these disagreements in the troposphere and middle/upper stratosphere, over 2000–2021 we find mostly significant negative lower stratospheric trends in observations that are on average more negative than those from both CCM1 experiments. The CCM1 multi-model mean ozone trends are negative (−0.6 DU for REF-C2, −1.1 DU for REF-D2), as are most of the DLM posterior trend distributions (69% for REF-C2 and 77% for REF-D2). These findings reinforce earlier findings of negative lower stratospheric ozone trends across multiple observational datasets (e.g., Ball et al., 2018; Godin-Beekman et al., 2022; Sofieva et al., 2026). To understand how disagreements among currently available observations impact the future understanding of ozone, we examined the relationship between early 21st-century recovery trends and end-of-century trends in models. We found a robust inter-model correlation in all layers of the stratosphere for REF-C2, with a particularly strong correlation in the middle stratosphere (30–10 hPa). This suggests that models with faster early-recovery rates tend to project higher end-of-century ozone levels. However, this "emergent constraint" relationship is largely absent in REF-D2, except for a borderline significant correlation in the middle stratosphere. Although the reasons for this lack of correlation in the newer model simulations are not known, it is possible that they are due to differences in late-21st century forcing between the two REF experiments, the representation of dynamical processes such as the QBO that potentially change over the course of the century, or as-yet undiagnosed model biases. Despite the presence of a relationship between recent and future ozone change in models, the practical application of an observational constraint is currently not possible. By propagating the spread of recent observational trends through the model-derived relationships, we demonstrated that the resulting uncertainty in "observationally-informed" end-of-century ozone is as large as, or larger than, the spread in the raw model projections. Depending on which observational dataset is treated as "truth," one could conclude that models either systematically overestimate or underestimate stratospheric ozone recovery for a given emissions scenario. Ultimately, these results indicate that uncertainties in recent tropical stratospheric and tropospheric ozone changes restrict our understanding of future total column ozone evolution. The current spread of observational trend estimates is insufficient for providing a robust constraint on future model projections. These results highlight a critical need for better reconciliation of observational ozone trend estimates among existing datasets, likely involving investigation of potential drifts in the source records that contribute to them, and further supports the utility of stable, trend quality stratospheric ozone measurements from both satellite and other platforms. Finally, the analysis here suggests a need to better understand inter-generational differences in chemistry climate models that lead to differing relationships between early and late 21st century ozone recovery.

Deleted: CCM1-1

Deleted: CCM1-2022

Deleted: REF-

Deleted: CCM1-2022

Deleted:

Deleted: CCM1-1

Deleted: CCM1-2022

Deleted: predicted

Deleted: highlight

1349 **Acknowledgements**

1350 The SBUV-COH dataset was developed with support by NOAA Grant NA24NESX432C0001 (Cooperative Institute for

1351 Satellite Earth System Studies - CISESS) at the University of Maryland/ESSIC. The development of GTO-ECV is funded by

1352 German Aerospace Center (DLR) and the EU Copernicus Climate Change Service (C3S). CA, AR, VS, and MW acknowledge

1353 the partial financial support from the ESA Project OREGANO (Ozone Recovery from Merged Observational Data and Model

1354 Analysis) under the contract number 4000137112/22/I-AG. [GC and AC acknowledge funding from the EU via the ERC](#)

1355 [Starting Grant 101078127](#), [SD acknowledges support from NASA Grant 80NSSC26K0049](#). AI was used for assistance in plot

1356 code and conclusion section initial drafting. CCSRNIES modeling group was supported by KAKENHI (JP24K00700,

1357 JP24H00751, JP25K07401 and JP25K00377) of the Ministry of Education, Culture, Sports, Science, and Technology, Japan.

1358 NEC SX-ACE and SX-AURORA TSUBASA at NIES were used to perform CCSRNIES model simulations.

1359 **Author contributions**

1360 Conceptualization: SD, WB, YJ, GC, JA, and JK.

1361 Providing satellite and reanalysis data: CA, M C-E, RD, SF, DH, NK, DL,E M-B, AR, VS, R VdA, RW, KW, MW, JW, JZ

1362 Providing model output and advisement on model output: HA, FD, SD, PJ, BJ, MM, OM, DP, TS, SW, YY

1363 Analysis and guidance: EB, AC, RD, MD, SD, RE, BH, MH, DH, RW, KR, JS, KT

1364 Writing (reviewing and editing): All coauthors

1365 **Competing interests**

1366 At least one of the coauthors is a member of the editorial board of Atmospheric Chemistry and Physics.

1367 **Code and data availability**

1368 All CCMI-1 and CCMI-2022 model output are archived and accessible at the British Atmospheric Data Centre. Individual

1369 merged satellite datasets and reanalyses are distributed by the dataset PIs, and access information can be found in the referenced

1370 works for each of the datasets listed in the observations and reanalyses section. The dlmme software package for DLM analysis

1371 is publicly available at <https://github.com/justinalsing/dlmme>. Plotting and analysis codes are available upon request to the

1372 first author.

Deleted:

1374 **References**

- 1375 Abalos, M., Birner, T., Chrysanthou, A., Davis, S., de la Cámara, A., Dhomse, S., Garny, H., Hegglin, M. I., Hubert, D.,
1376 Ivaniha, O., Keeble, J., Linz, M., Minganti, D., Neu, J., Plummer, D., Saunders, L., Shah, K., Stiller, G., Tourpali, K., Waugh,
1377 D., Abraham, N. L., Akiyoshi, H., Chipperfield, M. P., Jöckel, P., Josse, B., Morgenstern, O., Sukhodolov, T., Watanabe, S.,
1378 and Yamashita, Y.: Evaluation of stratospheric transport in three generations of Chemistry-Climate Models, EGUsphere, 1–
1379 61, <https://doi.org/10.5194/egusphere-2025-6549>, 2026.
- 1380
1381 Acquah, C., Stecher, L., Mertens, M., and Jöckel, P.: Effects of different emission inventories on tropospheric ozone and
1382 methane lifetime, Atmospheric Chemistry and Physics, 25, 13665–13686, <https://doi.org/10.5194/acp-25-13665-2025>, 2025.
- 1383
1384 Alsing, J.: dlmmc: Dynamical linear model regression for atmospheric time-series analysis, The Journal of Open Source
1385 Software, 4, 1157, <https://doi.org/10.21105/joss.01157>, 2019.
- 1386
1387 APARC, 2025: The Hunga Volcanic Eruption Atmospheric Impacts Report. Yunqian Zhu, Graham Mann, Paul A. Newman,
1388 and William Randel (Eds.), APARC Report No. 11, WCRP-10/2025, doi: 10.34734/FZJ-2025-05237, available at [www.aparc-](http://www.aparc-climate.org/publications/)
1389 [climate.org/publications/](http://www.aparc-climate.org/publications/)
- 1390
1391 Arosio, C., Rozanov, A., Malinina, E., Eichmann, K.-U., von Clarmann, T., and Burrows, J. P.: Retrieval of ozone profiles
1392 from OMPS limb scattering observations, Atmos. Meas. Tech., 11, 2135–2149, <https://doi.org/10.5194/amt-11-2135-2018>,
1393 2018.
- 1394
1395 Ball, W. T., Alsing, J., Mortlock, D. J., Rozanov, E. V., Tummon, F., and Haigh, J. D.: Reconciling differences in stratospheric
1396 ozone composites, Atmospheric Chemistry and Physics, 17, 12269–12302, <https://doi.org/10.5194/acp-17-12269-2017>, 2017.
- 1397
1398 Ball, W. T., Alsing, J., Mortlock, D. J., Staehelin, J., Haigh, J. D., Peter, T., Tummon, F., Stübi, R., Stenke, A., Anderson, J.,
1399 Bourassa, A., Davis, S. M., Degenstein, D., Frith, S., Froidevaux, L., Roth, C., Sofieva, V., Wang, R., Wild, J., Yu, P., Ziemke,
1400 J. R., and Rozanov, E. V.: Evidence for a continuous decline in lower stratospheric ozone offsetting ozone layer recovery,
1401 Atmospheric Chemistry and Physics, 18, 1379–1394, <https://doi.org/10.5194/acp-18-1379-2018>, 2018.
- 1402
1403 Ball, W. T., Alsing, J., Staehelin, J., Davis, S. M., Froidevaux, L., and Peter, T.: Stratospheric ozone trends for 1985–2018:
1404 sensitivity to recent large variability, Atmospheric Chemistry and Physics, 19, 12731–12748, [https://doi.org/10.5194/acp-19-](https://doi.org/10.5194/acp-19-12731-2019)
1405 [12731-2019](https://doi.org/10.5194/acp-19-12731-2019), 2019.
- 1406

1407 Ball, W. T., Chiodo, G., Abalos, M., Alsing, J., and Stenke, A.: Inconsistencies between chemistry–climate models and
1408 observed lower stratospheric ozone trends since 1998, *Atmospheric Chemistry and Physics*, 20, 9737–9752,
1409 <https://doi.org/10.5194/acp-20-9737-2020>, 2020. 

1410 

1411 Bodeker, G. E., Nitzbon, J., Tradowsky, J. S., Kremser, S., Schwertheim, A., and Lewis, J.: A global total column ozone
1412 climate data record, *Earth System Science Data*, 13, 3885–3906, <https://doi.org/10.5194/essd-13-3885-2021>, 2021.

1413

1414 Bourassa, A. E., Degenstein, D. A., Randel, W. J., Zawodny, J. M., Kyrölä, E., McLinden, C. A., Sioris, C. E., and Roth, C.
1415 Z.: Trends in stratospheric ozone derived from merged SAGE II and Odin-OSIRIS satellite observations, *Atmospheric*
1416 *Chemistry and Physics*, 14, 6983–6994, <https://doi.org/10.5194/acp-14-6983-2014>, 2014.

1417

1418 Bourassa, A. E., Roth, C. Z., Zawada, D. J., Rieger, L. A., McLinden, C. A., and Degenstein, D. A.: Drift-corrected Odin-
1419 OSIRIS ozone product: algorithm and updated stratospheric ozone trends, *Atmospheric Measurement Techniques*, 11, 489–
1420 498, <https://doi.org/10.5194/amt-11-489-2018>, 2018.

1421

1422 Butchart, N.: The Brewer-Dobson circulation, *Rev Geophys*, 52, 157–184, <https://doi.org/10.1002/2013RG000448>, 2014.

1423

1424 Davis, S. M., Rosenlof, K. H., Hassler, B., Hurst, D. F., Read, W. G., Vömel, H., Selkirk, H., Fujiwara, M., and Damadeo, R.:
1425 The Stratospheric Water and Ozone Satellite Homogenized (SWOOSH) database: a long-term database for climate studies,
1426 *Earth Syst. Sci. Data*, 8, 461–490, <https://doi.org/10.5194/essd-8-461-2016>, 2016.

1427

1428 Davis, S. M., Davis, N., Portmann, R. W., Ray, E., and Rosenlof, K.: The role of tropical upwelling in explaining discrepancies
1429 between recent modeled and observed lower-stratospheric ozone trends, *Atmospheric Chemistry and Physics*, 23, 3347–3361,
1430 <https://doi.org/10.5194/acp-23-3347-2023>, 2023.

1431

1432 Dietmüller, S., Garny, H., Eichinger, R., and Ball, W. T.: Analysis of recent lower-stratospheric ozone trends in chemistry
1433 climate models, *Atmospheric Chemistry and Physics*, 21, 6811–6837, <https://doi.org/10.5194/acp-21-6811-2021>, 2021.

1434

1435 Dobson, G. M. B.: Origin and distribution of the polyatomic molecules in the atmosphere, *Proc R Soc Lon Ser-A*, 236, 187–
1436 193, 1956.

1437

1438 Dobson, G. M. B., Harrison, D. N., and Lawrence, J.: Measurements of the Amount of Ozone in the Earth’s Atmosphere and
1439 Its Relation to Other Geophysical Conditions. Part III, *Proc. R. Soc. Lond. A*, 122, 456–486, 1929.

1440

Deleted: 

Deleted: Benito-Barca, S., Abalos, M., Calvo, N., Garny, H., Birner, T., Abraham, N. L., Akiyoshi, H., Dennison, F., Jöckel, P., Josse, B., Keeble, J., Kinnison, D., Marchand, M., Morgenstern, O., Plummer, D., Rozanov, E., Strode, S., Sukhodolov, T., Watanabe, S., and Yamashita, Y.: Recent Lower Stratospheric Ozone Trends in CCM1-2022 Models: Role of Natural Variability and Transport, *Journal of Geophysical Research: Atmospheres*, 130, e2024JD042412, <https://doi.org/10.1029/2024JD042412>, 2025. 

1450 Eyring, V., Lamarque, J.-F., Hess, P., Arfeuille, F., Bowman, K., Chipperfield, M. P., Duncan, B., Fiore, A., Gettelman, A.,
 1451 Giorgetta, M. A., Granier, C., Hegglin, M., Kinnison, D., Kunze, M., Langematz, U., Luo, B., Martin, R., Matthes, K.,
 1452 Newman, P. A., Peter, T., Robock, A., Ryerson, T., Saiz-Lopez, A., Salawitch, R., Schultz, M., Shepherd, T. G., Shindell, D.,
 1453 Staehelin, J., Tegtmeier, S., Thomason, L., Tilmes, S., Vernier, J.-P., Waugh, D. W., and Young, P. J.: Overview of
 1454 IGAC/SPARC Chemistry-Climate Model Initiative (CCMI) community simulations in support of upcoming ozone and climate
 1455 assessments, SPARC Newsletter, 40, 48–66, 2013.
 1456
 1457 Fioletov, V. E., Bodeker, G. E., Miller, A. J., McPeters, R. D., and Stolarski, R.: Global and zonal total ozone variations
 1458 estimated from ground-based and satellite measurements: 1964–2000, *Journal of Geophysical Research: Atmospheres*, 107,
 1459 ACH 21-1-ACH 21-14, <https://doi.org/10.1029/2001JD001350>, 2002.
 1460
 1461 Fioletov, V. E., Labow, G., Evans, R., Hare, E. W., Köhler, U., McElroy, C. T., Miyagawa, K., Redondas, A., Savastouk, V.,
 1462 Shalamyansky, A. M., Staehelin, J., Vanicek, K., and Weber, M.: Performance of the ground-based total ozone network
 1463 assessed using satellite data, *Journal of Geophysical Research: Atmospheres*, 113, <https://doi.org/10.1029/2008JD009809>,
 1464 2008.
 1465
 1466 Frith, S. M., Kramarova, N. A., Stolarski, R. S., McPeters, R. D., Bhartia, P. K., and Labow, G. J.: Recent changes in total
 1467 column ozone based on the SBUV Version 8.6 Merged Ozone Data Set, *Journal of Geophysical Research: Atmospheres*, 119,
 1468 9735–9751, <https://doi.org/10.1002/2014JD021889>, 2014.
 1469
 1470 Frith, S. M., Stolarski, R. S., Kramarova, N. A., and McPeters, R. D.: Estimating uncertainties in the SBUV Version 8.6 merged
 1471 profile ozone data set, *Atmospheric Chemistry and Physics*, 17, 14695–14707, <https://doi.org/10.5194/acp-17-14695-2017>,
 1472 2017.
 1473
 1474 Froidevaux, L., Anderson, J., Wang, H. J., Fuller, R. A., Schwartz, M. J., Santee, M. L., Livesey, N. J., Pumphrey, H. C.,
 1475 Bernath, P. F., Russell III, J. M., and McCormick, M. P.: Global OZone Chemistry And Related trace gas Data records for the
 1476 Stratosphere (GOZCARDS): methodology and sample results with a focus on HCl, H₂O, and O₃, *Atmospheric Chemistry and*
 1477 *Physics*, 15, 10471–10507, <https://doi.org/10.5194/acp-15-10471-2015>, 2015.
 1478 Funke, B., and Matthes, K.: Report on the SOLARIS-HEPPA Working Group Meeting, SPARC Newsletter, 54, 2020.
 1479
 1480 Gaudel, A., Cooper, O. R., Ancellet, G., Barret, B., Boynard, A., Burrows, J. P., Clerbaux, C., Coheur, P.-F., Cuesta, J., Cuevas,
 1481 E., Doniki, S., Dufour, G., Ebojje, F., Foret, G., Garcia, O., Granados-Muñoz, M. J., Hannigan, J. W., Hase, F., Hassler, B.,
 1482 Huang, G., Hurtmans, D., Jaffe, D., Jones, N., Kalabokas, P., Kerridge, B., Kulawik, S., Latter, B., Leblanc, T., Le Flochmoën,
 1483 E., Lin, W., Liu, J., Liu, X., Mahieu, E., McClure-Begley, A., Neu, J. L., Osman, M., Palm, M., Petetin, H., Petropavlovskikh,

I., Querel, R., Ralpoe, N., Rozanov, A., Schultz, M. G., Schwab, J., Siddans, R., Smale, D., Steinbacher, M., Tanimoto, H., Tarasick, D. W., Thouret, V., Thompson, A. M., Trickl, T., Weatherhead, E., Wespes, C., Worden, H. M., Vigouroux, C., Xu, X., Zeng, G., and Ziemke, J.: Tropospheric Ozone Assessment Report: Present-day distribution and trends of tropospheric ozone relevant to climate and global atmospheric chemistry model evaluation, *Elementa: Science of the Anthropocene*, 6, 39, <https://doi.org/10.1525/elementa.291>, 2018.

Godin-Beekmann, S., Azouz, N., Sofieva, V., Hubert, D., Petropavlovskikh, I., Effertz, P., Ancellet, G., Degenstein, D., Zawada, D., Froidevaux, L., Frith, S., Wild, J., Davis, S., Steinbrecht, W., Leblanc, T., Querel, R., Tourpali, K., Damadeo, R., Maillard Barras, E., Stübi, R., Vigouroux, C., Arosio, C., Nedoluha, G., Boyd, I., and van Malderen, R.: Updated trends of the stratospheric ozone vertical distribution in the 60° S–60° N latitude range based on the LOTUS regression model, *Atmospheric Chemistry and Physics*, 1–28, <https://doi.org/10.5194/acp-2022-137>, 2022.

Hegglin, Michaela I., and Theodore G. Shepherd. “Large Climate-Induced Changes in Ultraviolet Index and Stratosphere-to-Troposphere Ozone Flux.” *Nature Geoscience* 2, no. 10 (2009): 687–91. <https://doi.org/10.1038/ngeo604>.

[Hubert, D., Miyazaki, K., Dufour, G., Pennington, E. A., Sofieva, V., Arosio, C., Barret, B., Boynard, A., Coldewey-Egbers, M., Cuesta, J., Heue, K.-P., Keppens, A., Kramarova, N. A., Lambert, J.-C., Loyola, D., Orfanoz-Cheuquelaf, A. P., Siddans, R., Van Malderen, R., Veeffkind, P., Wespes, C., Ziemke, J. R., Tropospheric Ozone Assessment Report II: Past and present tropospheric ozone using satellite observations, *Philos. Trans. R. Soc. A.* \(submitted\).](#)

Jones, B. and O’Neill, B. C.: Spatially explicit global population scenarios consistent with the Shared Socioeconomic Pathways, *Environ. Res. Lett.*, 11, 084003, <https://doi.org/10.1088/1748-9326/11/8/084003>, 2016.

Laine, M., Latva-Pukkila, N., and Kyrölä, E.: Analysing time-varying trends in stratospheric ozone time series using the state space approach, *Atmospheric Chemistry and Physics*, 14, 9707–9725, <https://doi.org/10.5194/acp-14-9707-2014>, 2014.

Lamy, Kévin, Thierry Portafaix, Béatrice Josse, et al. “Clear-Sky Ultraviolet Radiation Modelling Using Output from the Chemistry Climate Model Initiative.” *Atmospheric Chemistry and Physics* 19, no. 15 (2019): 10087–110. <https://doi.org/10.5194/acp-19-10087-2019>.

[Lickley, Megan J., John S. Daniel, Laura A. McBride, Ross J. Salawitch, and Guus J. M. Velders. “The Return to 1980 Stratospheric Halogen Levels: A Moving Target in Ozone Assessments from 2006 to 2022.” *Atmospheric Chemistry and Physics* 24, no. 22 \(2024\): 13081–99. <https://doi.org/10.5194/acp-24-13081-2024>.](#)

1518 Maillard Barras, E., Haeefe, A., Stübi, R., Jouberton, A., Schill, H., Petropavlovskikh, I., Miyagawa, K., Stanek, M., and
1519 Froidevaux, L.: Dynamical linear modeling estimates of long-term ozone trends from homogenized Dobson Umkehr profiles
1520 at Arosa/Davos, Switzerland, *Atmospheric Chemistry and Physics*, 22, 14283–14302, [https://doi.org/10.5194/acp-22-14283-](https://doi.org/10.5194/acp-22-14283-2022)
1521 2022, 2022.

1522

1523 Meinshausen, M., Smith, S. J., Calvin, K., Daniel, J. S., Kainuma, M. L. T., Lamarque, J.-F., Matsumoto, K., Montzka, S. A.,
1524 Raper, S. C. B., Riahi, K., Thomson, A., Velders, G. J. M., and van Vuuren, D. P. P.: The RCP greenhouse gas concentrations
1525 and their extensions from 1765 to 2300, *Climatic Change*, 109, 213, <https://doi.org/10.1007/s10584-011-0156-z>, 2011.

1526

1527 Meinshausen, M., Vogel, E., Nauels, A., Lorbacher, K., Meinshausen, N., Etheridge, D. M., Fraser, P. J., Montzka, S. A.,
1528 Rayner, P. J., Trudinger, C. M., Krummel, P. B., Beyerle, U., Canadell, J. G., Daniel, J. S., Enting, I. G., Law, R. M., Lunder,
1529 C. R., O'Doherty, S., Prinn, R. G., Reimann, S., Rubino, M., Velders, G. J. M., Vollmer, M. K., Wang, R. H. J., and Weiss,
1530 R.: Historical greenhouse gas concentrations for climate modelling (CMIP6), *Geoscientific Model Development*, 10, 2057–
1531 2116, <https://doi.org/10.5194/gmd-10-2057-2017>, 2017.

1532

1533 Meinshausen, M., Nicholls, Z. R. J., Lewis, J., Gidden, M. J., Vogel, E., Freund, M., Beyerle, U., Gessner, C., Nauels, A.,
1534 Bauer, N., Canadell, J. G., Daniel, J. S., John, A., Krummel, P. B., Luderer, G., Meinshausen, N., Montzka, S. A., Rayner, P.
1535 J., Reimann, S., Smith, S. J., van den Berg, M., Velders, G. J. M., Vollmer, M. K., and Wang, R. H. J.: The shared socio-
1536 economic pathway (SSP) greenhouse gas concentrations and their extensions to 2500, *Geoscientific Model Development*, 13,
1537 3571–3605, <https://doi.org/10.5194/gmd-13-3571-2020>, 2020.

1538

1539 Nowack, P., Ceppi, P., Davis, S. M., Chiodo, G., Ball, W., Diallo, M. A., Hassler, B., Jia, Y., Keeble, J., and Joshi, M.:
1540 Response of stratospheric water vapour to warming constrained by satellite observations, *Nat. Geosci.*, 16, 577–583,
1541 <https://doi.org/10.1038/s41561-023-01183-6>, 2023.

1542

1543 O'Neill, B. C., Tebaldi, C., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J.-
1544 F., Lowe, J., Meehl, G. A., Moss, R., Riahi, K., and Sanderson, B. M.: The Scenario Model Intercomparison Project
1545 (ScenarioMIP) for CMIP6, *Geoscientific Model Development*, 9, 3461–3482, <https://doi.org/10.5194/gmd-9-3461-2016>,
1546 2016.

1547

1548 Petropavlovskikh, I., Wild, J. D., Abromitis, K., Effertz, P., Miyagawa, K., Flynn, L., Maillard Barras, E., Damadeo, R.,
1549 McConville, G., Johnson, B., Cullis, P., Godin-Beekmann, S., Ancellet, G., Querel, R., Van Malderen, R., Zawada, D., Ozone
1550 trends in homogenized Umkehr, ozonesonde, and COH overpass records, *Atmos. Chem. Phys.*, 25, 2895–2936,
1551 <https://doi.org/10.5194/acp-25-2895-2025>, 2025.

1552

1553 Morgenstern, O., Hegglin, M. I., Rozanov, E., O'Connor, F. M., Abraham, N. L., Akiyoshi, H., Archibald, A. T., Bekki, S.,
1554 Butchart, N., Chipperfield, M. P., Deushi, M., Dhomse, S. S., Garcia, R. R., Hardiman, S. C., Horowitz, L. W., Jöckel, P.,
1555 Josse, B., Kinnison, D., Lin, M., Mancini, E., Manyin, M. E., Marchand, M., Marécal, V., Michou, M., Oman, L. D., Pitari,
1556 G., Plummer, D. A., Revell, L. E., Saint-Martin, D., Schofield, R., Stenke, A., Stone, K., Sudo, K., Tanaka, T. Y., Tilmes, S.,
1557 Yamashita, Y., Yoshida, K., and Zeng, G.: Review of the global models used within phase 1 of the Chemistry–Climate Model
1558 Initiative (CCMI), *Geosci. Model Dev.*, 10, 639–671, <https://doi.org/10.5194/gmd-10-639-2017>, 2017.

1559

1560 Plummer, D., Nagashima, T., Tilmes, S., Archibald, A., Chiodo, G., Fadnavis, S., Garny, H., Josse, B., Kim, J., Lamarque, J.-
1561 F., Morgenstern, O., Murray, L., Orbe, C., Tai, A., Chipperfield, M., Funke, B., Juckes, M., Kinnison, D., Kunze, M., Luo, B.,
1562 Matthes, K., Newman, P. A., Pascoe, C., and Peter, T.: CCMI-2022: A new set of Chemistry-Climate Model Initiative (CCMI)
1563 Community Simulations to Update the Assessment of Models and Support Upcoming Ozone Assessment Activities, *SPARC*
1564 *newsletter*, 57, 2021.

1565

1566 Richter, J. H., Anstey, J. A., Butchart, N., Kawatani, Y., Meehl, G. A., Osprey, S., and Simpson, I. R.: Progress in Simulating
1567 the Quasi-Biennial Oscillation in CMIP Models, *Journal of Geophysical Research: Atmospheres*, 125, e2019JD032362,
1568 <https://doi.org/10.1029/2019JD032362>, 2020.

1569

1570 Santee, M. L., Lambert, A., Froidevaux, L., Manney, G. L., Schwartz, M. J., Millán, L. F., Livesey, N. J., Read, W. G., Werner,
1571 F., and Fuller, R. A.: Strong Evidence of Heterogeneous Processing on Stratospheric Sulfate Aerosol in the Extrapolar Southern
1572 Hemisphere Following the 2022 Hunga Tonga-Hunga Ha’apai Eruption, *Journal of Geophysical Research: Atmospheres*, 128,
1573 e2023JD039169, <https://doi.org/10.1029/2023JD039169>, 2023.

1574

1575 Shepherd, T. G., D. A. Plummer, J. F. Scinocca, et al. “Reconciliation of Halogen-Induced Ozone Loss with the Total-Column
1576 Ozone Record.” *Nature Geoscience* 7, no. 6 (2014): 443–49. <https://doi.org/10.1038/ngeo2155>.

1577

1578 Sofieva, V. F., Hänninen, R., Sofiev, M., Szelag, M., Lee, H. S., Tamminen, J., and Retscher, C.: Synergy of Using Nadir and
1579 Limb Instruments for Tropospheric Ozone Monitoring (SUNLIT), *Atmospheric Measurement Techniques*, 15, 3193–3212,
1580 <https://doi.org/10.5194/amt-15-3193-2022>, 2022.

1581

1582 Sofieva, V. F., Szelag, M., Tamminen, J., Arosio, C., Rozanov, A., Weber, M., Degenstein, D., Bourassa, A., Zawada, D.,
1583 Kiefer, M., Laeng, A., Walker, K. A., Sheese, P., Hubert, D., van Roozendaal, M., Retscher, C., Damadeo, R., and Lumpe, J.
1584 D.: Updated merged SAGE-CCI-OMPS+ dataset for the evaluation of ozone trends in the stratosphere, *Atmospheric*
1585 *Measurement Techniques*, 16, 1881–1899, <https://doi.org/10.5194/amt-16-1881-2023>, 2023.

1586

SPARC, 2017: The SPARC Data Initiative: Assessment of stratospheric trace gas and aerosol climatologies from satellite limb sounders. By M. I. Hegglin and S. Tegtmeier (eds.), SPARC Report No. 8, WCRP-5/2017, available at www.aparc-climate.org/publications/

Thompson, A. M., Stauffer, R. M., Wargan, K., Witte, J. C., Kollonige, D. E., and Ziemke, J. R.: Regional and Seasonal Trends in Tropical Ozone From SHADOZ Profiles: Reference for Models and Satellite Products, *J Geophys Res-Atmos*, 126, e2021JD034691, <https://doi.org/10.1029/2021JD034691>, 2021.

van der A, R. J., Allaart, M. a. F., and Eskes, H. J.: Extended and refined multi sensor reanalysis of total ozone for the period 1970–2012, *Atmospheric Measurement Techniques*, 8, 3021–3035, <https://doi.org/10.5194/amt-8-3021-2015>, 2015.

Van Malderen, R., Zang, Z., Chang, K.-L., Björklund, R., Cooper, O. R., Liu, J., Maillard Barras, E., Vigouroux, C., Petropavlovskikh, I., Leblanc, T., Thouret, V., Wolff, P., Effertz, P., Gaudel, A., Tarasick, D. W., Smit, H. G. J., Thompson, A. M., Stauffer, R. M., Kollonige, D. E., Poyraz, D., Ancellet, G., De Backer, M.-R., Frey, M. M., Hannigan, J. W., Hernandez, J. L., Johnson, B. J., Jones, N., Kivi, R., Mahieu, E., Morino, I., McConville, G., Müller, K., Murata, I., Notholt, J., Piters, A., Prignon, M., Querel, R., Rizi, V., Smale, D., Steinbrecht, W., Strong, K., and Sussmann, R.: Ground-based tropospheric ozone measurements: regional tropospheric ozone column trends from the TOAR-II/HEGIFTOM homogenized datasets, *Atmos. Chem. Phys.*, 25, 9905–9935, <https://doi.org/10.5194/acp-25-9905-2025>, 2025.

Villamayor, J., Iglesias-Suarez, F., Cuevas, C. A., Fernandez, R. P., Li, Q., Abalos, M., Hossaini, R., Chipperfield, M. P., Kinnison, D. E., Tilmes, S., Lamarque, J.-F., and Saiz-Lopez, A.: Very short-lived halogens amplify ozone depletion trends in the tropical lower stratosphere, *Nat. Clim. Chang.*, 13, 554–560, <https://doi.org/10.1038/s41558-023-01671-y>, 2023.

Wargan, K., Weir, B., Manney, G. L., Cohn, S. E., Knowland, K. E., Wales, P. A., and Livesey, N. J.: M2-SCREAM: A Stratospheric Composition Reanalysis of Aura MLS Data With MERRA-2 Transport, *Earth and Space Science*, 10, e2022EA002632, <https://doi.org/10.1029/2022EA002632>, 2023.

Weber, M., Coldewey-Egbers, M., Fioletov, V. E., Frith, S. M., Wild, J. D., Burrows, J. P., Long, C. S., and Loyola, D.: Total ozone trends from 1979 to 2016 derived from five merged observational datasets – the emergence into ozone recovery, *Atmospheric Chemistry and Physics*, 18, 2097–2117, <https://doi.org/10.5194/acp-18-2097-2018>, 2018.

Weber, M., Arosio, C., Coldewey-Egbers, M., Fioletov, V. E., Frith, S. M., Wild, J. D., Tourpali, K., Burrows, J. P., and Loyola, D.: Global total ozone recovery trends attributed to ozone-depleting substance (ODS) changes derived from five merged ozone datasets, *Atmospheric Chemistry and Physics*, 22, 6843–6859, <https://doi.org/10.5194/acp-22-6843-2022>, 2022.

Formatted: Font color: Black

1621
1622 [Wild, J. D.: A Merged Ozone Profile Dataset from SBUV, SBUV/2, OMPS: 1978 to 2024: A Technical Note,](#)
1623 https://www.star.nesdis.noaa.gov/data/smcd1/ozone/SBUV_OMPS_COH/COH_Tech_Note_v20250612.pdf (last access 13
1624 [May 2026\), 2025.](#)

1625

1626 WMO: Scientific Assessment of Ozone Depletion: 2010, Global Ozone Research and Monitoring Project, Geneva,
1627 Switzerland, Report No. 52, 2011.

1628

1629 WMO: Scientific Assessment of Ozone Depletion: 2022, GAW Report No. 278, 509, 2022.

1630

1631 Young, P. J., Harper, A. B., Huntingford, C., Paul, N. D., Morgenstern, O., Newman, P. A., Oman, L. D., Madronich, S., and
1632 Garcia, R. R.: The Montreal Protocol protects the terrestrial carbon sink, *Nature*, 596, 384–388,
1633 <https://doi.org/10.1038/s41586-021-03737-3>, 2021.

1634

1635 [Zawada, D. J., Rieger, L. A., Bourassa, A. E., and Degenstein, D. A.: Tomographic retrievals of ozone with the OMPS Limb](#)
1636 [Profiler: algorithm description and preliminary results, Atmospheric Measurement Techniques, 11, 2375–2393,](#)
1637 <https://doi.org/10.5194/amt-11-2375-2018>, 2018.

1638

1639 Zhang, Y., Cooper, O. R., Gaudel, A., Thompson, A. M., Nédélec, P., Ogino, S.-Y., and West, J. J.: Tropospheric ozone change
1640 from 1980 to 2010 dominated by equatorward redistribution of emissions, *Nature Geosci*, 9, 875–879,
1641 <https://doi.org/10.1038/ngeo2827>, 2016.

1642

1643 Ziemke, J. R., Chandra, S., and Bhartia, P. K.: Two new methods for deriving tropospheric column ozone from TOMS
1644 measurements: Assimilated UARS MLS/HALOE and convective-cloud differential techniques, *Journal of Geophysical*
1645 *Research: Atmospheres*, 103, 22115–22127, <https://doi.org/10.1029/98JD01567>, 1998.

1646

1647 Ziemke, J. R., Oman, L. D., Strode, S. A., Douglass, A. R., Olsen, M. A., McPeters, R. D., Bhartia, P. K., Froidevaux, L.,
1648 Labow, G. J., Witte, J. C., Thompson, A. M., Haffner, D. P., Kramarova, N. A., Frith, S. M., Huang, L.-K., Jaross, G. R.,
1649 Seftor, C. J., Deland, M. T., and Taylor, S. L.: Trends in global tropospheric ozone inferred from a composite record of
1650 TOMS/OMI/MLS/OMPS satellite measurements and the MERRA-2 GMI simulation, *Atmospheric Chemistry and Physics*,
1651 19, 3257–3269, <https://doi.org/10.5194/acp-19-3257-2019>, 2019.

1652

Page 21: [1] Deleted

Sean Davis

7/29/26 3:41:00 PM

▲

Page 23: [2] Deleted

Sean Davis

7/29/26 3:44:00 PM

X..