



Satellite nadir-viewing geometry affects the magnitude and detectability of long-term trends in stratospheric ozone

Louis Rivoire^{1,2,3}, Marianna Linz^{1,2}, Jessica L. Neu⁴, Pu Lin⁵, and Michelle L. Santee⁴

¹Department of Earth and Planetary Sciences, Harvard University, Cambridge, Massachusetts, USA

²School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts, USA

³Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

⁴Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA

⁵Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton, New Jersey, USA

Correspondence: Louis Rivoire (lrivoire@mit.edu)

Received: 20 August 2024 – Discussion started: 5 September 2024

Revised: 25 November 2024 – Accepted: 20 December 2024 – Published: 20 February 2025

Abstract. The continued monitoring of the ozone layer and its long-term evolution leans on comparative studies of merged satellite records. Comparing such records presents unique challenges due to differences in sampling, coverage, and retrieval algorithms between observing platforms, all of which complicate the detection of trends. Here we examine the effects of broad nadir averaging kernels on vertically resolved ozone trends, using one record as an example. We find errors as large as 1 % per decade and displacements in trend profile features by as much as 6 km in altitude due to the vertical redistribution of information by averaging kernels. Furthermore, we show that averaging kernels tend to increase (by 10 %–80 %, depending on the location) the length of the record needed to determine whether trend estimates are distinguishable from natural variability with good statistical confidence. We conclude that trend uncertainties may be underestimated, in part because averaging kernels misrepresent decadal to multidecadal internal variability, and in part because the removal of known modes of variability from the observed record can yield residual errors. The study provides a framework to reconcile differences between observing platforms and highlights the need for caution when using records from instruments with broad averaging kernels to quantify trends and their uncertainties.

1 Introduction

Since the discovery of the Antarctic ozone hole (Farman et al., 1985) and the advent of the World Meteorological Organization (WMO)/United Nations Environment Programme (UNEP) ozone assessment reports in 1989, there have been a number of activities dedicated to monitoring the state of the ozone layer and providing attribution of the long-term changes in ozone distribution. Notwithstanding declining levels of ozone-depleting substances (ODSs) (WMO, 2022), the search for evidence of the beginning of the recovery of the global ozone layer is ongoing. A number of chemistry–climate feedbacks complicate this search. The greenhouse-

driven cooling of the stratosphere is thought to aid the recovery in the mid-to-upper stratosphere by slowing down reactions that deplete ozone (Fels et al., 1980). In contrast, the acceleration of the Brewer–Dobson circulation widely predicted by chemistry–climate models (e.g., Eyring et al., 2013a; Meul et al., 2014; Abalos et al., 2021) is expected to redistribute stratospheric ozone from the tropics to the midlatitudes (e.g., Rind et al., 1990; Mahfouf et al., 1994; Shepherd, 2008) and to decrease stratospheric ozone abundances in the tropics (Shepherd, 2008; Li et al., 2009; Waugh et al., 2009; SPARC, 2010; Oman et al., 2010; Plummer et al., 2010; Meul et al., 2014; Banerjee et al., 2016; Chiodo et al., 2018; Dietmüller et al., 2021), with large impacts on

surface warming (Nowack et al., 2015) and ultraviolet radiation (Hegglin and Shepherd, 2009). Similarly, the expansion of the troposphere was also found to decrease stratospheric ozone by eroding the ozone layer from below (Match and Gerber, 2022). Because of these feedbacks, projected rates of recovery from chemistry–climate models are sensitive to model formulation (Dietmüller et al., 2014) and greenhouse gas emission scenarios (Revell et al., 2012). Lastly, emerging influences are adding new complexity to our understanding of the composition of the stratosphere, for instance, large wildfires and associated aerosols (Santee et al., 2022; Solomon et al., 2023); recent noncompliant ODS emissions (e.g., Montzka et al., 2018; Rigby et al., 2019; Chipperfield et al., 2020); new estimates for ODS lifetimes (Lickley et al., 2021), possibly increasing dynamical variability (Diallo et al., 2018, 2019); and large volcanic eruptions (e.g., the Hunga eruption in 2022; WMO, 2022; Wang et al., 2023; Evan et al., 2023; Santee et al., 2023; Wilmouth et al., 2023; Manney et al., 2023).

In light of such complexity, the analysis of observed trends continues to present challenges. The last WMO report (WMO, 2022) found a small but positive trend ($0.3 \pm 0.2\%$ per decade) in the near-global (60°S – 60°N) total column ozone. Vertically resolved trends reveal regional patterns hinting at competing influences on ozone abundances. The analysis of merged satellite products (e.g., Sofieva et al., 2017; Steinbrecht et al., 2017; Ball et al., 2017; Bourassa et al., 2018; WMO, 2018; Petropavlovskikh et al., 2019) found large positive post-2000 trends across the upper stratosphere (above 5 hPa), in agreement with the Chemistry–Climate Model Initiative (CCMI) simulations (Eyring et al., 2013b; Godin-Beekmann et al., 2022). However, agreement is lacking in the lower stratosphere (50–10 hPa): negative but uncertain trends are found in several satellite records, and although CCMI trends agree in sign below 30 hPa (Petropavlovskikh et al., 2019), the range across models and even across ensemble members of a single model is large (Stone et al., 2018). In the lowermost stratosphere (50–100 hPa), modeled and measured ozone and its trends remain highly uncertain. In addition, reanalysis products present varying degrees of realism in ozone trends (Davis et al., 2017) and should only be used with great caution (Box 3.2 in WMO, 2022). Whether models are flawed in their representation of stratospheric ozone or whether the statistical confidence placed in trend estimates is sufficient to address disagreements between models and observations, the need to improve our ability to distinguish trends from internal variability is clear.

In this paper, we turn to the nontrivial effects of error propagation in algorithms used to retrieve ozone abundances from spaceborne nadir measurements. Such errors are generally known to reduce the statistical confidence placed in trend estimates from merged records (Tummon et al., 2015; Hubert et al., 2016; Steinbrecht et al., 2017; Ball et al., 2017). This is especially concerning given that trends that were originally

assigned high statistical confidence can still suddenly change magnitude or even sign after the addition of just a few years of record (Chipperfield et al., 2018; Ball et al., 2019). The past literature provides extensive discussion about quantifying and reducing uncertainty in ozone trends (e.g., Stolarski and Frith, 2006; Harris et al., 2015; Petropavlovskikh et al., 2019), but methods have not yet been developed to systematically account for the limitations of satellite records. Using the example of the Solar Backscatter Ultraviolet (SBUV) merged record (McPeters et al., 2013), we utilize a novel method (Rivoire et al., 2024) to determine whether current trend estimates are distinguishable from natural variability, particularly when accounting for the effects of errors attributable to the SBUV averaging kernels.

2 Model output and observational data

In this section we introduce the datasets used in this study, namely:

- a pre-industrial chemistry–climate control simulation used as a reference for internal variability,
- chemistry–climate simulations used to estimate the time of emergence of future long-term ozone trends,
- the SBUV merged record, used as reference for the effects of satellite kernels, and
- other merged satellite records, used to compare against SBUV.

2.1 Simulations used as a reference for internal variability

In order to determine whether a trend derived from the observational record reflects natural variability or arises because of anthropogenic forcings, we need to compare said trend to a reference distribution of naturally occurring trends in a pre-industrial setting. Since the satellite era only covers the past 40–50 years, satellite records cannot be used to determine such a reference distribution. We therefore turn to a long simulation of the pre-industrial climate from a chemistry–climate model, building on standard practice (e.g., Kay et al., 2015).

Two runs from the Geophysical Fluid Dynamics Laboratory Earth System Model version 4.1 (ESM4.1; Dunne et al., 2020) are used as a reference: (1) a 500-year pre-industrial run (“piControl” in CMIP6 nomenclature) and (2) a historical run from 1850 to the end of 2014 (“historical”). The historical run provides a benchmark for the model’s performance given forcings and boundary conditions matching the historical record, including aerosol optical depths, 43 short-lived and long-lived greenhouse gases (GHGs), land use, solar forcing, sea surface temperatures, and sea ice concentrations. The pre-industrial run excludes such time-dependent forcings and is instead run with a prescribed global annual mean

atmospheric CO₂ concentration equal to that of 1850 (Eyring et al., 2016) and with a time-invariant volcanic aerosol forcing equivalent to the historical average (1850–2014). We use this simulation to quantify internal variability in the pre-industrial climate over any time period of a chosen length. Solar irradiance is also maintained constant following the specifications for the CMIP6 pre-industrial run (documented at <http://goo.gl/r8up31>, last access: 7 February 2025). The Quasi-Biennial Oscillation (QBO) in the model is internally generated.

The ESM4.1 model is the product of development efforts to capture coupled ocean–atmosphere and land–atmosphere interactions, biogeochemical cycles and ecosystem physics, sea ice, aerosol processes, and – importantly for this study – interactive ozone chemistry. The ozone chemistry includes an improved representation of ozone precursors (methane, carbon monoxide, nitrogen oxides, and volatile organic compounds) and accounts for heterogeneous reactions that occur on the surface of aerosols (Austin et al., 2013). ESM4.1 features 49 vertical levels with a model top at 1 Pa (~80 km) to better capture stratospheric chemistry and dynamics compared to the previous model generation (Horowitz et al., 2020). In the lower stratosphere, ESM4.1-simulated fields have a vertical resolution of about 2–3 km compared to the 4–6 km vertical resolution of SBUV (see Sect. 2.3). Coupled model simulations rarely include fully interactive chemistry in their extended control runs, making the ESM4.1 control run a unique asset for our study.

In order to ensure that the simulation provides suitable reference trend distributions, we assess the realism of the model's representation of stratospheric ozone variability. Simulated and observed variability is portrayed by power density spectra in Fig. 1. The spectra are estimated using the multitaper method (Thomson, 1982), which is suitable for short records. On interannual to multidecadal scales, it is inherently difficult to test whether the ozone variability produced by the model under pre-industrial conditions is realistic, as we have no historical record that spans the pre-industrial era. However, we can test whether the simulated ozone variability is on par with observations when the model is provided with historical radiative and chemical forcings that correspond to greenhouse gas and ODS emissions. If the model can capture the historical variability, we can at least say that it is skilled at capturing processes that are relevant to our current understanding of the atmosphere. One caveat in this reasoning is the fact that model improvements are tested against the historical record, which contains anthropogenic forcings; i.e., it is possible that some model improvements target modes of variability that are specific to external forcings rather than to natural climate variability. That being said, the historical model simulation exhibits improvements in the representation of total column ozone interannual variability and trends vs. previous versions of the model (Horowitz et al., 2020).

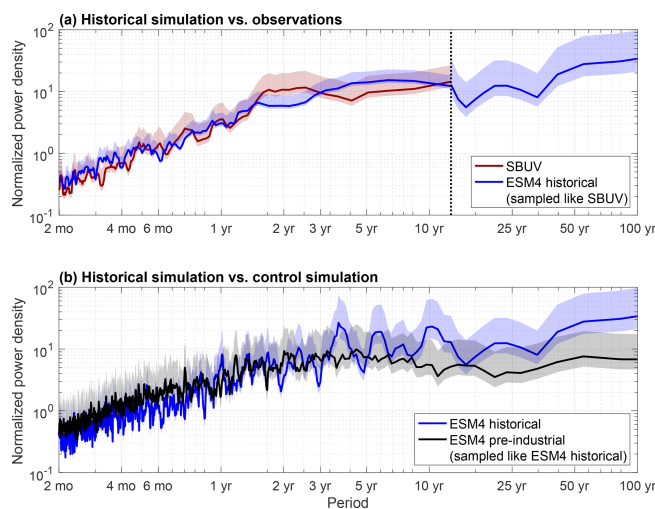


Figure 1. Comparison between observed and simulated normalized variability in the de-seasonalized zonal mean ozone for the 10–16 hPa layer in the Northern Hemisphere midlatitudes (37.5–57.5° N), pictured using power spectral densities (PSDs) estimated with the multitaper method. Shaded areas indicate chi-squared 95 % confidence intervals. To ensure comparability between spectra, the ESM4.1 historical run (1850–2014) is sub-sampled in 13-year periods, matching the longest continuous period given by the SBUV record (dashed line in **a**), and the resulting power spectra are subsequently averaged. Similarly, in **(b)**, the ESM4.1 pre-industrial run (500 years) is sub-sampled in 165-year periods, corresponding to the length of the historical run. Further, spectral density is normalized by the area under each power spectrum to remove changes in total power.

The pre-industrial simulation produces less ozone variability on decadal to centennial periods than the historical run. This behavior is expected since the pre-industrial run excludes forcings that affect ozone on these timescales (ODSs, GHGs, volcanic eruptions). This result means that trend uncertainties estimated from pre-industrial simulations are likely to be underestimated.

2.2 Simulations used to estimate trends

We use the average ozone trends from 2000 to 2020 from the same dataset as in Godin-Beekmann et al. (2022): the multi-model mean (16 models) from the CCM1's 1960–2100 period (REF-C2; Eyring et al., 2013b). The purpose of the REF-C2 experiments is to capture future climate trends, and as such, they follow the WMO (2011) A1 scenario for ODSs and the Representative Concentration Pathway (RCP) 6.0 for other greenhouse gas, ozone precursor, and aerosol precursor emissions. Depending on model capabilities, ocean boundary conditions and forcings for the 11-year solar cycle and the QBO are either simulated separately or generated internally. Although the recommendation was not to use volcanic forcings, some models do include them (Godin-Beekmann et al., 2022). More information about the forcings can be found in

Eyring et al. (2013a), Hegglin et al. (2016), and Morgenstern et al. (2017). Model runs are accessible at <http://catalogue.ceda.ac.uk/uuid/9cc6b94df0f4469d8066d69b5df879d5/> (last access: 7 February 2025) Hegglin et al. (2015).

2.3 SBUV satellite record

The SBUV Merged Ozone Dataset (MOD; version 2 release 1, https://acd-ext.gsfc.nasa.gov/Data_services/merged/, last access: 7 February 2025; see Frith et al., 2014, 2017) combines retrievals from backscattered ultraviolet radiation sensors on board a series of polar-orbiting satellites to provide total column and profile ozone products with monthly frequency. Retrievals include those from first- and second-generation SBUV sensors (algorithm version 8.7) and from the nadir profilers included in the Ozone Mapping and Profiling Suite (OMPS NP, algorithm version 2.8). We use the 2000–2020 portion (21 years) of the MOD (a) to focus on the time period after which the decline in ozone in the upper stratosphere stopped (Newchurch et al., 2003; WMO, 2007), (b) for consistency with other studies, and (c) to avoid issues with coverage gaps in the earlier parts of the record. Indeed, the MOD provides nearly global spatial coverage but suffers from substantial gaps in its temporal coverage. For instance, no reliable measurements are available from April 1976 through November 1978, and there are also none following large volcanic eruptions that injected aerosols in the stratosphere and affected the retrieval (e.g., mid-1991 through 1993; see Bhartia et al., 1993). Early portions of the dataset are also unsuitable for trend analysis due to partial instrument failure (May 1970 to April 1976). Note that among existing merged datasets, the SBUV record provides the densest and most spatially uniform sampling over the 2000–2020 time period (Tummon et al., 2015). Moreover, the MOD record merges data from similar SBUV instruments, and its sampling and retrieval characteristics are more homogeneous over the time period we use than other merged records that rely on more varied data sources.

Both total column and profile ozone products used in this study have a 5° horizontal resolution and are zonally averaged to minimize instrumental uncertainty. The ozone profile product provides ozone amounts in Dobson units (DU) for seven vertical layers spanning the mid-to-upper stratosphere, with layer edges near 25, 16, 10, 6.4, 4, 2.5, 1.6, and 1 hPa or approximately 23, 27, 31, 35, 40, 46, 50, and 54 km near the Equator (i.e., 4–6 km resolution, which is comparable to that of other nadir-viewing instruments). Data across these layers have accuracy suitable for long-term trend analysis and compare well with ozone records from other spaceborne and ground-based instruments (within 5%; Bhartia et al., 2013; McPeters et al., 2013). Data outside the 25–1 hPa range tend to be heavily influenced by the a priori ozone climatology used in the retrieval algorithm (Bhartia et al., 1996) and are therefore not relevant to this analysis (see Methods section).

Total column ozone data lie within 1 % of the ground-based instrumental record (McPeters et al., 2013).

2.4 Merged satellite trend estimates

We use trend estimates from the ninth report from the initiative on Stratosphere–Troposphere Processes and their Role in Climate (SPARC) on Long-term Ozone Trends and Uncertainties in the Stratosphere (LOTUS; Petropavlovskikh et al., 2019) to provide context for trends derived using the SBUV record alone. The homogenized satellite record used in the LOTUS report includes six merged datasets:

- two that rely on nadir-viewing instruments with relatively broad averaging kernels and low vertical resolution: the SBUV MOD (v8.6, Frith et al., 2014) and the SBUV cohesive dataset from NOAA, and
- four that rely on limb-viewing instruments and occultation techniques characterized by relatively narrow averaging kernels and high vertical resolution: the Global Ozone Chemistry And Related trace gas Data records for the Stratosphere (GOZCARDS v2.20, Froidevaux et al., 2015), the Stratospheric Water and Ozone Satellite Homogenized database (SWOOSH v2.6, Davis et al., 2016), the SAGE-OSIRIS-OMPS dataset corrected for sampling effects (corr-SAGE-OSIRIS-OMPS, Damadeo et al., 2018), and the SAGE-CCI-OMPS dataset from ESA (Sofieva et al., 2017).

The LOTUS trend estimates are based on multilinear regression and include proxies to correct for the effects of a range of known climate oscillations (El Niño–Southern Oscillation (ENSO), the QBO, the 11-year solar cycle, the Arctic and Antarctic Oscillations, the North Atlantic Oscillation), as well as changes in the stratospheric meridional circulation, tropopause height, stratospheric sulfate aerosol abundances, and long-term trends in chemically reactive halogens (Steinbrecht et al., 2017). A comprehensive list of references for these proxies is provided in the LOTUS report (Petropavlovskikh et al., 2019). Trends are calculated as the unweighted average of trends from each of the six merged datasets. Trend uncertainties account for the degree of dependency between the datasets and the propagation of errors in the regression model (see Sect. 5.3.2 in Petropavlovskikh et al., 2019). In this study, we use trend estimates for the 2000–2020 period, shown in Figs. 8 and 9 for three broad latitude bands.

3 Methods

3.1 Synthetic SBUV observations for the pre-industrial era

In order to quantify the impact of satellite averaging kernels on the observed variability in stratospheric ozone, we create

synthetic observations by applying the SBUV kernels to the ESM4.1 pre-industrial control run. As discussed in the previous section, the model run itself is used as a reference, and the “kernelized” run is used as equivalent observations of the atmosphere had satellites equipped with SBUV sensors flown during the pre-industrial era.

3.1.1 Averaging kernels for synthetic observations

Ozone profiles retrieved by remote sensing are not perfect measurements of the ozone concentrations in each layer of the atmosphere with independent errors but are rather best estimates made given measured radiances and some prior knowledge about the state of the atmosphere. Retrievals are an estimate of some smoothed function of the actual state, with errors that are correlated between different altitudes. As described in Rodgers (1990, 2000), the ozone profile \hat{x} that is retrieved is related to the true ozone profile x and an a priori profile x_a used in its retrieval by

$$\hat{x} - x_a = A(x - x_a) + \epsilon_x,$$

where A is the averaging kernel matrix, and ϵ_x represents random and systematic measurement errors. The i th row of A describes where in the column the information attributed to the i th vertical level actually comes from. Thus, averaging kernels are peaked functions whose half-widths correspond to the vertical resolution of the retrieval. Two examples for SBUV are shown in Fig. 2 (also see Fig. 2 in Kramarova et al., 2013a, for more examples). The blue line shows the row of the kernel matrix that corresponds to the 10.1–6.4 hPa layer, and the red line shows the same for the 101.3–63.9 hPa layer. The averaging kernel for the upper level has a peak at the target level, so the concentration attributed to the 10.1–6.4 hPa layer is the weighted average of concentrations between about 2 and 25 hPa but with the largest weight at the target level. The averaging kernel that is intended to represent the 101.3–63.9 hPa layer has a peak around 50 hPa and exhibits large sensitivity to tropospheric ozone, as discussed in Bhartia et al. (2012). Visualizing the kernels helps to demonstrate the limits of SBUV ozone retrievals in the lower stratosphere compared to the middle to upper stratosphere; for vertical levels below 25 hPa, averaging kernels have broad peaks that are significantly shifted away from their respective target levels, indicating that the retrieval for those levels heavily relies on a priori information (Rodgers, 2000). For this reason, as stated in Sect. 2, the analysis with SBUV is limited to pressure levels 25.45 hPa through 1.013 hPa (levels 9 through 15 in the MOD files) and to the total column. Kernel matrices are produced for each retrieval but are only available as monthly averages. We use the monthly averaged kernels for 1 representative year, chosen to be 2005. For simplicity and to reduce the computational cost of the study, we only use the SBUV kernels, even though the MOD record includes OMPS retrievals; note that MOD relies at least in part on SBUV until March 2018.

A priori information typically consists of a climatology obtained via independent measurements. For instance, a priori states for the SBUV merged dataset come from a combination of independent satellite retrievals and reanalysis data validated by comparison with balloon-borne ozonesondes from the Southern Hemisphere Additional Ozonesondes (SHADOZ) network (Ziemke et al., 2021). However, no such independent climatology exists when it comes to creating synthetic retrievals using a pre-industrial model run. Synthetic a priori information is therefore created using the model monthly mean state itself; that is to say, the averaging kernels are applied to the departure of the model state from its monthly climatology. The synthetic a priori is therefore unbiased by construction, and the synthetic retrievals represent a best-case scenario. As a consequence, variability in the synthetic retrievals may be underestimated, but this approach has the added benefit of isolating the effects of the SBUV kernels on the retrieved variability in stratospheric ozone from any effect of the a priori. The relationship between the synthetic retrieval and the simulated ozone (the reference or “true” quantity) is written as

$$\hat{x} = \bar{x} + A(x - \bar{x}) + \epsilon_x,$$

where \bar{x} denotes the monthly averaged zonal mean simulated ozone profile. As in Rodgers and Connor (2003), the simulated ozone profile is first interpolated onto the coarser SBUV vertical grid so that it can be convolved with the SBUV averaging kernel matrix to produce synthetic retrievals. While the vertical interpolation leads to a loss of some of the higher-resolution information from the model, it is necessary in order to isolate the effects of the averaging kernels, which are illustrated in Fig. 2. The synthetic profile appears smoother than the true profile (hence the expression “smoothing” errors) and is closer to the a priori than the true profile is. Below 25 hPa in the lower stratosphere, the averaging kernels become broader and the retrieval’s quality degrades accordingly, hindering the accurate retrieval of large vertical gradients in ozone concentration. Note that the SBUV kernels are not defined during polar night, which means that the synthetic SBUV data effectively capture the seasonal dependence of the SBUV sampling.

3.1.2 Consideration of known sources of errors

Random errors (e.g., instrumental noise) in retrieved SBUV profiles are on the order of 1 % (Natalya A. Kramarova, personal communication, 26 January 2023), but since retrievals are averaged zonally and monthly across ~ 1000 profiles, random errors in the MOD product are neglected, consistent with Kramarova et al. (2013b). This substantial sample size also allows us to ignore the effects of the spatial sampling of SBUV platforms: monthly zonal means in the pre-industrial model are equated with monthly zonal means in the SBUV dataset.

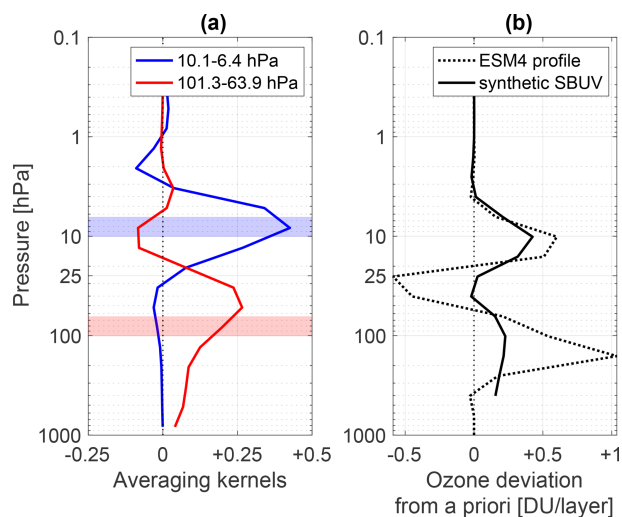


Figure 2. (a) Two rows of the averaging kernel matrix and their corresponding SBUV layers (see legend) for 47.5° N in July 2005, normalized using the SBUV a priori as in Kramarova et al. (2013a) (their Eq. 3). (b) One example of a model profile and its synthetic SBUV counterpart for the same time and location, shown as deviations from the model a priori profile to emphasize vertical structures; in this case, maximum deviations are on the order of 1 % of the a priori above (i.e., at pressures lower than) 25 hPa, where the a priori profile peaks, and 1 %–10 % below. The model profile is interpolated onto the SBUV vertical grid to isolate the effects of the averaging kernels.

Several non-random error sources in ϵ_x produce latitude- and month-dependent biases; these include errors in a priori profiles as well as calibration-related errors in atmospheric attenuation estimates (N values; see Bhartia et al., 2013). However, errors that normally affect a priori states are not relevant to this study since the a priori states are known with certainty (see Sect. 3.1.1). In addition, since we analyze decadal trends, data are de-seasonalized, and calibration errors that normally affect the seasonal cycle can thus be ignored. Other more sporadic sources of errors, including polar mesospheric clouds, volcanic aerosols, ash, and smoke, are also neglected. Accounting for errors resulting from the merging of different SBUV records and biases from non-uniform temporal sampling (orbital drift) is also challenging because of the nature of our analysis. Therefore, such errors are also neglected.

Given these considerations, the error term ϵ_x is considered negligible. What differs between the synthetic retrievals and the simulated ozone is their respective evolution over the course of the 500 years of simulation – any differences being attributable to the averaging kernels. The impact of these differences on ozone variability is shown in Fig. 3: decadal to multidecadal variability is misrepresented throughout the upper and middle stratosphere, with errors in integrated power frequently exceeding 25 % and occasionally approaching or exceeding 100 %. Unless these errors are explicitly accounted for, they will affect the apparent vari-

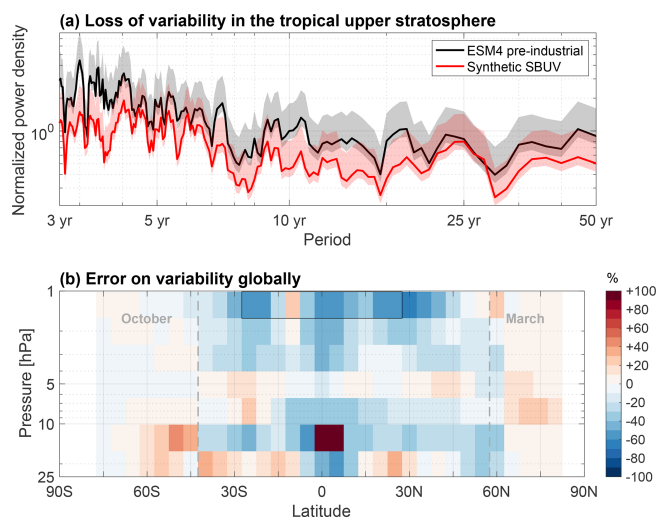


Figure 3. Effect of the SBUV averaging kernels on decadal to multidecadal ozone variability visualized (a) locally for the topmost SBUV layer, 1.6–1 hPa, between 27.5° S and 27.5° N (see the black box in panel b), and (b) globally, as the percent relative error in integrated power density between the pre-industrial run and the synthetic SBUV time series for periods of 10–60 years. Results for only October or March are shown in the high-latitude regions where SBUV data are not available year-round.

ability around SBUV trend estimates, and by extension the degree of statistical significance associated with those trend estimates in a general signal-to-noise-ratio sense. Regions in blue indicate where the statistical significance of trends will be overestimated since internal variability is underrepresented there. Regions in red indicate where statistical significance is underestimated; for instance, between 0 and 2.5° N in the 16.4–10 hPa layer, the variability around trends is inflated because the averaging kernels assign variability from adjacent layers (see Fig. 5a) to this layer.

3.1.3 Removal of known modes of variability

As in other studies (e.g., Steinbrecht et al., 2017; Petropavlovskikh et al., 2019), an attempt is made to remove the contributions of known modes of variability to the time series of ozone in order to better isolate the internal variability. Several known modes of variability are frequently considered: the seasonal cycle, ENSO, and the QBO on interannual timescales and the solar cycle of irradiance on decadal timescales. Solar irradiance is constant in the pre-industrial run, so no removal of that variability is needed. The seasonal component of variability is modeled as the long-term average ozone abundance for each month of the year and is subsequently removed. The Arctic, Antarctic, and North Atlantic Oscillations were found to have negligible impacts on trends (Petropavlovskikh et al., 2019) and are therefore ignored. Remaining ozone variations are then modeled using the following regression model, similar to that used by Stolarski et al.

(1991):

$$\Delta O_3(t) = \alpha QBO_1(t) + \beta QBO_2(t) + \gamma ENSO(t) + \epsilon(t),$$

where α , β , and γ are parameters determined using regression analysis. Unexplained ozone variability is represented by $\epsilon(t)$. $QBO_1(t)$ and $QBO_2(t)$ are the two leading empirical orthogonal functions (EOFs) of the de-seasonalized zonal mean monthly zonal wind between 10 and 70 hPa from the model simulations (see Fig. A1), as in Wallace et al. (1993). Following Oman et al. (2013), $ENSO(t)$ is based on the NOAA Oceanic Niño Index (ONI) and captures variations in (modeled) sea surface temperature in the Niño 3.4 region (5°S – 5°N , 170 – 120°W) in the central Pacific Ocean, lagged by 2 months as in Randel et al. (2009). In the pre-industrial model run, $QBO_1(t)$ and $QBO_2(t)$ account for 68 % of the normalized variance of the zonal wind time series. By construction, the correlation between the two EOFs is exactly zero. Thanks in part to the substantial length of the pre-industrial run, the correlation between the QBO EOFs and $ENSO(t)$ is very small (less than 0.05), yielding virtually independent regression terms.

The zonal wind data in Fig. A1 clearly show that the model's internally generated QBO is far from realistic: the weak QBO-like signal is confined to lower pressures, is much more prone to disruption, propagates vertically at more variable rates, and exhibits higher frequencies than the observed QBO. Given these differences, the appropriateness of the traditional EOF-based QBO removal could be called into question, but the model does produce a semi-periodic downward-propagating signal reminiscent of the QBO. Such signals produce large retrieval errors (1 %–6 % in observations) when convolved with the averaging kernels (Kramarova et al., 2013a). Since errors of this nature are the focus of this study, we use the traditional EOF-based approach to removing QBO variability. The QBO removal is performed after the simulated ozone is convolved with the SBUV averaging kernels, to mimic the residual errors produced by the removal of the QBO in SBUV observations.

While the use of EOFs allows us to account for variability due to changes in the amplitude and phase of zonal wind oscillations, it does not account for variability due to changes in the frequency of the oscillations. This is slightly problematic for the removal of the QBO signal in ozone since its frequencies differ between the Equator (where the EOFs are extracted) and higher latitudes (Tung and Yang, 1994). Nevertheless, the average contribution of the QBO to near-equatorial ozone variability is estimated to be about 0.5 % in the model (much smaller than the estimated 10 %–20 % in observations; see Brunner et al., 2006) and is expected to be even smaller at extratropical latitudes, where the atmospheric jets produce large variability. The EOF-based removal of QBO variability is therefore considered sufficient for the purposes of this study.

The methods discussed in this section provide a set of synthetic SBUV data in the absence of anthropogenic forcings

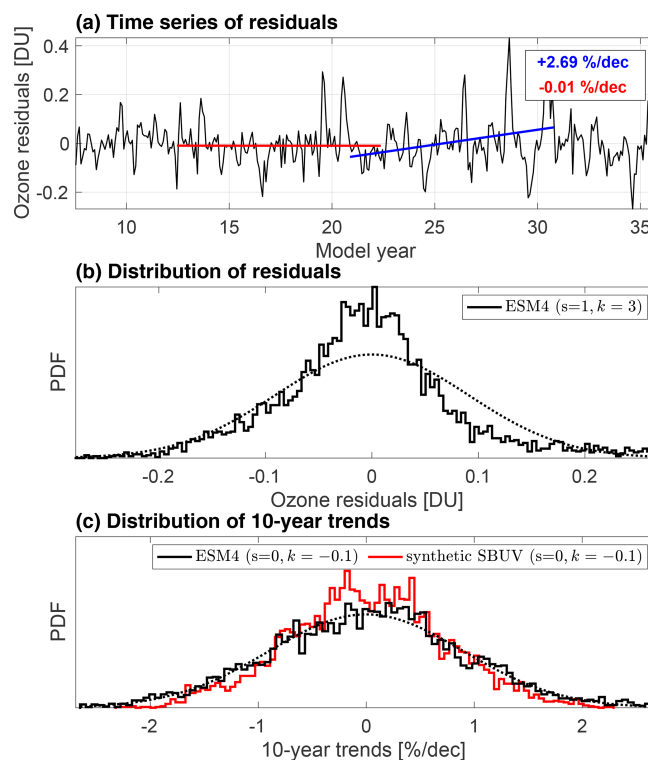


Figure 4. (a) Sample time series of pre-industrial ozone residuals at 2.5–1.6 hPa, 42.5°S , with 10-year trend examples calculated in percent of the mean ozone value at that location, ~ 4.5 DU. (b) Distribution of ozone residuals for the 500 years of pre-industrial simulation and (c) the corresponding reference distribution of 10-year unforced trends. Dashed lines show Gaussian distributions with moments identical to those from the model distributions for reference. The skewness and excess kurtosis of the distributions are provided in the legend.

and without known climate oscillations. We use this synthetic record to quantify the internal variability in the climate system relevant to stratospheric ozone, so as to estimate the lead time required for long-term trends in ozone to emerge in observations against the backdrop of internal variability. For simplicity, we assume that the modern climate system produces the same internal variability as that during the pre-industrial era. However, while this assumption has been widely used in the analysis of climate model output (e.g., IPCC, 2001; Hegerl et al., 2007; Deser et al., 2012), evidence showing that it may not hold for some variables in climate models exists (e.g., Schär et al., 2004; Scherrer et al., 2005; Rodgers et al., 2021), although different models exhibit different responses in their internal variability under forced change (Maher et al., 2021). Therefore, we suggest caution in interpreting statistical confidence.

3.2 Time of emergence of linear trends

Generally speaking, the smaller the trend, the longer it takes to become distinguishable from background noise. Several methods have been employed in the past to calculate the time of emergence of linear trends. Emergence has sometimes been said to occur when the ratio of the signal (trend) to the noise (internal variability) is large or exceeds a pre-determined threshold (Madden and Ramanathan, 1980; Wigley and Jones, 1981; Giorgi and Bi, 2009; Hawkins and Sutton, 2012). Such methods do not provide the formalism needed for a measure of statistical confidence that links the robustness of the results to the choice of the threshold to exceed. Another method (Mahlstein et al., 2011) defines the time of emergence as that when a statistical test deems an observation unlikely to belong to a distribution that represents an unperturbed climate state. A problem with this method, however, is that it does not account for autocorrelation in the time series, which means that the time of emergence for time series with nonzero autocorrelation is underestimated. Other approaches do take autocorrelation into account; examples include the methods of Tiao et al. (1990), Weatherhead et al. (1998), and Li et al. (2017). In the latter, emergence occurs when a predicted confidence interval of a cumulative trend excludes zero. The confidence interval is described analytically based on the work of Thompson et al. (2015) and accounts for lag-one autocorrelation. However, the method requires the residuals representing internal variability to be normally distributed. This is generally not the case for stratospheric ozone: after removing the seasonal cycle and the contributions from ENSO and the QBO, distributions of residuals often have heavier tails and a narrower peak than a normal distribution (positive excess kurtosis) and are not symmetric (skewness is nonzero; see an example in Fig. 4b).

3.2.1 Definition of the time of emergence

For these reasons, the time of emergence is calculated using the method of Rivoire et al. (2024) (R24 hereafter), which provides nearly identical results to that of Li et al. (2017) but can handle non-normally distributed residuals. R24 found empirical evidence that the method of Li et al. (2017) may still be used when residuals exhibit nonzero skewness or excess kurtosis, but without any formalism to prove this generally, we consider the method of R24 to be a safer choice. The general principle behind the method is to compare a linear trend of interest, denoted b , to a distribution of linear trends that arise purely as a result of internal variability. If b is located near the central quantiles of the reference distribution, then b aligns with typical fluctuations seen in the absence of external forcings. Conversely, if b is located near the outer quantiles, then it is unusually large compared to natural fluctuations and can reasonably be hypothesized to be the result of external forcings.

The reference distribution is calculated by sampling linear trends in the control simulation using a sampling window of fixed length (see Fig. 4a and c). The time of emergence, denoted y , is the length of the sampling window that yields a reference distribution such that $|b|$ exceeds $c = \frac{c_d+100}{2}$ % of unforced trends, with c_d as the statistical confidence threshold for a two-sided test. In the rest of the paper we use $c_d = 95$ % as the threshold for statistical significance. Although the use of this particular threshold is commonplace, we emphasize its arbitrary nature and caution against overinterpreting its significance.

In practice, the time of emergence is determined by iteration, starting with the shortest possible sampling window of two model time steps. The reference distribution of unforced trends over all possible two-step intervals (with overlap) is calculated, and its c th percentile is compared to $|b|$. The length of the sampling window is then progressively increased until the statistical criterion described above is met. For more details on this technique and its efficient implementation, the reader is referred to Sect. 2 of R24.

Since the time of emergence y is calculated using a trend distribution sampled directly from a model run, it neglects the effects of observational uncertainties. Figure 5 shows this “ideal” time of emergence – which is valid only if one possesses omniscient knowledge of the atmosphere – for both a homogeneous trend and the actual trends from CCM1-1 simulations. R24 shows that observational uncertainties can introduce large errors in y , and they provide the formalism to correct y accordingly. In our case, the correction for y accounts for the effects of the vertical redistribution of ozone variability by the averaging kernels (see Figs. 2 and 3). The corrected time of emergence is denoted y^* and is calculated using reference trend distributions sampled from the synthetic SBUV retrievals introduced earlier (e.g., the red curve in Fig. 4c) rather than the model run itself (the black curve in Fig. 4c). As a result, y^* quantifies the time of emergence of ozone trends as seen by SBUV platforms (see Sect. 4.2 and Fig. 6).

3.2.2 Notes on the realism of the pre-industrial control simulation

Since emergence is defined based on the representation of internal variability from a model, the realism of the model comes into question. On the one hand, the time of emergence y is subject to the realism – or lack thereof – of the model: misrepresentations in the magnitude or frequency spectrum of internal variability yield errors in the time of emergence. Although we provide an analysis of said realism to the extent possible (Sect. 2.1), quantifying these errors is inherently difficult given the absence of pre-industrial references for observed variability. On the other hand, R24 showed that the adjustment for the effect of observational uncertainties on the time of emergence, when expressed as a relative error ($\frac{y^*-y}{y}$), is accurate even if the model dramatically misrep-

resents internal variability. For this reason, we quantify the effect of the SBUV kernels on the time of emergence as a relative error (see Sect. 4.2).

4 Results

4.1 Time of emergence

Figure 5 provides estimates for the time of emergence (y , in years since 2000) with 95 % confidence for both a spatially homogeneous 1 % per decade trend (panel b) and the 2000–2020 trends from CCMI-1 (panel c). To put this into context, panel (a) shows the average ozone abundances and their standard deviation in the pre-industrial run. The spatially homogeneous trend serves to illustrate the effects of internal ozone variability on the time of emergence. Except for the lowermost stratosphere, a general correlation is seen between the time of emergence and ozone variability as quantified by its standard deviation: the largest values are found in the tropical middle-stratosphere and polar regions. Departures from this correlation arise where the spectrum of variability includes non-Gaussian behavior and is not well described by the standard deviation alone.

According to our method and based on ESM4.1 as a reference for internal variability, trends associated with the CCMI-modeled evolution of ozone under the Montreal Protocol may not emerge until the second half of the century in the heart of the midlatitude ozone layer (panel c). In the upper stratosphere (above 10 hPa), chemistry–climate model trends for 2000–2020 are largely distinguishable from ESM4.1's natural variability with ~ 20 years of model record, in agreement with the previous literature showing significantly increased ozone abundances (Gillett et al., 2011; Arblaster et al., 2014). Lower down, especially in the middle stratosphere (10–30 hPa), our method finds that trends have generally not emerged yet. In the midlatitudes (35–60° N/S), trends may require decades of additional record to emerge, owing in part to large variability and in part to small trends. This finding is generally consistent with the existing literature, which finds little to no significance to trends in the region (WMO, 2022, and references therein). Over the Arctic, we find that the emergence of positive trends could occur by 2030, compared to the literature expecting visible signs of recovery there in the mid-2040s. However, since our results are based on the analysis of a single model, we urge caution: previous analyses have concluded that large dynamical variability in this region precludes the detection of recovery until the late 21st century (WMO, 2022). It is possible that the ozone interannual variability in ESM4.1 still lacks realism despite recent improvements (Horowitz et al., 2020, their Fig. 11).

For vertically integrated ozone (panel d), despite the relatively small variability in the tropical region, emergence occurs later there because trends are small. Trends associated with chemical loss over the polar caps in spring are isolated using October- and March-only time series in the South-

ern Hemisphere and Northern Hemisphere, respectively. Although ozone loss is largest over Antarctica and is therefore expected to exhibit the clearest signs of recovery (Newchurch et al., 2003; Yang et al., 2008; Charlton-Perez et al., 2010), the analysis predicts that trends for October and March only emerge after ~ 40 years near the strongest trends (65° N/S, not shown), owing to large variability. Closer to the poles, October and March trends become very small and therefore take much longer to emerge.

Some caveats should be kept in mind: the time of emergence is based on a representation of unforced ozone variability that excludes the effects of ozone-depleting substances or volcanic aerosols (see Fig. 1). As a result, time of emergence estimates provide lower bounds for the timing of the detection of recovery (as long as the linear trend approximation holds). Further, in the previous literature, statistically significant trends associated with the recovery of the ozone layer are based on more than just ozone abundances; they include trends in other metrics, such as the minimum in 15 d average total ozone, maximum in 15 d average 220 DU ozone hole area, minimum in 15 d average ozone mass deficit, and the duration of the ozone hole (e.g., Tully et al., 2020). It is possible that these metrics provide more detectable trends than ozone abundances alone. Lastly, we highlight the fact that quantifying the emergence of a trend vs. that of an epoch difference may yield different results.

4.2 Effect of the SBUV kernels on the time of emergence

Results established so far apply to the “model world” only and do not factor in the limitations of observing platforms that make up the historical record. The synthetic SBUV record from Sect. 3 provides an example of the effect of observational limitations on the time of emergence. Figure 6 shows those limitations as the relative difference between the kernel-adjusted time of emergence y^* and the ideal time of emergence y . Large errors throughout the middle and upper tropical stratosphere are reminiscent of the patterns of variability associated with the QBO despite the prior removal of QBO variability (Sect. 3.1.3). This result arises because the convolution of the SBUV kernels with the ozone data affects the representation of the QBO (Kramarova et al., 2013a) but does not affect the wind field that is used to perform the QBO removal. Other error patterns arise where ozone variability is large, for example across the 16–10.1 hPa layer and above the polar caps (refer to Fig. 5a). The unavailability of SBUV retrievals during polar night may contribute to errors over the polar caps. Interestingly, errors are opposite in sign between the two polar caps at some levels, indicating opposite effects due to the combined SBUV averaging kernels and lack of measurements there. Errors are slightly larger over the southern polar cap than the northern, which is attributable to slightly larger internal variability in ESM4.1 in the Antarctic (as shown in Fig. 5a). This result is more specifically at-

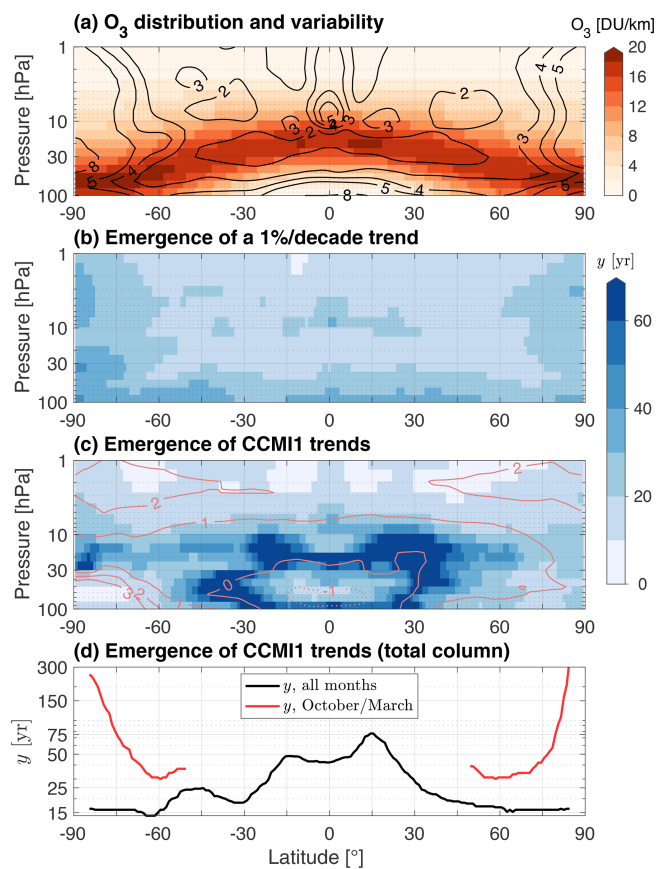


Figure 5. (a) Average ozone abundance in the ESM4.1 pre-industrial run (in DU per km, colors) and relative standard deviation of the de-seasonalized time series (black contours, in percent). (b) Ideal time of emergence (y) for a spatially homogeneous 1% per decade trend in years, with contours every 10 years. (c) Ideal time of emergence for 2000–2020 trends in CCMI-1 ref-C2 (shown by overlaid red contours, % per decade). (d) Ideal time of emergence (y) for 2000–2020 trends in CCMI-1 ref-C2 total column ozone. Results for October only (March only) are shown in the Southern (Northern) Hemisphere.

tributable to the shape of the power spectrum of internal variability in ESM4.1 around periods corresponding to the time of emergence. The largest negative errors (i.e., relative differences) found near the Equator just below 10 hPa are explained by the unusual overestimation of internal variability attributable to the SBUV kernels shown in Fig. 3b. At all latitudes, the error pattern is qualitatively unchanged whether trend magnitudes are homogeneous (panel a) or not (panel b), indicating that our method provides generally applicable understanding of the deficiencies of the SBUV ozone retrieval for detecting trends.

For total column ozone, the kernel-adjusted time of emergence estimates are very close to the ideal estimates (Fig. 6c). This is consistent with the result in Kramarova et al. (2013a) (their Fig. 6): combining SBUV layers improves the accuracy of ozone retrievals.

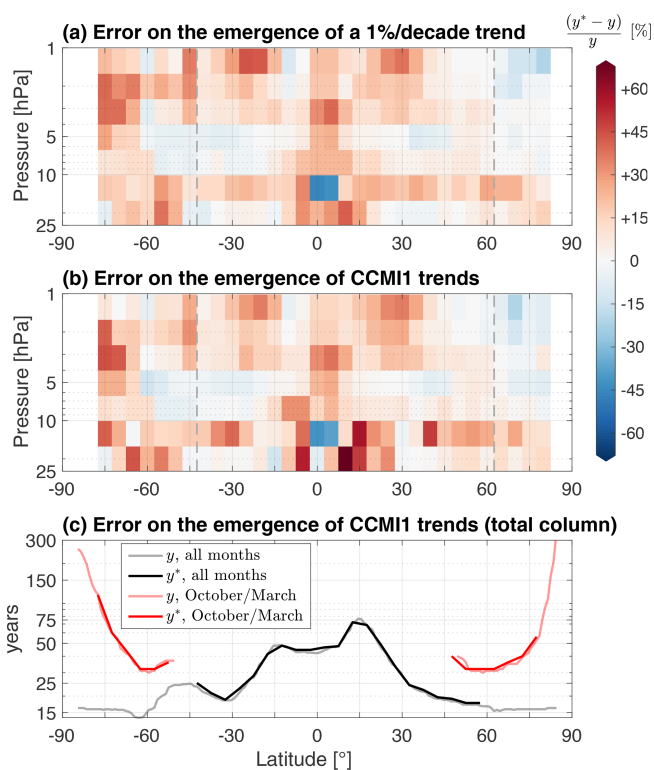


Figure 6. Relative error of the time of emergence due to the SBUV kernels ($\frac{y^* - y}{y}$) for (a) a spatially homogeneous 1% per decade trend and (b) the CCMI-1 trends from panel (c) in Fig. 5. Dashed gray lines denote the latitudes poleward of which the SBUV retrieval is not available year-round. (c) Ideal time of emergence (y) and kernel-adjusted time of emergence (y^*) for 2000–2020 CCMI-1 trends in total column ozone.

4.3 Effect of the SBUV kernels on vertically resolved trend estimates

Retrieval methods based on averaging kernels have previously been said not to affect trend estimates (Petropavlovskikh et al., 2019). The rationale behind this statement is straightforward: since the retrieval process remains unchanged over time, errors attributable to it are also constant over time (at least in a long-term sense), and trend estimates are therefore unbiased. However, this rationale fails to account for the vertical redistribution of information by the kernels. As trends affect the true profile (x in Sect. 3.1.1), the difference between the true profile and the a priori profile ($x - x_a$) also contains a trend signal. Once the kernel matrix is convolved with ($x - x_a$), the trend signal is redistributed vertically. As a result, trend estimates at each vertical level are affected by the trend signal from adjacent levels in a manner defined by the shape of the averaging kernels. This is true of any profile, including the hypothetical case of a vertically uniform profile in the context of trends expressed in percent per unit time, but errors can be especially large when trends are vertically heterogeneous.

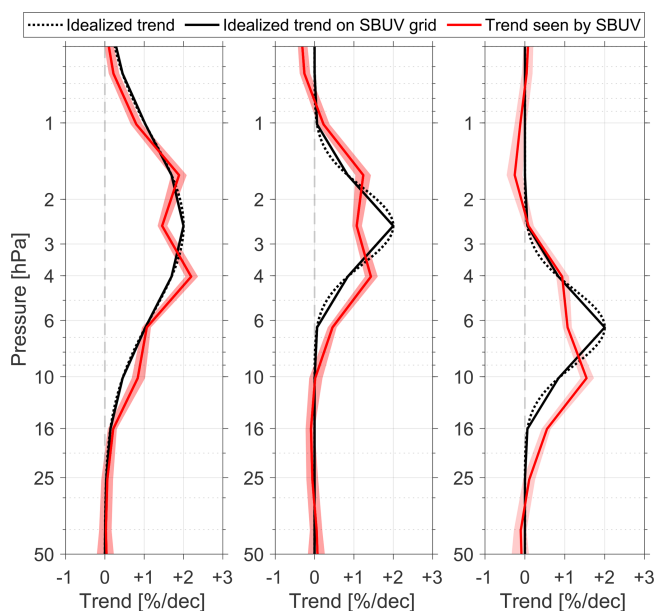


Figure 7. Idealized trend profiles representing three hypothetical scenarios for ozone recovery in the mid-to-upper stratosphere (black) and corresponding profiles that would be observed by SBUV at 42.5° N (red). The red shading shows the interquartile range for trends estimated using 24 years of data (roughly the length of the record since 2000).

Figure 7 shows such errors for a few idealized trend profiles representing possible ozone increases in the mid-to-upper stratosphere in the northern midlatitudes (Godin-Beekmann et al., 2022; see also Figs. 3–10 in WMO, 2022): one broad maximum and one narrower maximum centered near 2.5 hPa and one maximum centered near 6 hPa. The SBUV kernels attenuate peaks in the idealized trend profiles and shift the maxima to other altitudes above and below the true peaks. Errors as large as 100 % and displacements as large as 6 km are found, highlighting the importance of accounting for the retrieval process in estimating trend uncertainties. Negative correlation coefficients between layers in the averaging kernel matrix can yield negative trends even when the true trends are only positive.

The results suggest that trends in the middle and lower stratosphere may be affected by the presence of trends in the upper stratosphere. This is particularly relevant in the context of analyses that show negative trends in the lower stratosphere (e.g., Ball et al., 2018). While the errors near 25 hPa may seem small, they may induce large errors in the time of emergence. Altogether, this analysis highlights the importance of accounting for satellite averaging kernels when analyzing vertically resolved trends.

4.4 Smallest detectable trends

When trend estimates are in disagreement across model simulations or observational records, the method from R24 can

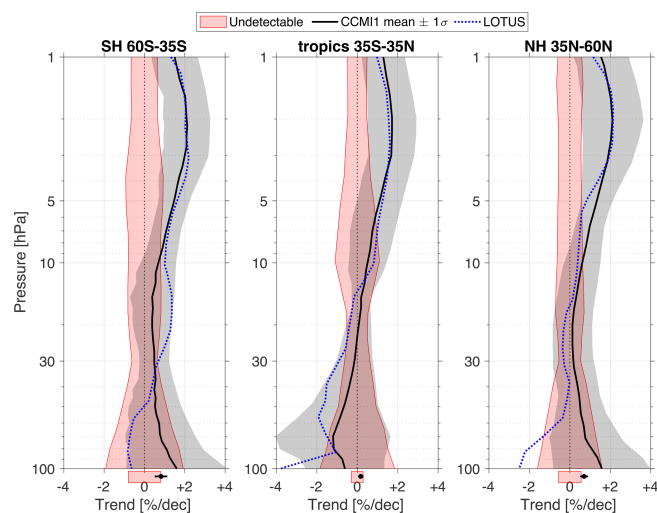


Figure 8. Simulated 2000–2020 ozone trends from CCM1-1 and the corresponding range of undetectable trends at the 95 % confidence level. Results for total column ozone are shown at the bottom outside each panel. Observed 2000–2020 trends from LOTUS are shown in blue for reference.

be used to determine the range of trend magnitudes that should be distinguishable from internal variability in the first place. Using the time of emergence γ as the length of the sampling window used to calculate trends from the control run (or its synthetic counterpart), b is the smallest detectable trend at confidence level c_d . Trends with a magnitude smaller than $|b|$ are not distinguishable from internal variability at the chosen degree of confidence. The magnitude of b provides an “envelope of undetectability”, shown in Fig. 8 for a few key latitude bands, assuming omniscient/model knowledge. Simulated and observed 2000–2020 trends in the upper stratosphere (above 10 hPa) largely lie outside the range of undetectable trends, indicating that the trend estimates have emerged from internal variability there. Interestingly, the inter-model spread only lies outside the envelope above 3–5 hPa. Below 10 hPa, mean trends are located within the range of undetectability, which means that the internal variability is so large that 21 years of data is not sufficient to ascertain whether ozone is increasing or decreasing with 95 % certainty there. Further, simulated and observed trends within the heart of the ozone layer (see Fig. 5a) are currently not detectable with 95 % statistical confidence according to the datasets used and again, assuming omniscient knowledge.

The same analysis is performed on the synthetic SBUV control simulation to obtain the envelope of undetectability accounting for observational uncertainties. Figure 9 overlays this SBUV envelope and that given by the unaltered model run, that is, the ideal envelope of undetectability (from Fig. 8). Overall, SBUV envelopes are slightly optimistic (i.e., narrower, by a few tenths of a percent per decade) because the SBUV kernels tend to reduce the true internal variability (in the latitude-band-average sense; compare to Fig. 3) or af-

fect its spectrum in a way that artificially decreases the time of emergence. The differences between the true and SBUV envelopes (Fig. 9) are most noticeable in the tropical region, where the SBUV kernels yield errors up to 25 %. Note that averaging the smallest trends over latitude bands reduces the SBUV kernel errors – those can be as large as 50 % locally (not shown). As was concluded from Fig. 6, the detectability of total column ozone trends is virtually unaffected by the SBUV kernels.

The SBUV envelope is most relevant to trends estimated using the SBUV record, and it provides useful context for comparative studies that include SBUV trends (e.g., Frith et al., 2014; Harris et al., 2015; Tummon et al., 2015; Solomon et al., 2016; Chipperfield et al., 2017; Steinbrecht et al., 2017; Weber et al., 2018; Ball et al., 2018; Brönnimann, 2022). We draw a comparison with the LOTUS trends since LOTUS is subject to the limitations of SBUV, among those of other datasets. The results are qualitatively similar to those in Fig. 8: trend estimates above 10 hPa should generally be distinguishable from internal variability, with the caveat that errors from the SBUV kernels can yield moderate overconfidence where trend estimates are close to the smallest detectable trends. Note that the large differences between SBUV-only and LOTUS trends are likely due to the heavy reliance of the LOTUS product on limb sounder retrievals, which have narrower averaging kernel peaks than those shown in Fig. 2, resulting in an ~ 2.5 km intrinsic vertical resolution in the lower stratosphere (compared to 6 to 10 km for SBUV). Below 10 hPa, most trends are predicted to be indistinguishable from internal variability.

Similar detectability envelopes could be quantified for other nadir sounders, for instance the Global Ozone Monitoring Experiment (GOME; Burrows et al., 1999), Ozone Monitoring Instrument (OMI; Levelt et al., 2006), or Tropospheric Emission Spectrometer (TES; Worden et al., 2004). Limb sounders are subject to similar errors given the extent that they also rely on averaging kernels, for instance OMPS-LP (Arosio et al., 2018), the Optical Spectrograph and Infrared Imager (OSIRIS; von Savigny et al., 2003), Atmospheric Chemistry Experiment Fourier Transform Spectrometer (ACE-FTS; Walker et al., 2005), Michelson Interferometer for Passive Atmospheric Sounding (MIPAS; Ridolfi et al., 2000), Microwave Limb Sounder (MLS; Livesey et al., 2006), and Stratospheric Aerosol and Gas Experiment III (SAGE III; Cisewski et al., 2014). However, since limb sounders typically achieve a much better vertical resolution than nadir sounders, kernel errors are likely to affect the detectability of trends to a lesser extent.

The results discussed here raise the following question: if not now, when will trend estimates eventually become distinguishable from internal variability? To answer this question, Fig. 10 shows the magnitude of the smallest detectable trends as a function of the length of the record. In the upper stratosphere (~ 1 hPa, Fig. 10a) where CCMI-1 trend estimates are 1 % per decade–2 % per decade, the current length

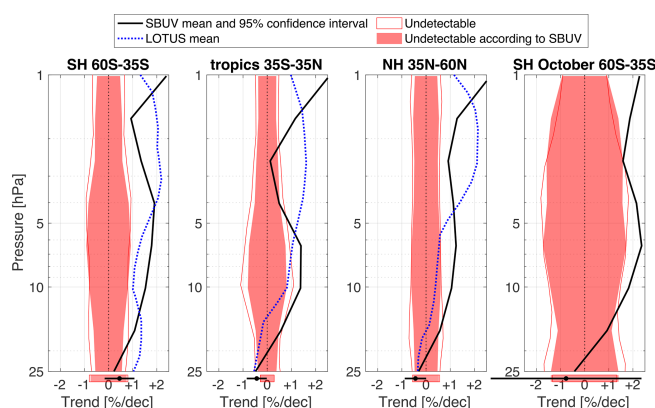


Figure 9. SBUV 2000–2020 ozone trends and the corresponding envelope of undetectability at the 95 % confidence level, including the adjustment for SBUV kernels, which tend to yield slightly narrower envelopes. Results for total column ozone are shown at the bottom outside each panel. Observed 2000–2020 trends from LOTUS are shown in blue for reference. The confidence interval around SBUV trends comes from linear regression coefficient estimates.

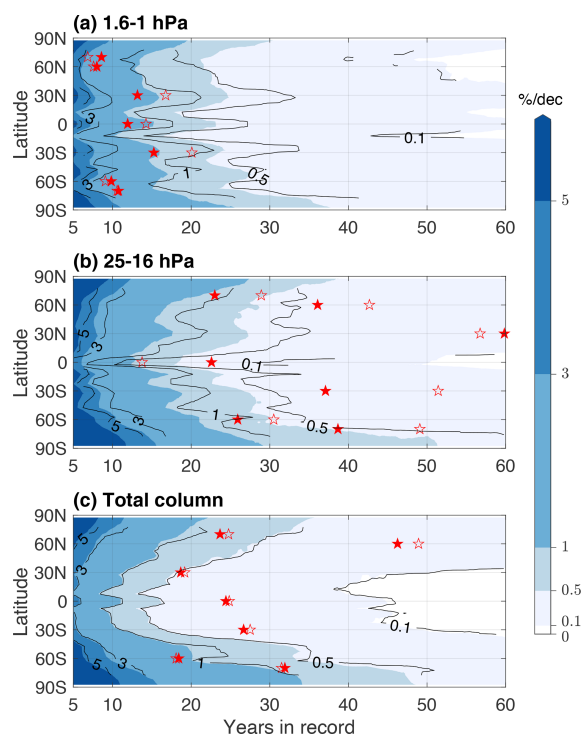


Figure 10. The smallest detectable ESM4.1 pre-industrial trends at the 95 % confidence level (shaded) and their SBUV kernel-adjusted counterparts (black contours) given as a function of the length of the available record for (a) 1.6–1 hPa, (b) 25–16 hPa, and (c) the total column. Filled and unfilled red stars indicate the emergence of CCMI-1 2000–2020 trends at a few latitudes based on the ESM4.1 smallest detectable trends and their SBUV kernel-adjusted counterparts, respectively.

of record since 2000 is sufficient to rule out the possibility that the observed increase in ozone is an artifact of internal variability, even when accounting for SBUV kernel errors. This result is in agreement with Fig. 5c and arises despite the relatively large underestimation of internal variability shown between 60°S and 60°N in Fig. 3b. Lower down (25–16 hPa, Fig. 10b), CCMI-1 trends are about a factor of 2 smaller, and ozone variability is larger; therefore, much longer records are needed. In the Northern Hemisphere midlatitudes, 35 to 60 years of record are needed. Over the southern polar cap, almost 40 years of data are needed, increasing to 50–60 years when accounting for SBUV kernel errors. In contrast, the SBUV kernels yield overestimated internal variability near the Equator (Fig. 3b), and accounting for such errors therefore decreases the length of the record needed to detect the CCMI-1 trend estimate (the unfilled red star is shifted toward earlier years than the filled one). For total column ozone (Fig. 10c), the results suggest that CCMI-1 trends have emerged or will soon, except in the Northern Hemisphere midlatitudes and over Antarctica. SBUV kernel errors have little effect on these estimates, consistent with previous findings (Kramarova et al., 2013a). We emphasize that these results are based on the analysis of a single model (ESM4.1) and that the definition of emergence used here differs from typical measures of statistical significance based on the standard error in regression coefficients. These caveats are important to keep in mind, given that observational studies have found signs of recovery in some column ozone metrics above Antarctica (WMO, 2022, Chap. 4).

5 Conclusions

We examined the statistical significance of long-term trends in stratospheric ozone. Statistical significance was tested against the null hypothesis of no *forced change* rather than the null hypothesis of no change, in order to include information about the magnitude of trends that arise purely as a result of internal variability in the climate system. Using the concept of time of emergence, we showed the potential for large errors in the statistical significance of trends estimated from satellite records: trends may appear to emerge from internal variability when in fact they are not robust and will require several more years to decades of observations to emerge. Two factors explain these results: (1) the averaging kernels inherent to the optimal estimation retrieval technique consistently misrepresent decadal to multidecadal variability in nadir-viewing satellite observations, and (2) known modes of variability (ENSO, QBO) interact with averaging kernels, and their removal from the observed record is therefore prone to residual errors.

Further, our analysis showed that by vertically redistributing information, averaging kernels can alter trend magnitudes, in addition to their uncertainties. This result contrasts with the assumption of Petropavlovskikh et al. (2019), who

stated that kernels do not affect trends since kernel errors are constant in time. We find instead that the shape of the vertical trend profile is subject to large errors: hypothetical scenarios of ozone recovery at a rate of 2 % per decade in the mid-to-upper stratosphere yield errors up to 1 % per decade, with possible vertical displacements in the altitude of local maxima of up to 6 km, consistent with Fig. 3–10 in WMO (2022). We also find errors in the sign of trends, with location and magnitude depending on the hypothetical profile used. Our analysis uses SBUV as an example, but other nadir sounders (e.g., GOME, OMI, TES) are in principle subject to similar effects, with magnitude dependent on their vertical resolution and the location of their vertical levels relative to vertical gradients in ozone concentrations and variability. Limb sounders (e.g., MLS, OMPS-LP, OSIRIS, ACE-FTS, MIPAS, SAGE III) may also be affected, albeit to a lesser extent.

We note a few caveats to our study. Time of emergence estimates are subject to the realism of the model we used (ESM4.1), although the relative adjustment to the time of emergence remains accurate. The reference simulation we used excludes the influence of volcanoes, which can be large. Additionally, the effects of non-uniform sampling and record-merging procedures are also excluded. The time of emergence estimates presented here should therefore be construed as lower bounds (as far as linear trends are concerned). Our time of emergence values may be further underestimated if the internal variability in the climate system is increasing under forced change (as discussed by Rodgers et al., 2021). In addition, we used an arbitrary 95 % threshold for statistical significance, and quantitative results are of course sensitive to this choice. Lastly, in this study we did not explore derived metrics related to ozone recovery (e.g., size and timing of the ozone hole; WMO, 2022, Chap. 4, Sect. 4.4.2.1), which may confer different detection power with respect to long-term trends.

Our results should be interpreted with these caveats in mind, but nonetheless, they provide useful context in light of the varying degrees of confidence placed in trends in the recent literature; notably, adding just a few years to the historical record can change the magnitude and even the sign of trends in some locations. Based on our results, we recommend systematically accounting for the effects of averaging kernels when calculating long-term trends using vertically resolved ozone records, particularly those from nadir-viewing instruments. We further recommend testing the statistical significance of trends against the null hypothesis of no *forced change* rather than no change at all. Future work in this direction would benefit from long reference simulations of chemistry–climate models with perpetual year-2000 conditions, to better characterize the variability relevant to the recovery of the ozone layer since then.

Appendix A: Representation of the QBO in the pre-industrial simulation

Figure A1 highlights differences between the QBO generated internally by the ESM4.1 pre-industrial model run and the observed QBO (see discussion in Sect. 3.1.3).

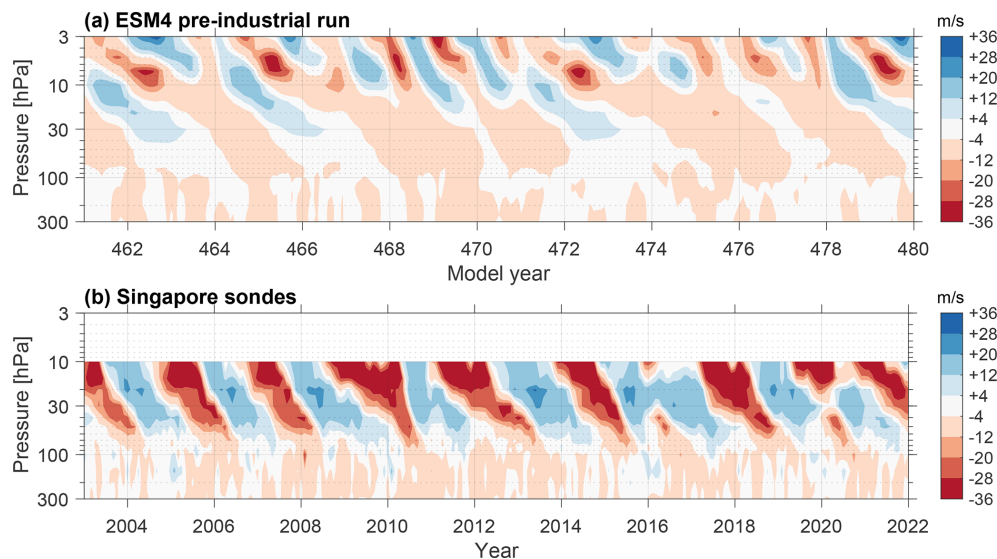


Figure A1. Sample time series of the simulated (a) and observed (b) de-seasonalized monthly zonal mean zonal wind. Positive values denote westerly wind anomalies.

Data availability. The SBUV merged ozone record is available at https://acd-ext.gsfc.nasa.gov/Data_services/merged/ (NASA Goddard Space Flight Center, 2025). The climate simulations from the Geophysical Fluid Dynamics Laboratory (GFDL) ESM4.1 model are available from the CMIP6 archive at <https://doi.org/10.22033/ESGF/CMIP6.8669> for the pre-industrial control run (Krasting et al., 2018a) and at <https://doi.org/10.22033/ESGF/CMIP6.8597> for the historical run (Krasting et al., 2018b). The CCM1 model archive is available at <http://catalogue.ceda.ac.uk/uuid/9cc6b94df0f4469d8066d69b5df879d5/> (Hegglin et al., 2015).

Author contributions. LR and ML designed the methods. LR performed the analysis and produced the figures. The text was primarily written by LR with feedback from all co-authors. This project was initiated by ML, JN, and PL (poster titled “Uncertainty in Ozone Trend Detection” presented at the American Meteorological Society’s 2020 annual meeting).

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher’s note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors.

Acknowledgements. We thank Natalya Kramarova (Goddard Space Flight Center) for help with applying the SBUV kernels. Ozone trend estimates from CCM1-1 were kindly prepared by Kleareti Tourpali (Aristotle University of Thessaloniki). We acknowledge discussions that shaped the analysis with Irina Petropavloskikh, Peter Zoogman, Aaron Match, Susann Tegtmeier, and Kane Stone, as well as early contributions to this project by Mary Frances Connors. Louis Rivoire was funded by the William F. Milton Fund and National Aeronautics and Space Administration (NASA) award 80NSSC21K0943. Marianna Linz was funded by NASA awards 80NSSC21K0943 and 80NSSC23K1005. Pu Lin was funded by award NA18OAR4320123 from the National Oceanic and Atmospheric Administration (NOAA), US Department of Commerce. Work at the Jet Propulsion Laboratory, California Institute of Technology, was done under contract with NASA (80NM0018D0004).

Financial support. This research has been supported by the National Aeronautics and Space Administration (grant no. 80NM0018D0004), the National Aeronautics and Space Administration (grant nos. 80NSSC21K0943 and 80NSSC23K1005), the National Oceanic and Atmospheric Administration (grant no. NA18OAR4320123), and the William F. Milton Fund.

Review statement. This paper was edited by Farahnaz Khosrawi and reviewed by three anonymous referees.

References

- Abalos, M., Calvo, N., Benito-Barca, S., Garny, H., Hardiman, S. C., Lin, P., Andrews, M. B., Butchart, N., Garcia, R., Orbe, C., Saint-Martin, D., Watanabe, S., and Yoshida, K.: The Brewer–Dobson circulation in CMIP6, *Atmos. Chem. Phys.*, 21, 13571–13591, <https://doi.org/10.5194/acp-21-13571-2021>, 2021.
- Arblaster, J. M., Gillett, N. P., Calvo, N., Forster, P. M., Polvani, L. M., Son, W. S., Waugh, D. W., Young, P. J., Barnes, E. A., Cionni, I., Garfinkel, C. I., Gerber, E. P., Hardiman, S. C., Hurst, D. F., Lamarque, J.-F., Lim, E.-P., Meredith, M. P., Perlwitz, J., Portmann, R. W., Previdi, M., Sigmond, M., Swart, N. C., Vernier, J.-P., and Wu, Y.: Stratospheric ozone changes and climate, in: *Scientific assessment of ozone depletion: 2014*, World Meteorological Organization, 416 pp., <https://csl.noaa.gov/assessments/ozone/2014/report/2014OzoneAssessment.pdf> (last access: 7 February 2025), 2014.
- Arosio, C., Rozanov, A., Malinina, E., Eichmann, K.-U., von Clarman, T., and Burrows, J. P.: Retrieval of ozone profiles from OMPS limb scattering observations, *Atmos. Meas. Tech.*, 11, 2135–2149, <https://doi.org/10.5194/amt-11-2135-2018>, 2018.
- Austin, J., Horowitz, L. W., Schwarzkopf, M. D., Wilson, R. J., and Levy, H.: Stratospheric ozone and temperature simulated from the preindustrial era to the present day, *J. Climate*, 26, 3528–3543, <https://doi.org/10.1175/jcli-d-12-00162.1>, 2013.
- Ball, W. T., Alsing, J., Mortlock, D. J., Rozanov, E. V., Tummon, F., and Haigh, J. D.: Reconciling differences in stratospheric ozone composites, *Atmos. Chem. Phys.*, 17, 12269–12302, <https://doi.org/10.5194/acp-17-12269-2017>, 2017.
- Ball, W. T., Alsing, J., Mortlock, D. J., Staehelin, J., Haigh, J. D., Peter, T., Tummon, F., Stübi, R., Stenke, A., Anderson, J., Bourassa, A., Davis, S. M., Degenstein, D., Frith, S., Froidevaux, L., Roth, C., Sofieva, V., Wang, R., Wild, J., Yu, P., Ziemke, J. R., and Rozanov, E. V.: Evidence for a continuous decline in lower stratospheric ozone offsetting ozone layer recovery, *Atmos. Chem. Phys.*, 18, 1379–1394, <https://doi.org/10.5194/acp-18-1379-2018>, 2018.
- Ball, W. T., Alsing, J., Staehelin, J., Davis, S. M., Froidevaux, L., and Peter, T.: Stratospheric ozone trends for 1985–2018: sensitivity to recent large variability, *Atmos. Chem. Phys.*, 19, 12731–12748, <https://doi.org/10.5194/acp-19-12731-2019>, 2019.
- Banerjee, A., Maycock, A. C., Archibald, A. T., Abraham, N. L., Telford, P., Braesicke, P., and Pyle, J. A.: Drivers of changes in stratospheric and tropospheric ozone between year 2000 and 2100, *Atmos. Chem. Phys.*, 16, 2727–2746, <https://doi.org/10.5194/acp-16-2727-2016>, 2016.
- Bhartia, P., McPeters, R., Mateer, C., Flynn, L., and Wellemeyer, C.: Algorithm for the estimation of vertical ozone profiles from the backscattered ultraviolet technique, *J. Geophys. Res.-Atmos.*, 101, 18793–18806, <https://doi.org/10.1029/96jd01165>, 1996.
- Bhartia, P. K., Herman, J., McPeters, R. D., and Torres, O.: Effect of Mount Pinatubo aerosols on total ozone measurements from backscatter ultraviolet (BUV) experiments, *J. Geophys. Res.-Atmos.*, 98, 18547–18554, <https://doi.org/10.1029/93jd01739>, 1993.
- Bhartia, P. K., McPeters, R. D., Flynn, L. E., Taylor, S., Kramarova, N. A., Frith, S., Fisher, B., and DeLand, M.: Solar Backscatter UV (SBUV) total ozone and profile algorithm, *Atmos. Meas. Tech.*, 6, 2533–2548, <https://doi.org/10.5194/amt-6-2533-2013>, 2013.
- Bhartia, P. K., McPeters, R. D., Flynn, L. E., Taylor, S., Kramarova, N. A., Frith, S., Fisher, B., and DeLand, M.: Solar Backscatter UV (SBUV) total ozone and profile algorithm, *Atmos. Meas. Tech.*, 6, 2533–2548, <https://doi.org/10.5194/amt-6-2533-2013>, 2013.
- Bourassa, A. E., Roth, C. Z., Zawada, D. J., Rieger, L. A., McLinden, C. A., and Degenstein, D. A.: Drift-corrected Odin-OSIRIS ozone product: algorithm and updated stratospheric ozone trends, *Atmos. Meas. Tech.*, 11, 489–498, <https://doi.org/10.5194/amt-11-489-2018>, 2018.
- Brönnimann, S.: Century-long column ozone records show that chemical and dynamical influences counteract each other, *Nat. Commun. Earth Environ.*, 3, 143, <https://doi.org/10.1038/s43247-022-00472-z>, 2022.
- Brunner, D., Staehelin, J., Maeder, J. A., Wohltmann, I., and Bodeker, G. E.: Variability and trends in total and vertically resolved stratospheric ozone based on the CATO ozone data set, *Atmos. Chem. Phys.*, 6, 4985–5008, <https://doi.org/10.5194/acp-6-4985-2006>, 2006.
- Burrows, J. P., Weber, M., Buchwitz, M., Rozanov, V., Ladstätter-Weißenmayer, A., Richter, A., DeBeek, R., Hoogen, R., Bramstedt, K., Eichmann, K.-U., Eisinger, M., and Perner, D.: The global ozone monitoring experiment (GOME): Mission concept and first scientific results, *J. Atmos. Sci.*, 56, 151–175, [https://doi.org/10.1175/1520-0469\(1999\)056<0151:TGOMEG>2.0.CO;2](https://doi.org/10.1175/1520-0469(1999)056<0151:TGOMEG>2.0.CO;2), 1999.
- Charlton-Perez, A. J., Hawkins, E., Eyring, V., Cionni, I., Bodeker, G. E., Kinnison, D. E., Akiyoshi, H., Frith, S. M., Garcia, R., Gettelman, A., Lamarque, J. F., Nakamura, T., Pawson, S., Yamashita, Y., Bekki, S., Braesicke, P., Chipperfield, M. P., Dhomse, S., Marchand, M., Mancini, E., Morgenstern, O., Pitari, G., Plummer, D., Pyle, J. A., Rozanov, E., Scinocca, J., Shibata, K., Shepherd, T. G., Tian, W., and Waugh, D. W.: The potential to narrow uncertainty in projections of stratospheric ozone over the 21st century, *Atmos. Chem. Phys.*, 10, 9473–9486, <https://doi.org/10.5194/acp-10-9473-2010>, 2010.
- Chiodo, G., Polvani, L. M., Marsh, D. R., Stenke, A., Ball, W., Rozanov, E., Muthers, S., and Tsigaridis, K.: The response of the ozone layer to quadrupled CO₂ concentrations, *J. Climate*, 31, 3893–3907, <https://doi.org/10.1175/JCLI-D-17-0492.1>, 2018.
- Chipperfield, M. P., Bekki, S., Dhomse, S., Harris, N. R., Hassler, B., Hossaini, R., Steinbrecht, W., Thiéblemont, R., and Weber, M.: Detecting recovery of the stratospheric ozone layer, *Nature*, 549, 211–218, <https://doi.org/10.1038/nature23681>, 2017.

- Chipperfield, M. P., Dhomse, S., Hossaini, R., Feng, W., Santee, M. L., Weber, M., Burrows, J. P., Wild, J. D., Loyola, D., and Coldewey-Egbers, M.: On the cause of recent variations in lower stratospheric ozone, *Geophys. Res. Lett.*, 45, 5718–5726, <https://doi.org/10.1029/2018gl078071>, 2018.
- Chipperfield, M. P., Hossaini, R., Montzka, S. A., Reimann, S., Sherry, D., and Tegtmeier, S.: Renewed and emerging concerns over the production and emission of ozone-depleting substances, *Nat. Rev. Earth Environ.*, 1, 251–263, <https://doi.org/10.1038/s43017-020-0048-8>, 2020.
- Cisewski, M., Zawodny, J., Gasbarre, J., Eckman, R., Topiwala, N., Rodriguez-Alvarez, O., Cheek, D., and Hall, S.: The stratospheric aerosol and gas experiment (SAGE III) on the International Space Station (ISS) Mission, in: *Sensors, Systems, and Next-Generation Satellites XVIII*, VI. 9241, 59–65, SPIE, <https://doi.org/10.1117/12.2073131>, 2014.
- Damadeo, R. P., Zawodny, J. M., Remsberg, E. E., and Walker, K. A.: The impact of nonuniform sampling on stratospheric ozone trends derived from occultation instruments, *Atmos. Chem. Phys.*, 18, 535–554, <https://doi.org/10.5194/acp-18-535-2018>, 2018.
- Davis, S. M., Rosenlof, K. H., Hassler, B., Hurst, D. F., Read, W. G., Vömel, H., Selkirk, H., Fujiwara, M., and Damadeo, R.: The Stratospheric Water and Ozone Satellite Homogenized (SWOOSH) database: a long-term database for climate studies, *Earth Syst. Sci. Data*, 8, 461–490, <https://doi.org/10.5194/essd-8-461-2016>, 2016.
- Davis, S. M., Hegglin, M. I., Fujiwara, M., Dragani, R., Harada, Y., Kobayashi, C., Long, C., Manney, G. L., Nash, E. R., Potter, G. L., Tegtmeier, S., Wang, T., Wargan, K., and Wright, J. S.: Assessment of upper tropospheric and stratospheric water vapor and ozone in reanalyses as part of S-RIP, *Atmos. Chem. Phys.*, 17, 12743–12778, <https://doi.org/10.5194/acp-17-12743-2017>, 2017.
- Deser, C., Phillips, A., Bourdette, V., and Teng, H.: Uncertainty in climate change projections: the role of internal variability, *Clim. Dynam.*, 38, 527–546, <https://doi.org/10.1007/s00382-010-0977-x>, 2012.
- Diallo, M., Riese, M., Birner, T., Konopka, P., Müller, R., Hegglin, M. I., Santee, M. L., Baldwin, M., Legras, B., and Ploeger, F.: Response of stratospheric water vapor and ozone to the unusual timing of El Niño and the QBO disruption in 2015–2016, *Atmos. Chem. Phys.*, 18, 13055–13073, <https://doi.org/10.5194/acp-18-13055-2018>, 2018.
- Diallo, M., Konopka, P., Santee, M. L., Müller, R., Tao, M., Walker, K. A., Legras, B., Riese, M., Ern, M., and Ploeger, F.: Structural changes in the shallow and transition branch of the Brewer–Dobson circulation induced by El Niño, *Atmos. Chem. Phys.*, 19, 425–446, <https://doi.org/10.5194/acp-19-425-2019>, 2019.
- Dietmüller, S., Ponater, M., and Sausen, R.: Interactive ozone induces a negative feedback in CO₂-driven climate change simulations, *J. Geophys. Res.-Atmos.*, 119, 1796–1805, <https://doi.org/10.1002/2013jd020575>, 2014.
- Dietmüller, S., Garny, H., Eichinger, R., and Ball, W. T.: Analysis of recent lower-stratospheric ozone trends in chemistry climate models, *Atmos. Chem. Phys.*, 21, 6811–6837, <https://doi.org/10.5194/acp-21-6811-2021>, 2021.
- Dunne, J. P., Horowitz, L. W., Adcroft, A. J., Ginoux, P., Held, I. M., John, J. G., Krasting, J. P., Malyshev, S., Naik, V., Paulot, F., Shevliakova, E., Stock, C. A., Zadeh, N., Balaji, V., Blanton, C., Dunne, K. A., Dupuis, C., Durachta, J., Dussin, R., Gauthier, P. P. G., Griffies, S. M., Guo, H., Hallberg, R. W., Harrison, M., He, J., Hurlin, W., McHugh, C., Menzel, R., Milly, P. C. D., Nikonov, S., Paynter, D. J., Ploshay, J., Radhakrishnan, A., Rand, K., Reichl, B. G., Robinson, T., Schwarzkopf, D. M., Sentman, L. T., Underwood, S., Vahlenkamp, H., Winton, M., Wittenberg, A. T., Wyman, B., Zeng, Y., and Zhao, M.: The GFDL Earth System Model version 4.1 (GFDL-ESM 4.1): Overall coupled model description and simulation characteristics, *J. Adv. Model Earth Sy.*, 12, e2019MS002015, <https://doi.org/10.1029/2019ms002015>, 2020.
- Evan, S., Brioude, J., Rosenlof, K. H., Gao, R.-S., Portmann, R. W., Zhu, Y., Volkamer, R., Lee, C. F., Metzger, J.-M., Lamy, K., Walter, P., Alvarez, S. L., Flynn, J. H., Asher, E., Todt, M., Davis, S. M., Thornberry, T., Vömel, H., Wienhold, F. G., Stauffer, R. M., Millán, L., Santee, M. L., Froidevaux, L., and Read, W. G.: Rapid ozone depletion after humidification of the stratosphere by the Hunga Tonga Eruption, *Science*, 382, eadg2551, <https://doi.org/10.1126/science.adg2551>, 2023.
- Eyring, V., Arblaster, J. M., Cionni, I., Sedláček, J., Perlwitz, J., Young, P. J., Bekki, S., Bergmann, D., Cameron-Smith, P., Collins, W. J., Faluvegi, G., Gottschaldt, K.-D., Horowitz, L. W., Kinnison, D. E., Lamarque, J.-F., Marsh, D. R., Saint-Martin, D., Shindell, D. T., Sudo, K., Szopa, S., and Watanabe, S.: Long-term ozone changes and associated climate impacts in CMIP5 simulations, *J. Geophys. Res.-Atmos.*, 118, 5029–5060, <https://doi.org/10.1002/jgrd.50316>, 2013a.
- Eyring, V., Lamarque, J.-F., Hess, P., Arfeuille, F., Bowman, K., Chipperfield, M. P., Duncan, B., Fiore, A., Gettelman, A., Giorgietta, M. A., Granier, C., Hegglin, M., Kinnison, D., Kunze, M., Langematz, U., Luo, B., Randall, M., Matthes, K., Newman, P. A., Peter, T., Robock, A., Ryerson, T., Saiz-Lopez, A., Salawitch, R., Schultz, M., Shepherd, T. G., Shindell, D., Staelin, J., Tegtmeier, S., Thomason, L., Tilmes, S., Vernier, J.-P., Waugh, D., and Young, P. Y.: Overview of IGAC/SPARC Chemistry-Climate Model Initiative (CCMI) community simulations in support of upcoming ozone and climate assessments, *SPARC newsletter*, 40, 48–66, <https://oceanrep.geomar.de/id/eprint/20227> (last access: 7 February 2025), 2013b.
- Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, *Geosci. Model Dev.*, 9, 1937–1958, <https://doi.org/10.5194/gmd-9-1937-2016>, 2016.
- Farman, J. C., Gardiner, B. G., and Shanklin, J. D.: Large losses of total ozone in Antarctica reveal seasonal ClO_x/NO_x interaction, *Nature*, 315, 207–210, <https://doi.org/10.1038/315207a0>, 1985.
- Fels, S. B., Mahlman, J. D., Schwarzkopf, M. D., and Sinclair, R. W.: Stratospheric sensitivity to perturbations in ozone and carbon dioxide- Radiative and dynamical response, *J. Atmos. Sci.*, 37, 2265–2297, [https://doi.org/10.1175/1520-0469\(1980\)037<2265:sstpio>2.0.co;2](https://doi.org/10.1175/1520-0469(1980)037<2265:sstpio>2.0.co;2), 1980.
- Frith, S. M., Kramarova, N. A., Stolarski, R. S., McPeters, R. D., Bhartia, P. K., and Labow, G. J.: Recent changes in total column ozone based on the SBUV Version 8.6 Merged Ozone Data Set, *J. Geophys. Res.-Atmos.*, 119, 9735–9751, <https://doi.org/10.1002/2014jd021889>, 2014.

- Frith, S. M., Stolarski, R. S., Kramarova, N. A., and McPeters, R. D.: Estimating uncertainties in the SBUV Version 8.6 merged profile ozone data set, *Atmos. Chem. Phys.*, 17, 14695–14707, <https://doi.org/10.5194/acp-17-14695-2017>, 2017.
- Froidevaux, L., Anderson, J., Wang, H.-J., Fuller, R. A., Schwartz, M. J., Santee, M. L., Livesey, N. J., Pumphrey, H. C., Bernath, P. F., Russell III, J. M., and McCormick, M. P.: Global OZone Chemistry And Related trace gas Data records for the Stratosphere (GOZCARDS): methodology and sample results with a focus on HCl, H₂O, and O₃, *Atmos. Chem. Phys.*, 15, 10471–10507, <https://doi.org/10.5194/acp-15-10471-2015>, 2015.
- Gillett, N. P., Akiyoshi, H., Bekki, S., Braesicke, P., Eyring, V., Garcia, R., Karpechko, A. Yu., McLinden, C. A., Morgenstern, O., Plummer, D. A., Pyle, J. A., Rozanov, E., Scinocca, J., and Shibata, K.: Attribution of observed changes in stratospheric ozone and temperature, *Atmos. Chem. Phys.*, 11, 599–609, <https://doi.org/10.5194/acp-11-599-2011>, 2011.
- Giorgi, F. and Bi, X.: Time of emergence (TOE) of GHG-forced precipitation change hot-spots, *Geophys. Res. Lett.*, 36, L06709, <https://doi.org/10.1029/2009gl0137593>, 2009.
- Godin-Beekmann, S., Azouz, N., Sofieva, V. F., Hubert, D., Petropavlovskikh, I., Effertz, P., Ancellet, G., Degenstein, D. A., 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., Van Malderen, R., Mahieu, E., Smale, D., and Sussmann, R.: Updated trends of the stratospheric ozone vertical distribution in the 60°S–60°N latitude range based on the LOTUS regression model, *Atmos. Chem. Phys.*, 22, 11657–11673, <https://doi.org/10.5194/acp-22-11657-2022>, 2022.
- Harris, N. R. P., Hassler, B., Tummon, F., Bodeker, G. E., Hubert, D., Petropavlovskikh, I., Steinbrecht, W., Anderson, J., Bhartia, P. K., Boone, C. D., Bourassa, A., Davis, S. M., Degenstein, D., Delcloo, A., Frith, S. M., Froidevaux, L., Godin-Beekmann, S., Jones, N., Kurylo, M. J., Kyrölä, E., Laine, M., Leblanc, S. T., Lambert, J.-C., Liley, B., Mahieu, E., Maycock, A., de Mazière, M., Parrish, A., Querel, R., Rosenlof, K. H., Roth, C., Sioris, C., Staehelin, J., Stolarski, R. S., Stübi, R., Tamminen, J., Vigouroux, C., Walker, K. A., Wang, H. J., Wild, J., and Zawodny, J. M.: Past changes in the vertical distribution of ozone – Part 3: Analysis and interpretation of trends, *Atmos. Chem. Phys.*, 15, 9965–9982, <https://doi.org/10.5194/acp-15-9965-2015>, 2015.
- Hawkins, E. and Sutton, R.: Time of emergence of climate signals, *Geophys. Res. Lett.*, 39, L01702, <https://doi.org/10.1029/2011gl0150087>, 2012.
- Hegerl, G. C., Zwiers, F. W., Braconnot, P., Gillett, N. P., Luo, Y., Marengo Orsini, J. A., Nicholls, N., Penner, J. E., and Stott, P. A.: Understanding and attributing climate change, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC), edited by: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M., and Miller, H. L., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 996 pp., <https://www.ipcc.ch/site/assets/uploads/2018/02/ar4-wg1-chapter9-1.pdf> (last access: 7 February 2025), 2007.
- Hegglin, M., Lamarque, J., Duncan, B., Eyring, V., Gettelman, A., Hess, P., Myhre, G., Nagashima, T., Plummer, D., Ryerson, T., Shepherd, T., and Waugh, D.: Report on the IGAC/S-PARC Chemistry-Climate Model Initiative (CCMI) 2015 Science Workshop, SPARC Newsletter, 46, 37–42, 2016.
- Hegglin, M. I. and Shepherd, T. G.: Large climate-induced changes in ultraviolet index and stratosphere-to-troposphere ozone flux, *Nat. Geosci.*, 2, 687–691, <https://doi.org/10.1038/ngeo604>, 2009.
- Hegglin, M. I., Lamarque, J. F., and Eyring, V.: The IGAC/S-PARC Chemistry-Climate Model Initiative Phase-1 (CCMI-1) model data output, NCAS British Atmospheric Data Centre [data set], date of citation, <http://catalogue.ceda.ac.uk/uuid/9cc6b94df0f4469d8066d69b5df879d5/> (last access: 7 February 2025), 2015.
- Horowitz, L. W., Naik, V., Paulot, F., Ginoux, P. A., Dunne, J. P., Mao, J., Schnell, J., Chen, X., He, J., John, J. G., Lin, M., Lin, P., Malyshev, S., Paynter, D., Shevliakova, E., and Zhao, M.: The GFDL global atmospheric chemistry-climate model AM4. 1: Model description and simulation characteristics, *J. Adv. Model Earth Sy.*, 12, e2019MS002032, <https://doi.org/10.1029/2019ms002032>, 2020.
- Hubert, D., Lambert, J.-C., Verhoelst, T., Granville, J., Keppens, A., Baray, J.-L., Bourassa, A. E., Cortesi, U., Degenstein, D. A., Froidevaux, L., Godin-Beekmann, S., Hoppel, K. W., Johnson, B. J., Kyrölä, E., Leblanc, T., Lichtenberg, G., Marchand, M., McElroy, C. T., Murtagh, D., Nakane, H., Portafaix, T., Querel, R., Russell III, J. M., Salvador, J., Smit, H. G. J., Stebel, K., Steinbrecht, W., Strawbridge, K. B., Stübi, R., Swart, D. P. J., Taha, G., Tarasick, D. W., Thompson, A. M., Urban, J., van Gijssel, J. A. E., Van Malderen, R., von der Gathen, P., Walker, K. A., Wolfram, E., and Zawodny, J. M.: Ground-based assessment of the bias and long-term stability of 14 limb and occultation ozone profile data records, *Atmos. Meas. Tech.*, 9, 2497–2534, <https://doi.org/10.5194/amt-9-2497-2016>, 2016.
- IPCC: Climate change 2001: the scientific basis, edited by: Houghton, J. T., Ding, Y., Griggs, D. J., Noguer, M., van der Linden, P. J., Dai, X., Maskell, K., and Johnson, C. A., Vol. 881, Cambridge University Press, <https://www.ipcc.ch/report/ar3/wg1/> (last access: 7 February 2025), 2001.
- Kay, J. E., Deser, C., Phillips, A., Mai, A., Hannay, C., Strand, G., Arblaster, J. M., Bates, S., Danabasoglu, G., Edwards, J., Holland, M., Kushner, P., Lamarque, J.-F., Lawrence, D., Lindsay, K., Middleton, A., Munoz, E., Neale, R., Oleson, K., Polvani, L., and Vertenstein, M.: The Community Earth System Model (CESM) large ensemble project: A community resource for studying climate change in the presence of internal climate variability, *B. Am. Meteorol. Soc.*, 96, 1333–1349, <https://doi.org/10.1175/BAMS-D-13-00255.1>, 2015.
- Kramarova, N. A., Bhartia, P. K., Frith, S. M., McPeters, R. D., and Stolarski, R. S.: Interpreting SBUV smoothing errors: an example using the quasi-biennial oscillation, *Atmos. Meas. Tech.*, 6, 2089–2099, <https://doi.org/10.5194/amt-6-2089-2013>, 2013a.
- Kramarova, N. A., Frith, S. M., Bhartia, P. K., McPeters, R. D., Taylor, S. L., Fisher, B. L., Labow, G. J., and DeLand, M. T.: Validation of ozone monthly zonal mean profiles obtained from the version 8.6 Solar Backscatter Ultraviolet algorithm, *Atmos. Chem. Phys.*, 13, 6887–6905, <https://doi.org/10.5194/acp-13-6887-2013>, 2013b.
- Krasting, J. P., John, J. G., Blanton, C., McHugh, C., Nikonov, S., Radhakrishnan, A., Rand, K., Zadeh, N. T., Balaji, V., Durachta, J., Dupuis, C., Menzel, R., Robinson, T., Underwood,

- S., Vahlenkamp, H., Dunne, K. A., Gauthier, P. P. G., Ginoux, P., Griffies, S. M., Hallberg, R., Harrison, M., Hurlin, W., Malyshev, S., Naik, V., Paulot, F., Paynter, D. J., Ploshay, J., Reichl, B. G., Schwarzkopf, D. M., Seman, C. J., Silvers, L., Wyman, B., Zeng, Y., Adcroft, A., Dunne, J. P., Dussin, R., Guo, H., He, J., Held, I. M., Horowitz, L. W., Lin, P., Milly, P. C. D., Shevliakova, E., Stock, C., Winton, M., Wittenberg, A. T., Xie, Y., and Zhao, M.: NOAA-GFDL GFDL-ESM4 model output prepared for CMIP6 CMIP piControl, Version 20191127, Earth System Grid Federation, WCRP [data set], <https://doi.org/10.22033/ESGF/CMIP6.8669>, 2018a.
- Krasting, J. P., John, J. G., Blanton, C., McHugh, C., Nikonov, S., Radhakrishnan, A., Rand, K., Zadeh, N. T., Balaji, V., Durachta, J., Dupuis, C., Menzel, R., Robinson, T., Underwood, S., Vahlenkamp, H., Dunne, K. A., Gauthier, P. P. G., Ginoux, P., Griffies, S. M., Hallberg, R., Harrison, M., Hurlin, W., Malyshev, S., Naik, V., Paulot, F., Paynter, D. J., Ploshay, J., Reichl, B. G., Schwarzkopf, D. M., Seman, C. J., Silvers, L., Wyman, B., Zeng, Y., Adcroft, A., Dunne, J. P., Dussin, R., Guo, H., He, J., Held, I. M., Horowitz, L. W., Lin, P., Milly, P. C. D., Shevliakova, E., Stock, C., Winton, M., Wittenberg, A. T., Xie, Y., and Zhao, M.: NOAA-GFDL GFDL-ESM4 model output prepared for CMIP6 CMIP historical, Version 20210526, Earth System Grid Federation, WCRP [data set], <https://doi.org/10.22033/ESGF/CMIP6.8597>, 2018b.
- Levelt, P. F., Hilsenrath, E., Leppelmeier, G. W., van den Oord, G. H., Bhartia, P. K., Tamminen, J., de Haan, J. F., and Veefkind, J. P.: Science objectives of the ozone monitoring instrument, *IEEE T. Geosci. Remote*, 44, 1199–1208, <https://doi.org/10.1109/tgrs.2006.872336>, 2006.
- Li, F., Stolarski, R. S., and Newman, P. A.: Stratospheric ozone in the post-CFC era, *Atmos. Chem. Phys.*, 9, 2207–2213, <https://doi.org/10.5194/acp-9-2207-2009>, 2009.
- Li, J., Thompson, D. W., Barnes, E. A., and Solomon, S.: Quantifying the lead time required for a linear trend to emerge from natural climate variability, *J. Climate*, 30, 10179–10191, <https://doi.org/10.1175/jcli-d-16-0280.1>, 2017.
- Lickley, M., Fletcher, S., Rigby, M., and Solomon, S.: Joint inference of CFC lifetimes and banks suggests previously unidentified emissions, *Nat. Commun.*, 12, 2920, <https://doi.org/10.1038/s41467-021-23229-2>, 2021.
- Livesey, N. J., Van Snyder, W., Read, W. G., and Wagner, P. A.: Retrieval algorithms for the EOS Microwave limb sounder (MLS), *IEEE T. Geosci. Remote*, 44, 1144–1155, <https://doi.org/10.1109/tgrs.2006.872327>, 2006.
- Madden, R. A. and Ramanathan, V.: Detecting climate change due to increasing carbon dioxide, *Science*, 209, 763–768, <https://doi.org/10.1126/science.209.4458.763>, 1980.
- Maher, N., Power, S. B., and Marotzke, J.: More accurate quantification of model-to-model agreement in externally forced climatic responses over the coming century, *Nat. Commun.*, 12, 788, <https://doi.org/10.1038/s41467-020-20635-w>, 2021.
- Mahfouf, J., Cariolle, D., Royer, J., Geleyn, J., and Timbal, B.: Response of the Meteo-France climate model to changes in CO₂ and sea surface temperature, *Clim. Dynam.*, 9, 345–362, <https://doi.org/10.1007/bf00223447>, 1994.
- Mahlstein, I., Knutti, R., Solomon, S., and Portmann, R. W.: Early onset of significant local warming in low latitude countries, *Environ. Res. Lett.*, 6, 034009, <https://doi.org/10.1088/1748-9326/6/3/034009>, 2011.
- Manney, G. L., Santee, M. L., Lambert, A., Millán, L. F., Minschwaner, K., Werner, F., Lawrence, Z. D., Read, W. G., Livesey, N. J., and Wang, T.: Siege in the Southern Stratosphere: Hunga Tonga-Hunga Ha’apai Water Vapor Excluded From the 2022 Antarctic Polar Vortex, *Geophys. Res. Lett.*, 50, e2023GL103855, <https://doi.org/10.1029/2023gl103855>, 2023.
- Match, A. and Gerber, E. P.: Tropospheric expansion under global warming reduces tropical lower stratospheric ozone, *Geophys. Res. Lett.*, 49, e2022GL099463, <https://doi.org/10.1029/2022gl099463>, 2022.
- McPeters, R. D., Bhartia, P. K., Haffner, D., Labow, G. J., and Flynn, L.: The version 8.6 SBUV ozone data record: An overview, *J. Geophys. Res.-Atmos.*, 118, 8032–8039, <https://doi.org/10.1002/jgrd.50597>, 2013.
- Meul, S., Langematz, U., Oberländer, S., Garny, H., and Jöckel, P.: Chemical contribution to future tropical ozone change in the lower stratosphere, *Atmos. Chem. Phys.*, 14, 2959–2971, <https://doi.org/10.5194/acp-14-2959-2014>, 2014.
- Montzka, S. A., Dutton, G. S., Yu, P., Ray, E., Portmann, R. W., Daniel, J. S., Kuijpers, L., Hall, B. D., Mondeel, D., Siso, C., Nance, J. D., Rigby, M., Manning, A. J., Hu, L., Moore, F., Miller, B. R., and Elkins, J. W.: An unexpected and persistent increase in global emissions of ozone-depleting CFC-11, *Nature*, 557, 413–417, <https://doi.org/10.1038/s41586-018-0106-2>, 2018.
- Morgenstern, O., Hegglin, M. I., Rozanov, E., O’Connor, F. M., Abraham, N. L., Akiyoshi, H., Archibald, A. T., Bekki, S., Butchart, N., Chipperfield, M. P., Deushi, M., Dhomse, S. S., Garcia, R. R., Hardiman, S. C., Horowitz, L. W., Jöckel, P., Josse, B., Kinnison, D., Lin, M., Mancini, E., Manyin, M. E., Marchand, M., Maréchal, V., Michou, M., Oman, L. D., Pitari, G., Plummer, D. A., Revell, L. E., Saint-Martin, D., Schofield, R., Stenke, A., Stone, K., Sudo, K., Tanaka, T. Y., Tilmes, S., Yamashita, Y., Yoshida, K., and Zeng, G.: Review of the global models used within phase 1 of the Chemistry–Climate Model Initiative (CCMI), *Geosci. Model Dev.*, 10, 639–671, <https://doi.org/10.5194/gmd-10-639-2017>, 2017.
- NASA Goddard Space Flight Center: SBUV Merged Ozone Data Set (MOD), NASA [data set], https://acd-ext.gsfc.nasa.gov/Data_services/merged/ (last access: 7 February 2025).
- Newchurch, M. J., Yang, E.-S., Cunnold, D. M., Reinsel, G. C., Zawodny, J., and Russell III, J. M.: Evidence for slowdown in stratospheric ozone loss: First stage of ozone recovery, *J. Geophys. Res.-Atmos.*, 108, 4507, <https://doi.org/10.1029/2003jd003471>, 2003.
- Nowack, P. J., Luke Abraham, N., Maycock, A. C., Braesicke, P., Gregory, J. M., Joshi, M. M., Osprey, A., and Pyle, J. A.: A large ozone-circulation feedback and its implications for global warming assessments, *Nat. Clim. Change*, 5, 41–45, <https://doi.org/10.1038/nclimate2451>, 2015.
- Oman, L., Plummer, D., Waugh, D., Austin, J., Scinocca, J., Douglass, A., Salawitch, R., Canty, T., Akiyoshi, H., Bekki, S., Braesicke, P., Butchart, N., Chipperfield, M. P., Cugnet, D., Dhomse, S., Eyring, V., Frith, S., Hardiman, S. C., Kinnison, D. E., Lamarque, J.-F., Mancini, E., Marchand, M., Michou, M., Morgenstern, O., Nakamura, T., Nielsen, J. E., Olivé, D., Pitari, G., Pyle, J., Rozanov, E., Shepherd, T. G., Shibata, K., Stolarski,

- R. S., Teyssède, H., Tian, W., Yamashita, Y., and Ziemke, J. R.: Multimodel assessment of the factors driving stratospheric ozone evolution over the 21st century, *J. Geophys. Res.-Atmos.*, 115, D24306, <https://doi.org/10.1029/2010jd014362>, 2010.
- Oman, L. D., Douglass, A. R., Ziemke, J. R., Rodriguez, J. M., Waugh, D. W., and Nielsen, J. E.: The ozone response to ENSO in Aura satellite measurements and a chemistry-climate simulation, *J. Geophys. Res.-Atmos.*, 118, 965–976, <https://doi.org/10.1029/2012jd018546>, 2013.
- Petropavlovskikh, I., Godin-Beekmann, S., Hubert, D., Damadeo, R., Hassler, B., and Sofieva, V.: SPARC/IO3C/GAW report on Long-term Ozone Trends and Uncertainties in the Stratosphere, SPARC Report No. 9, GAW Report No. 241, WCRP-17/2018, <https://doi.org/10.17874/f899e57a20b>, 2019.
- Plummer, D. A., Scinocca, J. F., Shepherd, T. G., Reader, M. C., and Jonsson, A. I.: Quantifying the contributions to stratospheric ozone changes from ozone depleting substances and greenhouse gases, *Atmos. Chem. Phys.*, 10, 8803–8820, <https://doi.org/10.5194/acp-10-8803-2010>, 2010.
- Randel, W. J., Garcia, R. R., Calvo, N., and Marsh, D.: ENSO influence on zonal mean temperature and ozone in the tropical lower stratosphere, *Geophys. Res. Lett.*, 36, L15822, <https://doi.org/10.1029/2009gl039343>, 2009.
- Revell, L. E., Bodeker, G. E., Huck, P. E., Williamson, B. E., and Rozanov, E.: The sensitivity of stratospheric ozone changes through the 21st century to N₂O and CH₄, *Atmos. Chem. Phys.*, 12, 11309–11317, <https://doi.org/10.5194/acp-12-11309-2012>, 2012.
- Ridolfi, M., Carli, B., Carlotti, M., von Clarmann, T., Dinelli, B. M., Dudhia, A., Flaud, J.-M., Höpfner, M., Morris, P. E., Raspollini, P., Stiller, G., and Wells, R. J.: Optimized forward model and retrieval scheme for MIPAS near-real-time data processing, *Appl. Opt.*, 39, 1323–1340, <https://doi.org/10.1117/12.317770>, 2000.
- Rigby, M., Park, S., Saito, T., Western, L. M., Redington, A. L., Fang, X., Henne, S., Manning, A. J., Prinn, R. G., Dutton, G. S., Fraser, P. J., Ganesan, A. L., Hall, B. D., Harth, C. M., Kim, J., Kim, K.-R., Krummel, P. B., Lee, T., Li, S., Liang, Q., Lunt, M. F., Montzka, S. A., Mühle, J., O'Doherty, S., Park, M.-K., Reimann, S., Salameh, P. K., Simmonds, P., Tunnicliffe, R. L., Weiss, R. F., Yokouchi, Y., and Young, D.: Increase in CFC-11 emissions from eastern China based on atmospheric observations, *Nature*, 569, 546–550, <https://doi.org/10.1038/s41586-019-1193-4>, 2019.
- Rind, D., Suozzo, R., Balachandran, N., and Prather, M.: Climate change and the middle atmosphere. Part I: The doubled CO₂ climate, *J. Atmos. Sci.*, 47, 475–494, [https://doi.org/10.1175/1520-0469\(1990\)047<0475:ccatma>2.0.co;2](https://doi.org/10.1175/1520-0469(1990)047<0475:ccatma>2.0.co;2), 1990.
- Rivoire, L., Linz, M., and Li, J.: Observational limitations to the emergence of climate signals, *Geophys. Res. Lett.*, 51, e2024GL109638, <https://doi.org/10.1029/2024gl109638>, 2024.
- Rodgers, C. D.: Characterization and error analysis of profiles retrieved from remote sounding measurements, *J. Geophys. Res.-Atmos.*, 95, 5587–5595, <https://doi.org/10.1029/jd095id05p05587>, 1990.
- Rodgers, C. D.: Inverse methods for atmospheric sounding: theory and practice, Vol. 2, World scientific, <https://doi.org/10.1142/3171>, 2000.
- Rodgers, C. D. and Connor, B. J.: Intercomparison of remote sounding instruments, *J. Geophys. Res.-Atmos.*, 108, <https://doi.org/10.1029/2002jd002299>, 2003.
- Rodgers, K. B., Lee, S.-S., Rosenbloom, N., Timmermann, A., Danabasoglu, G., Deser, C., Edwards, J., Kim, J.-E., Simpson, I. R., Stein, K., Stuecker, M. F., Yamaguchi, R., Bódai, T., Chung, E.-S., Huang, L., Kim, W. M., Lamarque, J.-F., Lombardozzi, D. L., Wieder, W. R., and Yeager, S. G.: Ubiquity of human-induced changes in climate variability, *Earth Syst. Dynam.*, 12, 1393–1411, <https://doi.org/10.5194/esd-12-1393-2021>, 2021.
- Santee, M. L., Lambert, A., Manney, G. L., Livesey, N. J., Froidevaux, L., Neu, J. L., Schwartz, M., Millán, L., Werner, F., Read, W. G., Park, M., Fuller, R. A., and Ward, B. M.: Prolonged and pervasive perturbations in the composition of the Southern Hemisphere midlatitude lower stratosphere from the Australian New Year's fires, *Geophys. Res. Lett.*, 49, e2021GL096270, <https://doi.org/10.1029/2021gl096270>, 2022.
- Santee, M. L., Lambert, A., Froidevaux, L., Manney, G. L., Schwartz, M. J., Millán, L., Livesey, N. J., Read, W. G., Werner, F., and Fuller, R. A.: Strong Evidence of Heterogeneous Processing on Stratospheric Sulfate Aerosol in the Extrapolated Southern Hemisphere Following the 2022 Hunga Tonga-Hunga Ha'apai Eruption, *J. Geophys. Res.-Atmos.*, 128, e2023JD039169, <https://doi.org/10.1029/2023jd039169>, 2023.
- Schär, C., Vidale, P. L., Lüthi, D., Frei, C., Häberli, C., Liniger, M. A., and Appenzeller, C.: The role of increasing temperature variability in European summer heatwaves, *Nature*, 427, 332–336, <https://doi.org/10.1038/nature02300>, 2004.
- Scherrer, S. C., Appenzeller, C., Liniger, M. A., and Schär, C.: European temperature distribution changes in observations and climate change scenarios, *Geophys. Res. Lett.*, 32, L19705, <https://doi.org/10.1029/2005gl024108>, 2005.
- Shepherd, T. G.: Dynamics, stratospheric ozone, and climate change, *Atmos.-Ocean*, 46, 117–138, <https://doi.org/10.3137/ao.460106>, 2008.
- Sofieva, V. F., Kyrölä, E., Laine, M., Tamminen, J., Degenstein, D., Bourassa, A., Roth, C., Zawada, D., Weber, M., Rozanov, A., Raupach, N., Stiller, G., Laeng, A., von Clarmann, T., Walker, K. A., Sheese, P., Hubert, D., van Roozendaal, M., Zehner, C., Damadeo, R., Zawodny, J., Kramarova, N., and Bhartia, P. K.: Merged SAGE II, Ozone_cci and OMPS ozone profile dataset and evaluation of ozone trends in the stratosphere, *Atmos. Chem. Phys.*, 17, 12533–12552, <https://doi.org/10.5194/acp-17-12533-2017>, 2017.
- Solomon, S., Ivy, D. J., Kinnison, D., Mills, M. J., Neely III, R. R., and Schmidt, A.: Emergence of healing in the Antarctic ozone layer, *Science*, 353, 269–274, <https://doi.org/10.1126/science.aae0061>, 2016.
- Solomon, S., Stone, K., Yu, P., Murphy, D. M., Kinnison, D., Ravishankara, A. R., and Wang, P.: Chlorine activation and enhanced ozone depletion induced by wildfire aerosol, *Nature*, 615, 259–264, <https://doi.org/10.1038/s41586-022-05683-0>, 2023.
- SPARC: SPARC CCMVal Report on the Evaluation of Chemistry-Climate Models, edited by: Eyring, V., Shepherd, T., and Waugh, D., SPARC Report No. 5, WCRP-30/2010, WMO/TD – No. 40, <http://www.sparc-climate.org/publications/sparc-reports/> (last access: 7 February 2025), 2010.
- Steinbrecht, W., Froidevaux, L., Fuller, R., Wang, R., Anderson, J., Roth, C., Bourassa, A., Degenstein, D., Damadeo, R., Zawodny,

- J., Frith, S., McPeters, R., Bhartia, P., Wild, J., Long, C., Davis, S., Rosenlof, K., Sofieva, V., Walker, K., Rahepo, N., Rozanov, A., Weber, M., Laeng, A., von Clarmann, T., Stiller, G., Kramarova, N., Godin-Beekmann, S., Leblanc, T., Querel, R., Swart, D., Boyd, I., Hocke, K., Kämpfer, N., Maillard Barras, E., Moreira, L., Nedoluha, G., Vigouroux, C., Blumenstock, T., Schneider, M., García, O., Jones, N., Mahieu, E., Smale, D., Kotkamp, M., Robinson, J., Petropavlovskikh, I., Harris, N., Hassler, B., Hubert, D., and Tummon, F.: An update on ozone profile trends for the period 2000 to 2016, *Atmos. Chem. Phys.*, 17, 10675–10690, <https://doi.org/10.5194/acp-17-10675-2017>, 2017.
- Stolarski, R. S. and Frith, S. M.: Search for evidence of trend slowdown in the long-term TOMS/SBUV total ozone data record: the importance of instrument drift uncertainty, *Atmos. Chem. Phys.*, 6, 4057–4065, <https://doi.org/10.5194/acp-6-4057-2006>, 2006.
- Stolarski, R. S., Bloomfield, P., McPeters, R. D., and Herman, J. R.: Total ozone trends deduced from Nimbus 7 TOMS data, *Geophys. Res. Lett.*, 18, 1015–1018, <https://doi.org/10.1029/91gl01302>, 1991.
- Stone, K. A., Solomon, S., and Kinnison, D. E.: On the identification of ozone recovery, *Geophys. Res. Lett.*, 45, 5158–5165, <https://doi.org/10.1029/2018gl077955>, 2018.
- Thompson, D. W., Barnes, E. A., Deser, C., Foust, W. E., and Phillips, A. S.: Quantifying the role of internal climate variability in future climate trends, *J. Climate*, 28, 6443–6456, <https://doi.org/10.1175/jcli-d-14-00830.1>, 2015.
- Thomson, D. J.: Spectrum estimation and harmonic analysis, *P. IEEE*, 70, 1055–1096, <https://doi.org/10.1109/proc.1982.12433>, 1982.
- Tiao, G. C., Reinsel, G. C., Xu, D., Pedrick, J. H., Zhu, X., Miller, A. J., DeLuisi, J. J., Mateer, C. L., and Wuebbles, D. J.: Effects of autocorrelation and temporal sampling schemes on estimates of trend and spatial correlation, *J. Geophys. Res.-Atmos.*, 95, 20507–20517, <https://doi.org/10.1029/jd095id12p20507>, 1990.
- Tully, M. B., Krummel, P. B., and Klekociuk, A. R.: Trends in Antarctic ozone hole metrics 2001–17, *J. Southern Hemisphere Earth Syst. Sci.*, 69, 52–56, <https://doi.org/10.1071/es19020>, 2020.
- Tummon, F., Hassler, B., Harris, N. R. P., Staehelin, J., Steinbrecht, W., Anderson, J., Bodeker, G. E., Bourassa, A., Davis, S. M., Degenstein, D., Frith, S. M., Froidevaux, L., Kyrölä, E., Laine, M., Long, C., Penckwitt, A. A., Sioris, C. E., Rosenlof, K. H., Roth, C., Wang, H.-J., and Wild, J.: Intercomparison of vertically resolved merged satellite ozone data sets: interannual variability and long-term trends, *Atmos. Chem. Phys.*, 15, 3021–3043, <https://doi.org/10.5194/acp-15-3021-2015>, 2015.
- Tung, K. and Yang, H.: Global QBO in circulation and ozone. Part I: Reexamination of observational evidence, *J. Atmos. Sci.*, 51, 2699–2707, [https://doi.org/10.1175/1520-0469\(1994\)051<2699:gqicao>2.0.co;2](https://doi.org/10.1175/1520-0469(1994)051<2699:gqicao>2.0.co;2), 1994.
- von Savigny, C., Haley, C. S., Sioris, C. E., McDade, I. C., Llewellyn, E. J., Degenstein, D., Evans, W. F. J., Gattinger, R. L., Griffioen, E., Kyrölä, E., Lloyd, N. D., McConnell, J. C., McLinden, C. A., Mégie, G., Murtagh, D. P., Solheim, B., and Strong, K.: Stratospheric ozone profiles retrieved from limb scattered sunlight radiance spectra measured by the OSIRIS instrument on the Odin satellite, *Geophys. Res. Lett.*, 30, 1755, <https://doi.org/10.1029/2002gl016401>, 2003.
- Walker, K. A., Randall, C. E., Trepte, C. R., Boone, C. D., and Bernath, P. F.: Initial validation comparisons for the Atmospheric Chemistry Experiment (ACE-FTS), *Geophys. Res. Lett.*, 32, L16S04, <https://doi.org/10.1029/2005gl022388>, 2005.
- Wallace, J. M., Panetta, R. L., and Estberg, J.: Representation of the equatorial stratospheric quasi-biennial oscillation in EOF phase space, *J. Atmos. Sci.*, 50, 1751–1762, [https://doi.org/10.1175/1520-0469\(1993\)050<1751:rotesq>2.0.co;2](https://doi.org/10.1175/1520-0469(1993)050<1751:rotesq>2.0.co;2), 1993.
- Wang, X., Randel, W., Zhu, Y., Tilmes, S., Starr, J., Yu, W., Garcia, R., Toon, O. B., Park, M., Kinnison, D., Zhang, J., Bourassa, A., Rieger, L., Warnock, T., and Li, J.: Stratospheric Climate Anomalies and Ozone Loss Caused by the Hunga Tonga-Hunga Ha’apai Volcanic Eruption, *J. Geophys. Res.-Atmos.*, 128, e2023JD039480, <https://doi.org/10.1029/2023jd039480>, 2023.
- Waugh, D., Oman, L., Kawa, S., Stolarski, R., Pawson, S., Douglass, A., Newman, P., and Nielsen, J.: Impacts of climate change on stratospheric ozone recovery, *Geophys. Res. Lett.*, 36, L03805, <https://doi.org/10.1029/2008gl036223>, 2009.
- Weatherhead, E. C., Reinsel, G. C., Tiao, G. C., Meng, X.-L., Choi, D., Cheang, W.-K., Keller, T., DeLuisi, J., Wuebbles, D. J., Kerr, J. B., Miller, A. J., Oltmans, S. J., and Frederick, J. E.: Factors affecting the detection of trends: Statistical considerations and applications to environmental data, *J. Geophys. Res.-Atmos.*, 103, 17149–17161, <https://doi.org/10.1029/98JD00995>, 1998.
- 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, *Atmos. Chem. Phys.*, 18, 2097–2117, <https://doi.org/10.5194/acp-18-2097-2018>, 2018.
- Wigley, T. and Jones, P.: Detecting CO₂-induced climatic change, *Nature*, 292, 205–208, <https://doi.org/10.1038/292205a0>, 1981.
- Wilmouth, D. M., Østerstrøm, F. F., Smith, J. B., Anderson, J. G., and Salawitch, R. J.: Impact of the Hunga Tonga volcanic eruption on stratospheric composition, *P. Natl. Acad. Sci. USA*, 120, e2301994120, <https://doi.org/10.1073/pnas.2301994120>, 2023.
- WMO: Scientific assessment of ozone depletion: 2006, World Meteorological Organisation, Global Ozone Research and Monitoring Project, 50, 572, <https://ozone.unep.org/sites/default/files/2023-04/SAP-report-2006.pdf> (last access: 7 February 2025), 2007.
- WMO: Scientific Assessment of Ozone Depletion, Global Ozone Research and Monitoring Project – Report No. 52, Geneva, Switzerland, 516 pp., <https://ozone.unep.org/sites/default/files/2019-05/00-SAP-2010-Assement-report.pdf> (last access: 7 February 2025), 2011.
- WMO: Scientific Assessment of Ozone Depletion, Global Ozone Research and Monitoring Project – Report No. 58, Geneva, Switzerland, 588 pp., <https://ozone.unep.org/sites/default/files/2019-05/SAP-2018-Assessment-report.pdf> (last access: 7 February 2025), 2018.
- WMO: Scientific Assessment of Ozone Depletion, Global Ozone Research and Monitoring Project – GAW Report No. 278, p. 56 pp., <https://ozone.unep.org/sites/default/files/2023-02/Scientific-Assessment-of-Ozone-Depletion-2022.pdf> (last access: 7 February 2025), 2022.
- Worden, J., Kulawik, S. S., Shephard, M. W., Clough, S. A., Worden, H., Bowman, K., and Goldman, A.: Predicted errors of

- tropospheric emission spectrometer nadir retrievals from spectral window selection, *J. Geophys. Res.-Atmos.*, 109, D09308, <https://doi.org/10.1029/2004jd004522>, 2004.
- Yang, E.-S., Cunnold, D. M., Newchurch, M. J., Salawitch, R. J., McCormick, M. P., Russell III, J. M., Zawodny, J. M., and Oltmans, S. J.: First stage of Antarctic ozone recovery, *J. Geophys. Res.-Atmos.*, 113, D20308, <https://doi.org/10.1029/2007jd009675>, 2008.
- Ziemke, J. R., Labow, G. J., Kramarova, N. A., McPeters, R. D., Bhartia, P. K., Oman, L. D., Frith, S. M., and Haffner, D. P.: A global ozone profile climatology for satellite retrieval algorithms based on Aura MLS measurements and the MERRA-2 GMI simulation, *Atmos. Meas. Tech.*, 14, 6407–6418, <https://doi.org/10.5194/amt-14-6407-2021>, 2021.