ML-TOMCAT: Machine-Learning-Based Satellite-Corrected Global Stratospheric Ozone Profile Dataset from a Chemical Transport Model

Sandip S. Dhomse[1,2], Carlo Arosio[3], Wuhu Feng[1,4], Alexei Rozanov[3], Mark Weber[3], and Martyn P. Chipperfield[1,2]

[1]School of Earth and Environment, University of Leeds, Leeds, UK
[2]National Centre for Earth Observation, University of Leeds, Leeds, UK
[3]Institute for Environmental Physics, University of Bremen, Bremen, Germany
[4]National Centre for Atmospheric Science, University of Leeds, Leeds, UK

Correspondence: Sandip S. Dhomse (s.s.dhomse@leeds.ac.uk)

Abstract.

High quality stratospheric ozone profile datasets are a key requirement for accurate quantification and attribution of long-term ozone changes. Satellite instruments obtain stratospheric ozone profile measurements over typical mission durations of 5-15 years. Various methodologies have then been applied to merge and homogenise the different satellite data in order to

5   create longer term observation-based ozone profile datasets with minimal data gaps. However, individual satellite instruments use different measurement methods, sampling patterns and retrieval algorithms which complicate the merging of these different datasets. In contrast, atmospheric chemical models can produce chemically consistent long-term ozone simulations based on specified changes in external forcings, but they are subject to the deficiencies associated with incomplete understanding of complex atmospheric processes and uncertain photochemical parameters.

10   Here, we use chemically self-consistent output from the TOMCAT 3-D chemical transport model (CTM) and a Random-Forest (RF) ensemble learning method to create a merged 42-year (1979-2020) stratospheric ozone profile dataset (ML-TOMCAT V1.0). The underlying CTM simulation was forced by meteorological reanalyses, specified trends in long-lived source gas, solar flux and aerosol variations. The RF is trained using the Stratospheric Water and OzOne Satellite Homogenized (SWOOSH) dataset over the time periods of the Microwave Limb Sounder (MLS) from the Upper Atmosphere Research

15   Satellite (UARS) (1991-1998) and Aura (2005-2016) missions. We find that ML-TOMCAT shows excellent agreement with available independent satellite-based datasets which use pressure as the vertical coordinate (e.g. GOZCARDS, SWOOSH for non-MLS periods) but weaker agreement with the datasets which are height-based (e.g. SAGE–CCI–OMPS, SCIAMACHY-OMPS). We find that at almost all stratospheric levels ML-TOMCAT ozone concentrations are well within uncertainties in the observational datasets. The ML-TOMCAT dataset is thus ideally suited for the evaluation of model ozone profiles from the

20   tropopause to 0.1 hPa. ML-TOMCAT data is freely available via https://zenodo.org/record/4997959#.YNzleUlKiUk (Dhomse et al., 2021).

1 Introduction

With the successful implementation of the Montreal Protocol, various observations confirm reductions in the concentrations of ozone depleting substances (ODSs) in the atmosphere (WMO, 2014, 2018). Satellite data records also confirm a peak in
25 upper stratospheric HCl around 1997, followed by a steady decline (Anderson et al., 2000; Froidevaux et al., 2006a; Hossaini et al., 2019). Hence, attention has turned towards the detection and attribution of ozone recovery (e.g. Dhomse et al., 2006; Solomon et al., 2016; Chipperfield et al., 2017; Steinbrecht et al., 2017; Dhomse et al., 2018; Szeląg et al., 2020). However, accurate quantification of ozone changes is challenging because of the quality of long-term ozone profile datasets, where measurement errors are of similar magnitude or larger than the long-term ozone trends. In addition, complex coupling between
30 various physical and chemical processes controlling stratospheric ozone concentrations cause large short-term ozone changes. Complications also arise because there are some non-linear changes in stratospheric dynamics as well as chemical constituents. For example, between 2018 and 2021, some of the largest and smallest ozone losses of the recent decades were recorded in both the Arctic and Antarctic polar stratospheres (e.g. Wargan et al., 2020; Wohltmann et al., 2020; Bognar et al., 2021; Weber et al., 2021). Some observational data suggest that there has been a continuous decline in lower stratospheric ozone (Ball et al.,
35 2018, 2020), which could be attributed to changes in stratospheric dynamics (e.g. Chipperfield et al., 2018; Wargan et al., 2018; Orbe et al., 2020; Abalos and de la Cámara, 2020). Atmospheric concentrations of ODSs such as CFC-11 are decreasing at uneven rates (Montzka et al., 2018, 2021) which could variability in ozone trends. Additionally, significant positive trends have been detected in very short-lived substances (VSLS) containing chlorine and bromine that are not controlled by the Montreal Protocol (e.g. Hossaini et al., 2015, 2019).

40 As there is no long-term ozone profile data from a single satellite instrument, various attempts have been made to merge such data from different instruments. However, individual satellite instruments have different temporal and spatial resolution depending on the measurement techniques and retrieval algorithms (e.g. Sofieva et al., 2014; Damadeo et al., 2018). For example, solar occultation instruments (e.g. Stratospheric Aerosol and Gas Experiment (SAGE, McCormick et al., 1989), Halogen Occultation Experiment (HALOE, Russell III et al., 1993)) provide high quality measurements but are constrained by
45 limited spatial coverage. Limb-scanning instruments such as the Microwave Limb Sounder (MLS, Froidevaux et al., 2006b), Scanning Imaging Absorption Spectrometer for Atmospheric Cartography (SCIAMACHY, Bovensmann et al., 1999), Optical Spectrograph and InfraRed Imager System (OSIRIS, Murtagh et al., 2002) provide better spatial coverage but have coarser vertical resolution. A key constraining factor is that only few satellite datasets cover enough overlapping years to remove inter-instrument biases with minimal uncertainty.

50 Hence, Randel and Wu (2007) adopted a novel approach to create a gap-free stratospheric ozone profile data for the 1979–2005 time period. They used SAGE (I and II) satellite profile measurements and polar ozonesondes, together with seasonally varying ozone climatology from Paul et al. (1998) to fill the gaps, to generate multi-variate regression based gap-free ozone anomalies. Later, Cionni et al. (2011) used a similar methodology along with climate model simulations to extend the time series backwards to 1850. The Cionni et al. (2011) data was recommended for the historical CMIP5 simulations, in
55 order to enforce time-dependent ozone variations, for the models that did not include stratospheric chemistry. Hassler et al.

(2008) used a different methodology to create a satellite-based long-term ozone profile dataset. Along with SAGE I and II measurements, they used HALOE and POAM II and III satellite measurements, as well as ozonesonde data from 130 stations, to create a collection of binary data files; also known as the "Binary DataBase of Profiles" (BDBP) version 1.0. Bodeker et al. (2013), updated the BDBP dataset to construct "Bodeker Scientific" or "BS" data. They updated BDBP data by including

60 measurements from the Limb Infrared Monitor of the Stratosphere (LIMS), the Improved Limb Array Spectrometer (ILAS), and ILAS II. They used a multivariate regression model to create different versions of the ozone profile dataset ranging from the surface to 70 km for the 1979-2008 time period. Hassler et al. (2018) revised the Bodeker et al. (2013) dataset by using the TOMCAT chemical transport model (CTM) ozone profiles as a transfer function to capture ozone variability for the period without satellite observations.

65 Another widely used merged data is the Global OZone Chemistry And Related trace gas Data records for the Stratosphere (GOZCARDS, Froidevaux et al., 2015). These are monthly mean zonal mean time series constructed using ozone profiles measurements from several NASA satellite instruments and the Atmospheric Chemistry Experiment - Fourier Transform Spectrometer (ACE-FTS, Bernath et al., 2005). Merging is done primarily by removing average biases between SAGE II and individual data records for overlap periods (Froidevaux et al., 2015). The GOZCARDS data files contain mixing ratios on

70 a pressure–latitude grid (300 hPa to 0.1 hPa), updated later to GOZCARDS v2.2 (Froidevaux et al., 2019).

Davis et al. (2016) adopted a slightly different approach to construct the Stratospheric Water and Ozone Satellite Homogenized (SWOOSH) data. SWOOSH merges stratospheric ozone profile data from solar occultation instruments (SAGE-II/III, HALOE, ACE-FTS) as well as limb-scanning instruments (UARS-MLS, and Aura-MLS). The measurements are homogenized by applying corrections that are calculated from data taken during time periods of instrument overlap. The primary SWOOSH

75 data product consists of monthly mean zonal-mean values on a pressure grid at 2.5, 5 and 10 degree resolution. One of major characteristics of SWOOSH data is that when merging greater weight is given to the instruments that sample more frequently (e.g. Aura-MLS). Filled and unfilled versions of the dataset exist on both geographical and equivalent latitude coordinates.

Several other attempts have been made to merge satellite time series from limb and occultation instruments. In this study we consider the SAGE–CCI–OMPS dataset, described in Sofieva et al. (2017), which includes SAGE II time series and several

80 limb datasets. The OMPS-LP dataset used is produced at the University of Saskatoon (Zawada et al., 2018). The CCI data sets were firstly screened and homogenized in the HARMOZ format and then merged in terms of ozone anomalies with the SAGE II and OMPS-LP observations.

Arosio et al. (2019) also created a merged SCIAMACHY-OMPS limb merged dataset (SCIA-OMPS), which combines these two time series produced at the University of Bremen. They use MLS data series as a transfer function to merge SCIAMACHY

85 with OMPS-LP as these instruments share only two months of overlap, but MLS was not included in the merged dataset. This time series is monthly averaged, covers the period 2002-present and is longitudinally resolved, with a 5° latitude × 20° longitude grid. Due to the similarities in the measurement geometries and techniques, and in the retrieval approaches, a plain debiasing approach was implemented, directly obtaining a long-term ozone time series in appropriate units.

In this paper we present a new data-model method for producing a long-term dataset of stratospheric ozone. We use ozone

90 profile output from a CTM to create a machine-learning-based satellite-corrected long-term chemically (and dynamically)

consistent ozone profile dataset (hereafter, ML-TOMCAT) for the 1979–2020 time period. The CTM setup is described in Section 2, followed by our methodology in Section 3. Comparisons of ML-TOMCAT with some of the other available merged ozone profile datasets are presented in Section 4, with a summary of our key results in Section 5.

2 Model Setup

95 We use chemically consistent monthly mean zonal mean ozone profiles from the TOMCAT chemical transport model (CTM) as the basis dataset. TOMCAT is an off-line three-dimensional (3D) CTM that includes a comprehensive stratospheric chemistry scheme (Chipperfield, 2006). For the present study, the CTM setup is similar to the control simulations used in our recent studies such as Feng et al. (2021); Bognar et al. (2021) and Weber et al. (2021). Briefly, the model is forced with meteorological fields from ERA-5 reanalyses (Hersbach et al., 2020), starting from 1979. Simulations are performed at a 2.8×2.8 degree

100 horizontal resolution with 32 hybrid sigma-pressure levels extending from the surface to about 60 km. For major ODSs and GHGs the model uses time-dependent observed global mean surface mixing ratios (Carpenter et al., 2018) that are treated as well-mixed throughout troposphere. The model also includes the effects of solar flux variability and heterogeneous chemistry on volcanically enhanced stratospheric aerosol as described in Dhomse et al. (2015, 2016). Solar irradiance data are from the NRL2 (Coddington et al., 2016) empirical model and the sulphate aerosol surface area density (SAD) from Luo (2016). The

105 model also includes chlorine and bromine contributions from VSLS as described in Hossaini et al. (2019). The model has been regularly used to study long-term changes in stratospheric trace gases, showing good agreement with various ground-based and satellite datasets (e.g. Mahieu et al., 2014; Chipperfield et al., 2015; Wales et al., 2018; Harrison et al., 2021).

3 Methodology

We use the Random Forest (RF) regression analysis to generate a long-term chemically consistent dataset. The RF is a su-
110 pervised machine learning (ML) algorithm that uses an ensemble of decision trees (e.g. Breiman, 2001; Svetnik et al., 2003). Decision trees use a binary recursive classifying algorithm by splitting observations into two homologous groups.The recursive nature of the algorithm means splitting could be repeated until only two observations are left in the final split. RF uses a bagging technique that means the model consists of many individual trees or learners and aggregated predictions are used for the final prognosis. A distinct advantage of RF regression is that it can behave like a non-linear regression method. As RF adds random-
115 ness to the model, instead of relying on the most important feature, it searches for the best feature among random subsets. This ensures that the final output does not rely heavily on a single learner, thereby avoiding overfitting (e.g. Kotsiantis, 2013). We use Random Forest (RF) Regression from the Python package sklearn (Pedregosa et al., 2011) with two options: *random_state=0*, and *bootstrap=True*. Initially, TOMCAT zonal mean ozone profiles are linearly interpolated in log-pressure space on to 43 equidistant (12 per decade) pressure levels (1000–0.1 hPa, MLS pressure levels), followed by spatial interpolation on to 72
120 SWOOSH latitude bins at $2.5°$ resolution. SWOOSH data is obtained via https://csl.noaa.gov/groups/csl8/swoosh/. Then, we calculate the ozone difference (dO_3) between SWOOSH and model ozone profiles for the 1991-1998 and 2005-2016 time pe-

riods (total 20 years). For the calculation of dO_3 values, we use the gap-filled SWOOSH data product. SWOOSH data ranges from 316 to 1 hPa (31 pressure levels), hence dO_3 for pressure levels below 316 hPa are linearly extrapolated by approximating dO_3 at 1000 hPa to be about 0.01 ppm. Similarly, for levels above 1 hPa, we use linear interpolation assuming that dO_3 at
125 0.1 hPa is about -0.1 ppm, based on bias seen with respect to Aura MLS measurements (Livesey et al., 2020).

For the regression analysis, a 20-year (largely MLS covering) time period is selected in order to avoid heteroscedasticity (i.e. effect of different sampling frequencies/methodologies (e.g. Sofieva et al., 2014; Millán et al., 2016) between different types of satellite datasets) as SWOOSH relies heavily on MLS (UARS and Aura) data records. The regression model has 5 terms: Passive ozone (PO_3), HCl mixing ratio (HCl), methane mixing ratio (CH_4) as well as observation-model total column
130 difference ($dTCO$) and Mg II solar flux term ($MgII$). The PO_3, HCl and CH_4 terms account for possible biases in CTM profiles due to transport in different stratospheric regions (e.g. Strahan et al., 2011; Feng et al., 2021). $dTCO$ is an ideal learner for the lower stratospheric ozone transport as total column ozone measurements have much smaller retrieval errors (e.g. Petropavlovskikh et al., 2019), hence they provide a good constraint for the possible biases in ERA-5 stratospheric transport (e.g. Ploeger et al., 2021). TOMCAT has 203 spectral bins in the photolysis scheme (e.g. Dhomse et al., 2016). Hence, the
135 $MgII$ solar flux term is included to account for possible biases in the representation of the 11-year solar flux variability (e.g. Haigh et al., 2010; Dhomse et al., 2013) or the use of coarse spectral bins (e.g. Sukhodolov et al., 2016).

Overall, there are thus five learners in the regression model that can be represented as:

$$dO_3 = PO_3 + HCl + CH_4 + dTCO + MgII + residuals \qquad (1)$$

PO_3, HCl, CH_4 are TOMCAT monthly mean zonal mean tracers. For the calculation of $dTCO$ we use Copernicus Climate
140 Change Service (C3S) total ozone data (1979–2018). The C3S total column product is a combination of total column data from 15 sensors using gap-filling assimilation methods and is obtained via https://cds.climate.copernicus.eu/cdsapp#!/dataset/satellite-ozone?tab=overview (last access: 1 May 2021). For the years 2019 and 2020, we use level 3 total column data from the Ozone Monitoring instrument (OMI) V3 that is obtained via https://search.earthdata.nasa.gov (last access: 1 May 2021). The Mg II index is obtained from http://www.iup.uni-bremen.de/UVSAT/Datasets/mgii (last access: 1 May 2021). We assume
145 long-term chemical ozone changes are realistically represented by TOMCAT chemistry (e.g. Feng et al., 2007; Chipperfield et al., 2017; Dhomse et al., 2019) and so all the predictor time series are detrended and normalised between 0 to 1.

4 Results

Atmospheric chemical models are ideal tools for understanding/simulating past (and future) ozone changes, as they combine up-to-date knowledge about various physical and chemical processes using a mathematically consistent framework. Different
150 models use different combinations of chemical and dynamical schemes to represent important processes in the atmosphere. However, some of these processes are computationally expensive, hence they are represented by somewhat simplified parameterisations. For example, many chemical models prescribe observation-based sulphate surface area density (SAD) to represent the effects of volcanically enhanced stratospheric aerosol to simulate heterogeneous chemistry which leads to ozone loss

(Dhomse et al., 2015). Many models also prescribe surface concentrations of GHGs and ODSs rather than emission fluxes.
CTMs such as TOMCAT use dynamical forcing fields from (re)analyses datasets such as ERA-Interim or ERA-5. Hence CTMs
are subject to the possible inhomogeneities due changes in the number of assimilated observations, as well as other deficiencies
(e.g. missing processes) in the forward model used in the assimilation system. On the other hand, observational datasets are also
subject to errors associated with the measurement techniques, instrument degradation and retrieval algorithms. Hence, almost
all models may be expected to show a bias against observational data records, either because of model deficiencies or errors
in the observational datasets. However, chemical models do use a consistent chemical scheme, so we can assume that model-
observations differences are largely due to uncertainties in the forcing fields such as meteorology (e.g. winds, temperature) and
chemical parameterisations (e.g. reaction rates, solar fluxes, photolysis schemes). CTMs have the distinct advantage in terms of
dynamics as they are forced with up-to-date reanalysis data, although with the above-noted caveat of possible inhomogeneities
in observations used in the assimilation systems. Therefore, we train the ML algorithm on the model-observation differences
for the period that has relatively good temporal sampling. Estimated parameters are then used to simulate differences for the
entire (1979–2020) time period. In this section, we analyse model-observation biases associated with individual predictors and
compare the ML-corrected data against a variety of observation-based datasets.

4.1 Model biases

Figure 1 shows climatological (2006–2020) monthly zonal mean differences between SWOOSH and TOMCAT ozone profiles
(SWOOSH minus TOMCAT). Ozone differences show almost symmetrical structure with positive biases in the upper strato-
sphere and negative biases in the lower stratosphere. The largest positive biases (up to 0.8 ppm) occur in the tropical upper
stratosphere (around 3 hPa) and they remain positive throughout the year. The ozone lifetime at these altitudes is less than
a day, hence the observed biases might be associated with deficiencies in the photochemical reactions in the model. At this
altitude, ozone production is largely controlled by fluxes below 240 nm while longer wavelengths control ozone destruction
(e.g. Haigh et al., 2010). Therefore, negative ozone biases in the upper stratosphere are most probably due to uncertainties in
the solar irradiances and/or photolysis cross sections that control ozone production (e.g. Brasseur and Solomon, 2006). Fur-
thermore, in this region of the atmosphere, ozone chemistry is very temperature-dependent (e.g. Stolarski et al., 2010; Dhomse
et al., 2013, 2016), hence model ozone biases could be due to uncertainties in temperature-dependent reaction rates (e.g. Ghosh
et al., 1997).

In the lower stratosphere the ozone lifetime ranges from months to years, hence positive biases in the modelled ozone could
be due to a combination of both dynamics and chemistry. First, reduced overhead ozone could increase lower stratospheric
ozone via the self-healing effect, i.e. increased ultra-violet radiation increases ozone production at lower altitudes (e.g. Haigh,
1994). Second, ozone is primarily produced in the tropical stratosphere, and its downward transport is controlled by the QBO
(e.g. Tian et al., 2006), whereas transport towards mid-high latitudes is determined by the strength of the Brewer-Dobson (BD)
circulation (e.g. Holton et al., 1995; Weber et al., 2003, 2011; Dhomse et al., 2006) which increases its lifetime considerably.
Hence, ozone biases in the lower stratosphere are likely due to the incomplete representation of various circulation pathways

in TOMCAT either due to model resolution or missing representation of key physical process in the ERA-5 reanalysis scheme (e.g. Mitchell et al., 2020) which impacts the meteorology used in the CTM.

4.2 Learner contributions

190 As the exact causes of model ozone biases are still not well understood, we use the RF model to remove these biases in TOMCAT. The RF regression model coefficients are derived using 20 years (1991–1998, 2006–2018) for which SWOOSH data includes a large number of observational profiles especially from MLS on the UARS and Aura satellite platforms. The RF regression model uses 20-years of monthly data with 80% and 20% of data points being used for training and testing, respectively. Estimated RF regression coefficients are then used to calculate model biases for the entire 42-year time period

195 (1979–2020). RF-calculated ozone biases are then added to the TOMCAT time series to create the long-term gap-free dataset, hereafter labelled ML-TOMCAT.

Figure 2 shows how much variance (or R^2) of the data the RF regression model is able to explain, along with regression coefficients for individual learners. For example, R^2 equal to 0.8 indicates that RF regression model is able to explain 80% of the biases in modelled ozone relative to SWOOSH data for the 20 years of the training period. Overall, the RF regression model

200 performance is consistently high ($R^2 > 0.8$) throughout the stratosphere, except for the mid-stratosphere which is a transition region where the TOMCAT ozone biases change from positive to negative. At high northern latitudes, mid-stratospheric R^2 values decrease to 0.6. However, since TOMCAT-SWOOSH differences are much smaller here, a RF-based correction has a minimal impact on the quality of ML-TOMCAT ozone profiles.

Additionally, as expected, the RF regression coefficients are significant in different regions of the stratosphere for various

205 learners. The passive ozone tracer seems to show the largest coefficients in the tropical mid-stratosphere, as well as varying contributions in different regions of the stratosphere. The passive ozone contribution in the tropical mid-stratospheric could be linked to the incomplete representation of NO_x-related chemical changes in TOMCAT and/or seasonal changes in the stratospheric transport in the re-analysis (e.g. Galytska et al., 2019). The HCl tracer shows significant coefficients in the upper stratosphere, where the ClO ozone loss cycle is important as well as in the tropical lower stratosphere. The CH_4 tracer term

210 seems to show significant coefficients in the lowermost stratosphere (just above the tropopause) as well as a significant contribution around mid-latitude sub-tropics. The CH_4 tracer contribution resembles a QBO-induced secondary circulation pattern. Interestingly, the solar term shows the largest coefficients in the mid-latitude upper stratosphere rather than in the tropical upper stratosphere, suggesting solar flux variability has only a minor contribution to the model-observation biases. As expected the $dTCO$ term shows the largest contribution in the lowermost stratosphere, especially in the tropical and polar regions. Inter-

215 estingly, ozone anomalies in these regions show good agreement with various satellite-based data sets (e.g. Chipperfield et al., 2017, 2018; Li et al., 2020; Feng et al., 2021), and model biases are much smaller. This means although $dTCO$ coefficients are largest in the lowermost stratosphere, the overall bias correction contribution remains relatively small.

4.3 Comparison against merged datasets

After analysing the regression coefficients, we now present a comparison between ML-TOMCAT and available satellite-based
220 long-term datasets. Due to key differences between satellite measurement techniques, ozone profiles are retrieved either at
height or pressure levels and either in units of mixing ratio or number density. For example, MLS retrieves profiles of ozone
mixing ratio on pressure levels whereas SAGE retrieves profiles of number density on height levels. Hence, merging these
different datasets needs pressure, temperature or height information at a given co-location from an external source such as
reanalysis data. GOZCARDS and SWOOSH datasets use MERRA2 reanalysis data to convert SAGE II ozone number den-
225 sity profiles on fixed pressure levels (Damadeo et al., 2013). Since, ML-TOMCAT is based on modelled ozone profiles as a
function of pressure, conversion to height coordinates is straightforward. In particular, ML-TOMCAT data was processed on
corresponding grids/units using ERA-5 geopotential, temperature and pressure fields that are used to drive TOMCAT.

This subsection consists of two parts. First we compare ML-TOMCAT profiles with datasets using pressure co-ordinate
systems (e.g. SWOOSH, GOZCARDS), followed by comparisons with height-based datasets (SAGE–CCI–OMPS, SCIA-
230 OMPS, BSVert).

4.3.1 Comparison with pressure level data

As noted earlier, we used only 20 years of SWOOSH data to train the RF model. Hence, the next obvious step is to compare
ML-TOMCAT ozone with SWOOSH over the full time period. Figure 3 compares relative differences (in percentage) of ML-
TOMCAT with GOZCARDS and SWOOSH, respectively, as a function of latitude and pressure. ML-TOMCAT shows slightly
235 positive biases in the middle stratosphere and somewhat negative biases in the upper and lower stratosphere with respect to both
SWOOSH and GOZCARDS data. The largest biases (up to 10%) are observed in the tropical lowermost stratosphere as well
as polar latitudes. However, these largest differences in the tropical lowermost stratosphere (and upper troposphere) cannot be
correctly validated as most satellite retrievals show largest uncertainties in this region (Rahpoe et al., 2015; Steinbrecht et al.,
2017; Sofieva et al., 2021). Similarly, for non-MLS period, the biases in the polar stratosphere could be due to the lack of
240 observational ozone profiles during polar nights.

Figure 4 shows TOMCAT, ML-TOMCAT, SWOOSH and GOZCARDS ozone time series over the equator (0° lat) at 3
pressure levels (1, 10 and 50 hPa). Supplementary Figures S1 to S10 show similar comparisons at 15°N, 15°S, 30°N, 30°S,
45°N, 45°S, 60°N, 60°S, 75°N, and 75°S latitude bins. The grey shaded area indicates the standard deviation of the ozone
values within each bin for the GOZCARDS time series. The green shaded areas indicates the root mean square uncertainty
245 of the combined datasets for each bin in SWOOSH data (σ_{rmss} in Davis et al. (2016)). Overall, there is a good agreement
between the ML-TOMCAT, GOZCARDS and SWOOSH time series. As seen in Figure 1, ML-TOMCAT shows significant
improvements in the tropical stratosphere.

A peculiar detail of Figure 4 is that the standard deviation in the SWOOSH time series is largest during the 1991-1999 time
period which could be due to a combination of various factors. First, for this time period, SWOOSH largely relies on UARS
250 MLS ozone profiles. As UARS was launched in a near-circular orbit, the MLS latitudinal coverage was either 34°S – 80°N or

34°N – 80°S, depending on the spacecraft yaw position which alternated every 36 days (Barath et al., 1993). Hence, for lower latitude ozone profiles during 1990s, SWOOSH relies on SAGE II and HALOE profiles with a low spatial sampling. Second, UARS MLS ozone profiles are retrieved at only six levels per pressure decade (Livesey et al., 2003) instead of 12 levels per decade for Aura MLS (see https://mls.jpl.nasa.gov/data/v5-0_data_quality_document.pdf). Third, significant enhancement in
255   the stratospheric aerosol loading following the Mt. Pinatubo eruption in June 1991, led to larger retrieval errors. Even with those uncertainties in SWOOSH (and GOZCARDS), ML-TOMCAT is generally close to the satellite-based datasets for the entire time period and the agreement with satellite data is greatly improved in comparison with the original TOMCAT model data. Supplementary Figures S1 to S10 also show an excellent agreement between ML-TOMCAT and GOZCARDS/SWOOSH datasets.

260   As discussed above, the ML-TOMCAT training dataset (SWOOSH) has large uncertainties in the 1990s. Next we scrutinise percentage differences between GOZCARDS and ML-TOMCAT on the same pressure levels. Figure 5 shows relative differences between TOMCAT, ML-TOMCAT and SWOOSH ozone time series with respect to GOZCARDS. As seen earlier, TOMCAT ozone shows up to 40% positive biases in the lower stratosphere and 10% negative biases in the upper stratosphere (also seen in Figure 1). In contrast, ML-TOMCAT biases are well below 5% at all levels. At 50 hPa, TOMCAT biases seems
265   to follow QBO-type oscillations that are correctly removed in ML-TOMCAT. Similarly, at 1 hPa TOMCAT differences show some uneven variations that could be linked to the inhomogeneities in the ERA-5 dynamical fields that are used to force TOMCAT. Furthermore, ML-TOMCAT differences show much smaller and almost linear biases at 1 hPa and lie well within the spread of GOZCARDS data.

    Interestingly, although both GOZCARDS and SWOOSH are created by merging nearly identical datasets, there are differ-
270   ences between them which are largest for the 1984 to 2004 time period. This indicates that even slight differences in merging methods leads to large differences in the merged dataset. Although we use completely independent output from a CTM as a basis dataset, GOZCARDS-ML-TOMCAT differences are not significantly different to GOZCARDS-SWOOSH differences, especially at 10 and 50 hPa.

    Another notable feature in Figure 5 is that at 50 hPa, ML-TOMCAT shows largest differences during 2020, which could
275   be associated with the biases in ERA-5 dynamics during that period. A TOMCAT sensitivity simulation forced with ECMWF operational analysis data shows better agreement with MLS ozone variation during this period (not shown). In addition, larger differences seen during 1984 (50 hPa), 1988 (10 hPa) and 1996-1999 (1 hPa) are most probably associated with SAGE II sampling issues and/or inhomogeneities in ERA-5 dynamical fields.

4.3.2 Comparison with height level data

280   In this section, we compare ML-TOMCAT ozone profiles against height-based merged satellite datasets. Figure 6 shows relative difference between TOMCAT/ML-TOMCAT vs SAGE-CCI-OMPS (Sofieva et al., 2017), BSVert (Hassler et al., 2018) as well as SCIA-OMPS (Arosio et al., 2018) datasets as a function of altitude and latitude. The top panels (a and b) compare the mean relative difference between the SAGE-CCI-OMPS dataset, TOMCAT and ML-TOMCAT, respectively. Here TOMCAT shows large positive biases (up to 20%) in the lowermost stratosphere and negative biases (up to 15%) in the upper stratosphere. On

285 the other hand, ML-TOMCAT shows only ±10% biases throughout the stratosphere. Larger biases are seen in the Antarctic stratosphere that could be attributed to the limited observational ozone profiles used to construct the height-based merged satellite data products. Interestingly, ML-TOMCAT shows largest biases (up to 30%) w.r.t. BSVert data set, though TOMCAT profiles (forced with ERA-Interim) are used to construct the BSVert data set (Hassler et al., 2018). In addition, in the lowermost stratosphere, biases are negative in the SH latitudes and positive biases in the NH mid-high latitudes. Hence, a contributing

290 factor for these hemispherically asymmetric biases with respect to BSVert ozone profiles might be differences between ERA-Interim and ERA-5 reanalysis data (e.g. Ploeger et al., 2021) that are used to force these two datasets. The negative values in relative differences in the lower tropical stratosphere shown with respect to SCIA-OMPS dataset in the fourth panel is systematic throughout the time series and is thought to be related to two reasons. The first one is the coarse vertical grid (corresponding to SCIAMACHY vertical resolution of 3.3 km) which makes it sensitive to the interpolation onto TOMCAT

295 grid. The second is a different usage of MLS in the merging procedure implemented for SCIA-OMPS w.r.t. SWOOSH, so that ML-TOMCAT trained over MLS period using SWOOSH shows a negative bias w.r.t. SCIA-OMPS, which however doesn't show such large bias w.r.t. MLS.

Figure 7 compares TOMCAT and ML-TOMCAT profiles with the three height-based ozone datasets with a focus on the equator (0° latitude). Supplementary Figures S11 to S20 show similar comparisons for 15°N, 15°S, 30°N, 30°S, 45°N, 45°S,

300 60°N, 60°S, 75°N, and 75°S latitude bins. Figure 8 displays the respective relative differences with respect to the SAGE-CCI-OMPS dataset which in this case is taken as a reference. In this way it is possible to evaluate the improvement introduced by applying the ML algorithm but also have an estimation of the discrepancies between different merged datasets, which is expected to be on the order of 5-10%. With respect to the comparison with the datasets on pressure vertical coordinate, the scatter between the time series is larger here, due to the larger variety of different satellites available to produce the merged

305 products and the fact that they have not been used in the ML training.

At about 45km in the tropics the ML algorithm seems to over-correct the negative bias shown by TOMCAT, leading to generally higher ozone values with respect to the other datasets, especially in the first half of the time series. In the middle stratosphere we find the best agreement between SAGE-CCI-OMPS and ML-TOMCAT; here the expected discrepancies among the merged datasets are comparable to the differences observed between ML-TOMCAT and SAGE-CCI-OMPS. At the peak

310 of the ozone number profile around 25 km, we notice generally lower values for ML-TOMCAT, on average by 5%. Similar biases are observed at mid-high latitude as well as seen in Supplementary Figures S11 to S20. The strong seasonal cycle seen in the TOMCAT difference with respect to the merged datasets is largely reduced by ML-TOMCAT at this altitude.

4.3.3 Polar regions

The use of ML-TOMCAT helps to fill the observational gaps especially in atmospheric regions with lack of observations and

315 before the beginning of the 21st century, when satellite measurements were sparser. For example, polar regions during local winter cannot be observed by limb observations based on scattered sunlight. Instruments such as Aura MLS and the Sounding of the Atmosphere using Broadband Emission Radiometry (SABER, Rong et al., 2008) are generally used to fill this gap over the last two decades. Also for chemical models, complexities are also associated with the denitrification and dehydration (or

chlorine activation) schemes that determine heterogeneous ozone losses (Grooß et al., 2018). Though most of our earlier studies
320 showed that TOMCAT is able to polar ozone losses quite realistically (e.g. Feng et al., 2007; Chipperfield et al., 2015, 2017;
Dhomse et al., 2019), some systematic biases in polar stratosphere were noted in Feng et al. (2021); Weber et al. (2021). Figure
9 compares ozone at 18 km over the North Pole which also demonstrates the good agreement between ML-TOMCAT and MLS
in this region, both during local summer and winter. In the bottom panel, the ozone sub-column over the South Pole (poleward
of 70°S latitude) intergrated between 12 and 20 km, for TOMCAT, ML-TOMCAT and MLS averaged over September-October
325 months. The good agreement between MLS and ML-TOMCAT during the ozone hole period is observed for most of the years.
ML-TOMCAT enables the reconstruction of the large ozone losses which occurred in the 1980s during a phase when ozone
depleting substances were on a rapid rise before the implementation of the Montreal Protocol and their phase out.

5 Summary and Conclusions

Stratospheric ozone concentrations are affected by many short- and long-term processes, hence high quality ozone profile
330 datasets are needed for accurate attribution studies. Though satellite instruments provide global measurements, due to their
short mission durations various merging methodologies have been adopted to create homogeneous and gap-free long-term
ozone profile datasets. Individual merging methodologies have distinct advantages and disadvantages. Atmospheric chemical
models are also able to simulate chemically consistent long-term datasets, but they are prone to the deficiencies associated with
the simplified parameterisations and uncertain parameters.

335   Here we have used TOMCAT CTM ozone profiles and a RF regression model to create gap-free ozone profile dataset (ML-
TOMCAT) for 1979-2020. The RF is trained for the ozone difference between the SWOOSH and TOMCAT ozone profiles
by selecting 20 years of MLS measurement time periods (UARS-MLS and AURA-MLS). RF show consistent performance
throughout the stratosphere, except at high latitudes and mid-latitude mid-stratosphere. Overall, ML-TOMCAT shows excellent
agreement with SWOOSH for the entire time period (1984–2020), though somewhat larger differences are visible for the period
340 where limited ozone measurements are available for SWOOSH construction. We also find that ML-TOMCAT shows better
agreement with satellite-based merged datasets which use pressure as the vertical coordinate (e.g. SWOOSH, GOZCARDS)
but weaker agreement with the datasets which use height (e.g. SAGE–CCI–OMPS, SCIA-OMPS). We find that at almost
all stratospheric levels ML-TOMCAT ozone concentrations are well within uncertainties in the observational datasets. The
ML-TOMCAT dataset is thus ideally suited for the evaluation of model ozone profiles from the tropopause to 0.1 hPa. ML-
345 TOMCAT V1.0 ozone profile data on pressure and height levels in mixing ratios and number density units is publicly available
via https://zenodo.org/record/4997959#.YNzleUlKiUk.

Data availability. We thank Sean Davies for SWOOSH data which is publicly available via https://csl.noaa.gov/groups/csl8/swoosh/. We
also thank Lucien Froidevaux (Lucien.Froidevaux@jpl.nasa.gov) for GOZCARDS v2 data. SAGE–CCI–OMPS was obtained via http://www.esa-

ozone-cci.org. SCIA-OMPS data is available upon email request to AR or MW. BSVert data was obtained from https://zenodo.org/record/1217184
350 (Birgit et al., 2018). ML-TOMCAT V1.0 data is publicly available via https://zenodo.org/record/#.YNzleUlKiUk (Dhomse et al., 2021)

Acknowledgements. SD was supported by the NERC SISLAC project (NE/R001782/1). The financial support of part of this work from the
State of Bremen, DAAD grant (CA), and ESA SOLVE Living Planet Fellowship (CA) is gratefully acknowledged. The UB contribution
also benefited from the DFG VolImpact and BMBF SynopSys Projects. We thank NASA, NOAA and ESA for GOZCARDS, SWOOSH and
SAGE-CCI-OMPS data products. We thank the European Centre for Medium-Range Weather Forecasts for providing their analyses. The
355 model simulations were performed on the UK national Archer and Leeds Arc4 HPC systems.

References

Abalos, M. and de la Cámara, A.: Twenty-First Century Trends in Mixing Barriers and Eddy Transport in the Lower Stratosphere, Geophysical Research Letters, 47, e2020GL089 548, 2020.

Anderson, J., Russell Iii, J., Solomon, S., and Deaver, L.: Halogen Occultation Experiment confirmation of stratospheric chlorine decreases
360   in accordance with the Montreal Protocol, Journal of Geophysical Research: Atmospheres, 105, 4483–4490, 2000.

Arosio, C., Rozanov, A., Malinina, E., Eichmann, K.-U., von Clarmann, T., and Burrows, J. P.: Retrieval of ozone profiles from OMPS limb
scattering observations, Atmospheric Measurement Techniques, 11, 2135–2149, https://doi.org/10.5194/amt-11-2135-2018, 2018.

Arosio, C., Rozanov, A., Malinina, E., Weber, M., and Burrows, J. P.: Merging of ozone profiles from SCIAMACHY, OMPS and SAGE II
observations to study stratospheric ozone changes, Atmospheric Measurement Techniques, 12, 2423–2444, https://doi.org/10.5194/amt-
365   12-2423-2019, 2019.

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, Atmospheric Chemistry and Physics, 18,
1379–1394, https://doi.org/10.5194/acp-18-1379-2018, 2018.

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

Barath, F. T., Chavez, M. C., Cofield, R. E., Flower, D. A., Frerking, M. A., Gram, M. B., Harris, W. M., Holden, J. R., Jarnot,
R. F., Kloezeman, W. G., Klose, G. J., Lau, G. K., Loo, M. S., Maddison, B. J., Mattauch, R. J., McKinney, R. P., Peckham,
375   G. E., Pickett, H. M., Siebes, G., Soltis, F. S., Suttie, R. A., Tarsala, J. A., Waters, J. W., and Wilson, W. J.: The Upper Atmo-
sphere Research Satellite microwave limb sounder instrument, Journal of Geophysical Research: Atmospheres, 98, 10 751–10 762,
https://doi.org/https://doi.org/10.1029/93JD00798, 1993.

Bernath, P. F., McElroy, C. T., Abrams, M. C., Boone, C. D., Butler, M., Camy-Peyret, C., Carleer, M., Clerbaux, C., Coheur, P.-F., Colin, R.,
DeCola, P., DeMazière, M., Drummond, J. R., Dufour, D., Evans, W. F. J., Fast, H., Fussen, D., Gilbert, K., Jennings, D. E., Llewellyn,
380   E. J., Lowe, R. P., Mahieu, E., McConnell, J. C., McHugh, M., McLeod, S. D., Michaud, R., Midwinter, C., Nassar, R., Nichitiu, F.,
Nowlan, C., Rinsland, C. P., Rochon, Y. J., Rowlands, N., Semeniuk, K., Simon, P., Skelton, R., Sloan, J. J., Soucy, M.-A., Strong,
K., Tremblay, P., Turnbull, D., Walker, K. A., Walkty, I., Wardle, D. A., Wehrle, V., Zander, R., and Zou, J.: Atmospheric Chemistry
Experiment (ACE): Mission overview, Geophysical Research Letters, 32, https://doi.org/https://doi.org/10.1029/2005GL022386, 2005.

Birgit, B., Stefanie, K., Greg, B., Jared, L., Kage, N., Sean, D., Martyn, C., Sandip, D., and Martin, D.: BSVerticalOzone database,
385   https://doi.org/10.5281/zenodo.1217184, 2018.

Bodeker, G., Hassler, B., Young, P., and Portmann, R. W.: A vertically resolved, global, gap-free ozone database for assessing or constraining
global climate model simulations, Earth System Science Data, 5, 31–43, 2013.

Bognar, K., Alwarda, R., Strong, K., Chipperfield, M. P., Dhomse, S. S., Drummond, J. R., Feng, W., Fioletov, V., Goutail, F., Herrera, B.,
Manney, G. L., McCullough, E. M., Millán, L. F., Pazmino, A., Walker, K. A., Wizenberg, T., and Zhao, X.: Unprecedented Spring 2020
390   Ozone Depletion in the Context of 20 Years of Measurements at Eureka, Canada, Journal of Geophysical Research: Atmospheres, 126,
e2020JD034 365, https://doi.org/https://doi.org/10.1029/2020JD034365, e2020JD034365 2020JD034365, 2021.

Bovensmann, H., Burrows, J., Buchwitz, M., Frerick, J., Noël, S., Rozanov, V., Chance, K., and Goede, A.: SCIAMACHY: Mission objectives and measurement modes, Journal of the atmospheric sciences, 56, 127–150, 1999.

Brasseur, G. P. and Solomon, S.: Aeronomy of the middle atmosphere: Chemistry and physics of the stratosphere and mesosphere, vol. 32, Springer Science & Business Media, 2006.

Breiman, L.: Random forests, Machine learning, 45, 5–32, 2001.

Carpenter, L. J., Daniel, J. S., Fleming, E. L., Hanaoka, T., Ju, H., Ravishankara, A. R., Ross, M. N., Tilmes, S., Wallington, T. J., and Wuebbles, D. J.: Scenarios and information for policy makers, in: Scientific Assessment of Ozone Depletion: 2018, World Meteorological Organization, Global Ozone Research and Monitoring Project - Report No. 58, chap. 6, World Meteorological Organization/UNEP, https://public.wmo.int/en/resources/library/scientific-assessment-of-ozone-depletion-2018, 2018.

Chipperfield, M. P.: New version of the TOMCAT/SLIMCAT off-line chemical transport model: Intercomparison of stratospheric tracer experiments, Quarterly Journal of the Royal Meteorological Society, 132, 1179–1203, https://doi.org/10.1256/qj.05.51, 2006.

Chipperfield, M. P., Dhomse, S. S., Feng, W., McKenzie, R. L., Velders, G. J. M., and Pyle, J. A.: Quantifying the ozone and ultraviolet benefits already achieved by the Montreal Protocol., Nature communications, 6, 7233, https://doi.org/10.1038/ncomms8233, 2015.

Chipperfield, M. P., Bekki, S., Dhomse, S., Harris, N. R. P., 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, Geophysical Research Letters, 45, 5718–5726, https://doi.org/10.1029/2018GL078071, 2018.

Cionni, I., Eyring, V., Lamarque, J.-F., Randel, W., Stevenson, D., Wu, F., Bodeker, G., Shepherd, T., Shindell, D., and Waugh, D.: Ozone database in support of CMIP5 simulations: results and corresponding radiative forcing, Atmospheric Chemistry and Physics, 11, 11 267–11 292, 2011.

Coddington, O., Lean, J. L., Pilewskie, P., Snow, M., and Lindholm, D.: A solar irradiance climate data record, Bulletin of the American Meteorological Society, 97, 1265–1282, https://doi.org/10.1175/BAMS-D-14-00265.1, 2016.

Damadeo, R., Petropavlovskikh, I., Godin-Beekmann, S., Hubert, D., Sofieva, V., and Hassler, B.: The Long-term Ozone Trends and Uncertainties in the Stratosphere (LOTUS) SPARC activity: Lessons Learned, in: EGU General Assembly Conference Abstracts, p. 9953, 2018.

Damadeo, R. P., Zawodny, J. M., Thomason, L. W. L., and Iyer, N.: SAGE version 7.0 algorithm: application to SAGE II, Atmospheric Measurement Techniques, 6, 3539–3561, https://doi.org/10.5194/amt-6-3539-2013, 2013.

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 System Science Data, 8, 461–490, https://doi.org/10.5194/essd-8-461-2016, 2016.

Dhomse, S., Weber, M., Wohltmann, I., Rex, M., and Burrows, J. P.: On the possible causes of recent increases in northern hemispheric total ozone from a statistical analysis of satellite data from 1979 to 2003, Atmospheric chemistry and physics, 6, 1165–1180, https://doi.org/10.5194/acp-6-1165-2006, 2006.

Dhomse, S., Feng, W., Montzka, S. A., Hossaini, R., Keeble, J., Pyle, J., Daniel, J., and Chipperfield, M.: Delay in recovery of the Antarctic ozone hole from unexpected CFC-11 emissions, Nature communications, 10, 1–12, 2019.

Dhomse, S. S., Chipperfield, M. P., Feng, W., Ball, W. T., Unruh, Y. C., Haigh, J. D., Krivova, N. A., Solanki, S. K., and Smith, A. K.:
Stratospheric O3 changes during 2001—2010: the small role of solar flux variations in a chemical transport model, Atmospheric Chemistry
and Physics, 13, 10 113–10 123, https://doi.org/10.5194/acp-13-10113-2013, 2013.

Dhomse, S. S., Chipperfield, M. P., Feng, W., Hossaini, R., Mann, G. W., and Santee, M. L.: Revisiting the hemispheric asymmetry in
midlatitude ozone changes following the Mount Pinatubo eruption: A 3-D model study, Geophysical Research Letters, 42, 3038–3047,
https://doi.org/10.1002/2015GL063052, 2015.

Dhomse, S. S., Chipperfield, M. P., Damadeo, R. P., Zawodny, J. M., Ball, W. T., Feng, W., Hossaini, R., Mann, G. W., and Haigh, J. D.:
On the ambiguous nature of the 11-year solar cycle signal in upper stratospheric ozone, Geophysical Research Letters, 43, 7241–7249,
https://doi.org/10.1002/2016GL069958, 2016.

Dhomse, S. S., Kinnison, D., Chipperfield, M. P., Salawitch, R. J., Cionni, I., Hegglin, M. I., Abraham, N. L., Akiyoshi, H., Archibald, A. T.,
Bednarz, E. M., Bekki, S., Braesicke, P., Butchart, N., Dameris, M., Deushi, M., Frith, S., Hardiman, S. C., Hassler, B., Horowitz, L. W.,
Hu, R.-M., Jöckel, P., Josse, B., Kirner, O., Kremser, S., Langematz, U., Lewis, J., Marchand, M., Lin, M., Mancini, E., Marécal, V.,
Michou, M., Morgenstern, O., O'Connor, F. M., Oman, L., Pitari, G., Plummer, D. A., Pyle, J. A., Revell, L. E., Rozanov, E., Schofield,
R., Stenke, A., Stone, K., Sudo, K., Tilmes, S., Visioni, D., Yamashita, Y., and Zeng, G.: Estimates of ozone return dates from Chemistry-
Climate Model Initiative simulations, Atmospheric Chemistry and Physics, 18, 8409–8438, https://doi.org/10.5194/acp-18-8409-2018,
2018.

Dhomse, S. S., Chipperfield, M. P., Feng, W., Arosio, C., Weber, M., and Rozanov, A.: ML-TOMCAT V1.0: Machine-Learning-Based
Satellite- Corrected Global Stratospheric Ozone Profile Dataset, https://doi.org/10.5281/zenodo.4997959, 2021.

Feng, W., Chipperfield, M. P., Davies, S., von der Gathen, P., Kyrö, E., Volk, C. M., Ulanovsky, A., and Belyaev, G.: Large chemical ozone
loss in 2004/2005 Arctic winter/spring, Geophysical Research Letters, 34, https://doi.org/10.1029/2006GL029098, 2007.

Feng, W., Dhomse, S. S., Arosio, C., Weber, M., Burrows, J. P., Santee, M. L., and Chipperfield, M. P.: Arctic ozone depletion in 2019/20:
Roles of chemistry, dynamics and the Montreal Protocol, Geophysical Research Letters, https://doi.org/10.1029/2020GL091911, 2021.

Froidevaux, L., Livesey, N., Read, W., Salawitch, R., Waters, J., Drouin, B., MacKenzie, I., Pumphrey, H., Bernath, P., Boone, C., et al.:
Temporal decrease in upper atmospheric chlorine, Geophysical Research Letters, 33, 2006a.

Froidevaux, L., Livesey, N. J., Read, W. G., Jiang, Y. B., Jimenez, C., Filipiak, M. J., Schwartz, M. J., Santee, M. L., Pumphrey, H. C., Jiang,
J. H., et al.: Early validation analyses of atmospheric profiles from EOS MLS on the Aura satellite, IEEE Transactions on Geoscience and
Remote Sensing, 44, 1106–1121, 2006b.

Froidevaux, L., Anderson, J., Wang, H.-J., Fuller, R., Schwartz, M., Santee, M., Livesey, N., Pumphrey, H., Bernath, P., Russell III, J., et al.:
Global OZone Chemistry And Related trace gas Data records for the Stratosphere (GOZCARDS): methodology and sample results with
a focus on HCl, H 2 O, and O 3, Atmospheric Chemistry and Physics, 15, 10 471–10 507, 2015.

Froidevaux, L., Kinnison, D. E., Wang, R., Anderson, J., and Fuller, R. A.: Evaluation of CESM1 (WACCM) free-running and specified
dynamics atmospheric composition simulations using global multispecies satellite data records, Atmospheric Chemistry and Physics, 19,
4783–4821, 2019.

Galytska, E., Rozanov, A., Chipperfield, M. P., Dhomse, S. S., Weber, M., Arosio, C., Feng, W., and Burrows, J. P.: Dynamically con-
trolled ozone decline in the tropical mid-stratosphere observed by SCIAMACHY, Atmospheric Chemistry and Physics, 19, 767–783,
https://doi.org/10.5194/acp-19-767-2019, 2019.

Ghosh, S., Pyle, J. A., and Good, P.: Temperature dependence of the ClO concentration near the stratopause, Journal of Geophysical Research:
Atmospheres, 102, 19 207–19 216, https://doi.org/https://doi.org/10.1029/97JD01099, 1997.

Grooß, J.-U., Müller, R., Spang, R., Tritscher, I., Wegner, T., Chipperfield, M. P., Feng, W., Kinnison, D. E., and Madronich, S.: On the discrepancy of HCl processing in the core of the wintertime polar vortices, Atmospheric Chemistry and Physics, 18, 8647–8666, https://doi.org/10.5194/acp-18-8647-2018, 2018.

Haigh, J.: The role of stratospheric ozone in modulating the solar radiative forcing of climate, Nature, 370, 544–546, https://doi.org/https://doi.org/10.1038/370544a0, 1994.

Haigh, J., Winning, A., Toumi, R., and Harder, J.: An influence of solar spectral variations on radiative forcing of climate, Nature, 467, https://doi.org/https://doi.org/10.1038/370544a0, 2010.

Harrison, J. J., Chipperfield, M. P., Boone, C. D., Dhomse, S. S., and Bernath, P. F.: Fifteen Years of HFC-134a Satellite Observations: Comparisons With SLIMCAT Calculations, Journal of Geophysical Research: Atmospheres, 126, e2020JD033 208, https://doi.org/https://doi.org/10.1029/2020JD033208, e2020JD033208 2020JD033208, 2021.

Hassler, B., Bodeker, G., and Dameris, M.: A new global database of trace gases and aerosols from multiple sources of high vertical resolution measurements, Atmospheric Chemistry and Physics, 8, 5403–5421, 2008.

Hassler, B., Kremser, S., Bodeker, G. E., Lewis, J., Nesbit, K., Davis, S. M., Chipperfield, M. P., Dhomse, S. S., and Dameris, M.: An updated version of a gap-free monthly mean zonal mean ozone database, Earth System Science Data, 10, 1473–1490, 2018.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/https://doi.org/10.1002/qj.3803, 2020.

Holton, J. R., Haynes, P. H., McIntyre, M. E., Douglass, A. R., Rood, R. B., and Pfister, L.: Stratosphere-troposphere exchange, Reviews of Geophysics, 33, 403–439, https://doi.org/10.1029/95RG02097, 1995.

Hossaini, R., Chipperfield, M. P., Montzka, S. A., Rap, A., Dhomse, S., and Feng, W.: Efficiency of short-lived halogens at influencing climate through depletion of stratospheric ozone, Nature Geoscience, advance on, 186–190, https://doi.org/10.1038/ngeo2363, 2015.

Hossaini, R., Atlas, E., Dhomse, S. S., Chipperfield, M. P., Bernath, P. F., Fernando, A. M., Mühle, J., Leeson, A. A., Montzka, S. A., Feng, W., Harrison, J. J., Krummel, P., Vollmer, M. K., Reimann, S., O'Doherty, S., Young, D., Maione, M., Arduini, J., and Lunder, C. R.: Recent Trends in Stratospheric Chlorine From Very Short-Lived Substances, Journal of Geophysical Research: Atmospheres, 124, 2318–2335, https://doi.org/10.1029/2018JD029400, 2019.

Kotsiantis, S. B.: Decision trees: a recent overview, Artificial Intelligence Review, 39, 261–283, 2013.

Li, Y., Chipperfield, M. P., Feng, W., Dhomse, S. S., Pope, R. J., Li, F., and Guo, D.: Analysis and attribution of total column ozone changes over the Tibetan Plateau during 1979–2017, Atmospheric Chemistry and Physics, 20, 8627–8639, https://doi.org/10.5194/acp-20-8627-2020, 2020.

Livesey, N. J., Read, W. G., Froidevaux, L., Waters, J. W., Santee, M. L., Pumphrey, H. C., Wu, D. L., Shippony, Z., and Jarnot, R. F.: The UARS Microwave Limb Sounder version 5 data set: Theory, characterization, and validation, Journal of Geophysical Research: Atmospheres, 108, https://doi.org/https://doi.org/10.1029/2002JD002273, 2003.

Livesey, N. J., Read, W. G., Wagner, P. A., Froidevaux, L., Santee, M. L., Schwartz, M. J., Lambert, A., Millan, L., Pumphrey, H. C., Manney, G. L., Fulle, r. R. A., Jarnot, R. F., Knosp, B. W., and Lay, R. R.: EOS MLS Version 5.0x Level 2 and 3 data quality and description document., Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, 2020.

Luo, B.: Stratospheric aerosol data for use in CMIP6 models, ftp://iacftp.ethz.ch/pub{_}read/luo/CMIP6/Readme{_}Data{_}Description.
505   pdf, 2016.

Mahieu, E., Chipperfield, M. P., Notholt, J., Reddmann, T., Anderson, J., Blumenstock, T., Coffey, M. T., Dhomse, S., Feng, W., and Franco,
B.: Increase in northern hemisphere stratospheric hydrogen chloride over recent years, in: Reklim Conference 2014: Our Climate-Our
Future, 2014.

McCormick, M., Zawodny, J., Veiga, R., Larsen, J., and Wang, P.: An overview of SAGE I and II ozone measurements, Planetary and Space
510   Science, 37, 1567–1586, 1989.

Millán, L. F., Livesey, N. J., Santee, M. L., Neu, J. L., Manney, G. L., and Fuller, R. A.: Case studies of the impact of orbital sampling
on stratospheric trend detection and derivation of tropical vertical velocities: solar occultation vs. limb emission sounding, Atmospheric
Chemistry and Physics, 16, 11 521–11 534, https://doi.org/10.5194/acp-16-11521-2016, 2016.

Mitchell, D. M., Lo, Y. E., Seviour, W. J., Haimberger, L., and Polvani, L. M.: The vertical profile of recent tropical temperature trends:
515   Persistent model biases in the context of internal variability, Environmental Research Letters, 15, 1040b4, 2020.

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., et al.: An
unexpected and persistent increase in global emissions of ozone-depleting CFC-11, Nature, 557, 413–417, 2018.

Montzka, S. A., Dutton, G. S., Portmann, R. W., Chipperfield, M. P., Davis, S., Feng, W., Manning, A. J., Ray, E., Rigby, M., Hall, B. D.,
et al.: A decline in global CFC-11 emissions during 2018- 2019, Nature, pp. 1–5, 2021.

520 Murtagh, D., Frisk, U., Merino, F., Ridal, M., Jonsson, A., Stegman, J., Witt, G., Eriksson, P., Jiménez, C., Megie, G., de La Noë, J., Ricaud,
P., Baron, P., Pardo, J. R., Hauchcorne, A., Llewellyn, E. J., Degenstein, D. A., Gattinger, R. L., Lloyd, N. D., Evans, W. F. J., McDade,
I. C., Haley, C. S., Sioris, C., von Savigny, C., Solheim, B. H., McConnell, J. C., Strong, K., Richardson, E. H., Leppelmeier, G. W.,
Kyrölä, E., Auvinen, H., and Oikarinen, L.: Review: An overview of the Odin atmospheric mission, Canadian Journal of Physics, 80,
309–319, https://doi.org/10.1139/p01-157, 2002.

525 Orbe, C., Wargan, K., Pawson, S., and Oman, L. D.: Mechanisms linked to recent ozone decreases in the Northern Hemisphere lower
stratosphere, Journal of Geophysical Research: Atmospheres, 125, e2019JD031 631, 2020.

Paul, J., Fortuin, F., and Kelder, H.: An ozone climatology based on ozonesonde and satellite measurements, Journal of Geophysical Research:
Atmospheres, 103, 31 709–31 734, 1998.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al.:
530   Scikit-learn: Machine learning in Python, the Journal of machine Learning research, 12, 2825–2830, 2011.

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, Tech. rep., SPARC, https://elib.dlr.de/126666/, 9th assessment report of the SPARC
project, published by the International Project Office at DLR-IPA. also: GAW Report No. 241; WCRP Report 17/2018, 2019.

Ploeger, F., Diallo, M., Charlesworth, E., Konopka, P., Legras, B., Laube, J. C., Grooß, J.-U., Günther, G., Engel, A., and Riese, M.: The
535   stratospheric Brewer–Dobson circulation inferred from age of air in the ERA5 reanalysis, Atmospheric Chemistry and Physics Discus-
sions, 2021, 1–27, https://doi.org/10.5194/acp-2020-1253, 2021.

Rahpoe, N., Weber, M., Rozanov, A. V., Weigel, K., Bovensmann, H., Burrows, J. P., Laeng, A., Stiller, G., von Clarmann, T., Kyrölä, E.,
Sofieva, V. F., Tamminen, J., Walker, K., Degenstein, D., Bourassa, A. E., Hargreaves, R., Bernath, P., Urban, J., and Murtagh, D. P.:
Relative drifts and biases between six ozone limb satellite measurements from the last decade, Atmospheric Measurement Techniques, 8,
540   4369–4381, https://doi.org/10.5194/amt-8-4369-2015, 2015.

Randel, W. J. and Wu, F.: A stratospheric ozone profile data set for 1979–2005: Variability, trends, and comparisons with column ozone data, Journal of Geophysical Research: Atmospheres, 112, 2007.

Rong, P., Russell III, J., Mlynczak, M., Remsberg, E., Marshall, B., Gordley, L., and Lopez-Puertas, M.: Validation of TIMED/SABER v1. 07 ozone at 9.6 μm in the altitude range 15–70km, J. Geophys. Res, 114, D04 306, 2008.

545 Russell III, J. M., Gordley, L. L., Park, J. H., Drayson, S. R., Hesketh, W. D., Cicerone, R. J., Tuck, A. F., Frederick, J. E., Harries, J. E., and Crutzen, P. J.: The halogen occultation experiment, Journal of Geophysical Research: Atmospheres, 98, 10 777–10 797, 1993.

Sofieva, V. F., Kalakoski, N., Päivärinta, S.-M., Tamminen, J., Laine, M., and Froidevaux, L.: On sampling uncertainty of satellite ozone profile measurements, Atmospheric Measurement Techniques, 7, 1891–1900, https://doi.org/10.5194/amt-7-1891-2014, 2014.

Sofieva, V. F., Kyrölä, E., Laine, M., Tamminen, J., Degenstein, D., Bourassa, A., Roth, C., Zawada, D., Weber, M., Rozanov, A., Rahpoe, N.,
550   Stiller, G., Laeng, A., von Clarmann, T., Walker, K. A., Sheese, P., Hubert, D., van Roozendael, 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, Atmospheric Chemistry and Physics, 17, 12 533–12 552, https://doi.org/10.5194/acp-17-12533-2017, 2017.

Sofieva, V. F., Szeląg, M., Tamminen, J., Kyrölä, E., Degenstein, D., Roth, C., Zawada, D., Rozanov, A., Arosio, C., Burrows, J. P., Weber, M., Laeng, A., Stiller, G. P., von Clarmann, T., Froidevaux, L., Livesey, N., van Roozendael, M., and Retscher, C.: Measurement report:
555   regional trends of stratospheric ozone evaluated using the MErged GRIdded Dataset of Ozone Profiles (MEGRIDOP), Atmospheric Chemistry and Physics, 21, 6707–6720, https://doi.org/10.5194/acp-21-6707-2021, 2021.

Solomon, S., Ivy, D. J., Kinnison, D., Mills, M. J., Neely, R. R., and Schmidt, A.: Emergence of healing in the Antarctic ozone layer, Science, http://science.sciencemag.org/content/early/2016/06/30/science.aae0061, 2016.

Steinbrecht, W., Froidevaux, L., Fuller, R., Wang, R., Anderson, J., Roth, C., Bourassa, A., Degenstein, D., Damadeo, R., Zawodny, J.,
560   Frith, S., McPeters, R., Bhartia, P., Wild, J., Long, C., Davis, S., Rosenlof, K., Sofieva, V., Walker, K., Rahpoe, 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, Atmospheric Chemistry and Physics, 17, 10 675–10 690, https://doi.org/10.5194/acp-
565   17-10675-2017, 2017.

Stolarski, R. S., Douglass, A. R., Newman, P. A., Pawson, S., and Schoeberl, M. R.: Relative Contribution of Greenhouse Gases and Ozone-Depleting Substances to Temperature Trends in the Stratosphere: A Chemistry–Climate Model Study, Journal of Climate, 23, 28–42, https://doi.org/10.1175/2009JCLI2955.1, 2010.

Strahan, S. E., Douglass, A. R., Stolarski, R. S., Akiyoshi, H., Bekki, S., Braesicke, P., Butchart, N., Chipperfield, M. P., Cugnet, D., Dhomse,
570   S., Frith, S. M., Gettelman, A., Hardiman, S. C., Kinnison, D. E., Lamarque, J. F., Mancini, E., Marchand, M., Michou, M., Morgenstern, O., Nakamura, T., Olivie, D., Pawson, S., Pitari, G., Plummer, D. A., Pyle, J. A., Scinocca, J. F., Shepherd, T. G., Shibata, K., Smale, D., Teyssedre, H., Tian, W., and Yamashita, Y.: Using transport diagnostics to understand chemistry climate model ozone simulations, Journal of Geophysical ResearchAtmospheres, 116, https://doi.org/10.1029/2010jd015360, 2011.

Sukhodolov, T., Rozanov, E., Ball, W., Bais, A., Tourpali, K., Shapiro, A., Telford, P., Smyshlyaev, S., Fomin, B., Sander, R., Bossay, S.,
575   Bekki, S., Marchand, M., Chipperfield, M., Dhomse, S., Haigh, J., Peter, T., and Schmutz, W.: Evaluation of simulated photolysis rates and their response to solar irradiance variability, Journal of Geophysical Research: Atmospheres, 121, https://doi.org/10.1002/2015JD024277, 2016.

Svetnik, V., Liaw, A., Tong, C., Culberson, J. C., Sheridan, R. P., and Feuston, B. P.: Random forest: a classification and regression tool for compound classification and QSAR modeling, Journal of chemical information and computer sciences, 43, 1947–1958, 2003.

580 Szeląg, M. E., Sofieva, V. F., Degenstein, D., Roth, C., Davis, S., and Froidevaux, L.: Seasonal stratospheric ozone trends over 2000–2018 derived from several merged data sets, Atmospheric Chemistry and Physics, 20, 7035–7047, 2020.

Tian, W., Chipperfield, M. P., Gray, L. J., and Zawodny, J. M.: Quasi-biennial oscillation and tracer distributions in a coupled chemistry-climate model, Journal of Geophysical Research: Atmospheres, 111, https://doi.org/https://doi.org/10.1029/2005JD006871, 2006.

Wales, P. A., Salawitch, R. J., Nicely, J. M., Anderson, D. C., Canty, T. P., Baider, S., Dix, B., Koenig, T. K., Volkamer, R., Chen, D.,
585 Huey, L. G., Tanner, D. J., Cuevas, C. A., Fernandez, R. P., Kinnison, D. E., Lamarque, J.-F., Saiz-Lopez, A., Atlas, E. L., Hall, S. R., Navarro, M. A., Pan, L. L., Schauffler, S. M., Stell, M., Tilmes, S., Ullmann, K., Weinheimer, A. J., Akiyoshi, H., Chipperfield, M. P., Deushi, M., Dhomse, S. S., Feng, W., Graf, P., Hossaini, R., Jöckel, P., Mancini, E., Michou, M., Morgenstern, O., Oman, L. D., Pitari, G., Plummer, D. A., Revell, L. E., Rozanov, E., Saint-Martin, D., Schofield, R., Stenke, A., Stone, K., Visioni, D., Yamashita, Y., and Zeng, G.: Stratospheric Injection of Brominated Very Short-Lived Substances Inferred from Aircraft Observations of Organic Bromine and BrO
590 in the Western Pacific, Journal of Geophysical Research: Atmospheres, 0, https://doi.org/10.1029/2017JD027978, 2018.

Wargan, K., Orbe, C., Pawson, S., Ziemke, J. R., Oman, L. D., Olsen, M. A., Coy, L., and Emma Knowland, K.: Recent decline in extratropical lower stratospheric ozone attributed to circulation changes, Geophysical research letters, 45, 5166–5176, 2018.

Wargan, K., Weir, B., Manney, G. L., Cohn, S. E., and Livesey, N. J.: The anomalous 2019 Antarctic ozone hole in the GEOS Constituent Data Assimilation System with MLS observations, Journal of Geophysical Research: Atmospheres, 125, e2020JD033 335, 2020.

595 Weber, M., Dhomse, S., Wittrock, F., Richter, A., Sinnhuber, B.-M., and Burrows, J. P.: Dynamical control of NH and SH winter/spring total ozone from GOME observations in 1995–2002, Geophysical Research Letters, 30, 1583, https://doi.org/10.1029/2002GL016799, 2003.

Weber, M., Dikty, S., Burrows, J. P., Garny, H., Dameris, M., Kubin, A., Abalichin, J., and Langematz, U.: The Brewer-Dobson circulation and total ozone from seasonal to decadal time scales, Atmospheric Chemistry and Physics, 11, 11 221–11 235, https://doi.org/10.5194/acp-11-11221-2011, 2011.

600 Weber, M., Arosio, C., Feng, W., Dhomse, S. S., Chipperfield, M. P., Meier, A., Burrows, J. P., Eichmann, K.-U., Richter, A., and Rozanov, A.: The Unusual Stratospheric Arctic Winter 2019/20: Chemical Ozone Loss From Satellite Observations and TOMCAT Chemical Transport Model, Journal of Geophysical Research: Atmospheres, 126, e2020JD034 386, https://doi.org/https://doi.org/10.1029/2020JD034386, e2020JD034386 2020JD034386, 2021.

WMO: Scientific Assessment of Ozone Depletion:2014, Tech. rep., World Meteorological Organization, Global Ozone Research and Moni-
605 toring Project, Report No. 55, Geneva, Switzerland, https://www.wmo.int/pages/prog/arep/gaw/ozone{_}2014/, 2014.

WMO: Scientific Assessment of Ozone Depletion:2018, Tech. rep., World Meteorological Organization, Global Ozone Research and Monitoring Project, Report No. 58, Geneva, Switzerland, https://www.wmo.int/pages/prog/arep/gaw/ozone{_}2018/, 2018.

Wohltmann, I., Gathen, P., Lehmann, R., Maturilli, M., Deckelmann, H., Manney, G. L., Davies, J., Tarasick, D., Jepsen, N., Kivi, R., Lyall, N., and Rex, M.: Near-complete local reduction of Arctic stratospheric ozone by severe chemical loss in spring 2020, Geophysical
610 Research Letters, 47, https://doi.org/10.1029/2020GL089547, 2020.

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

Author contributions. SD conceived the idea and initiated the study in discussion with MPC. The model runs were performed and analysed
615   by SD, MPC and WF. The figures were prepared by CA and SD. The paper was written by SD, MPC and MW who included comments from
all of the other coauthors.

Competing interests. The authors have no competing interests
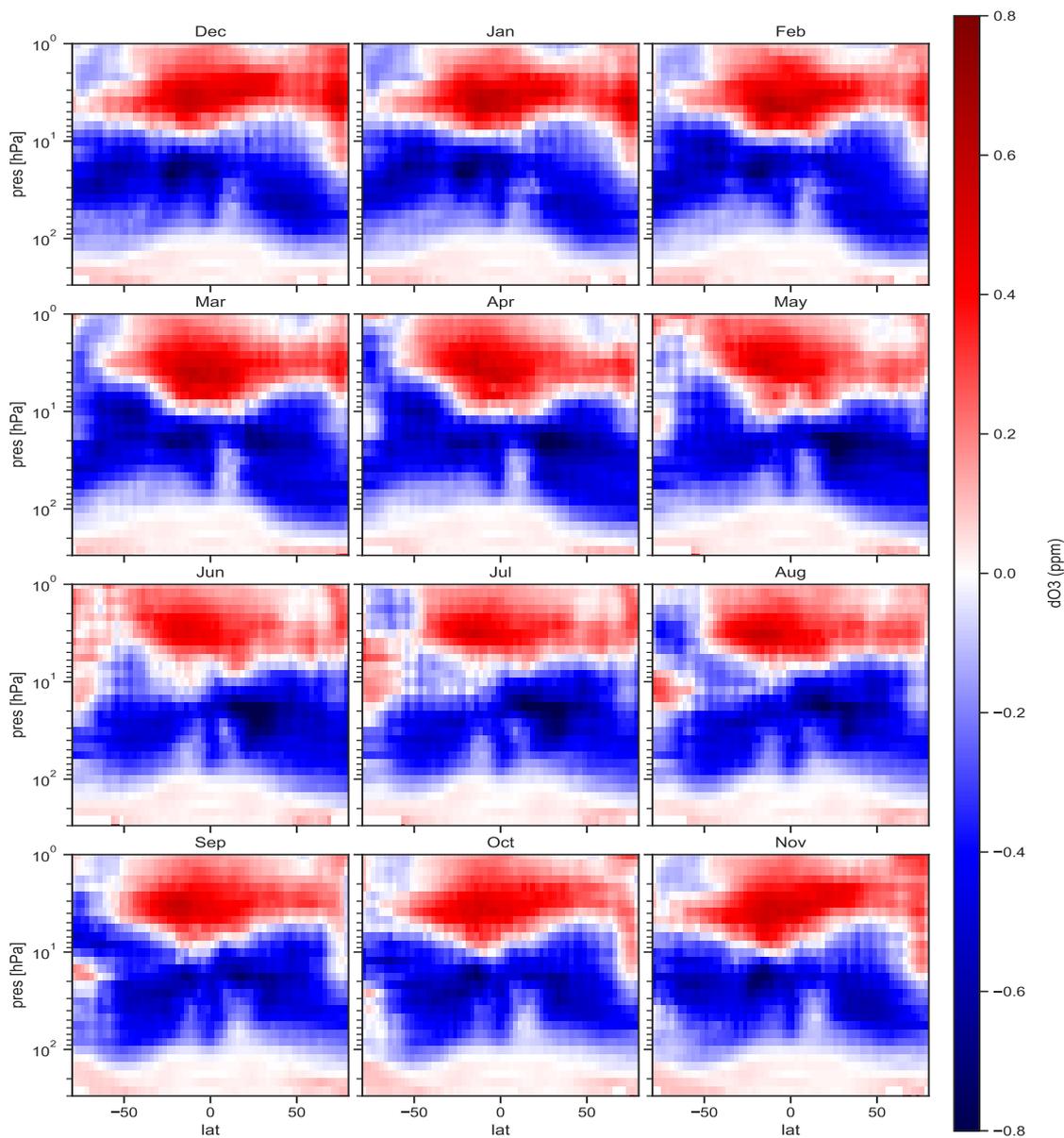
Figure 1. Latitude-pressure cross sections of the climatological (2001-2020) monthly mean difference (ppm) between SWOOSH (Davis et al., 2016) and TOMCAT ozone profiles.

o
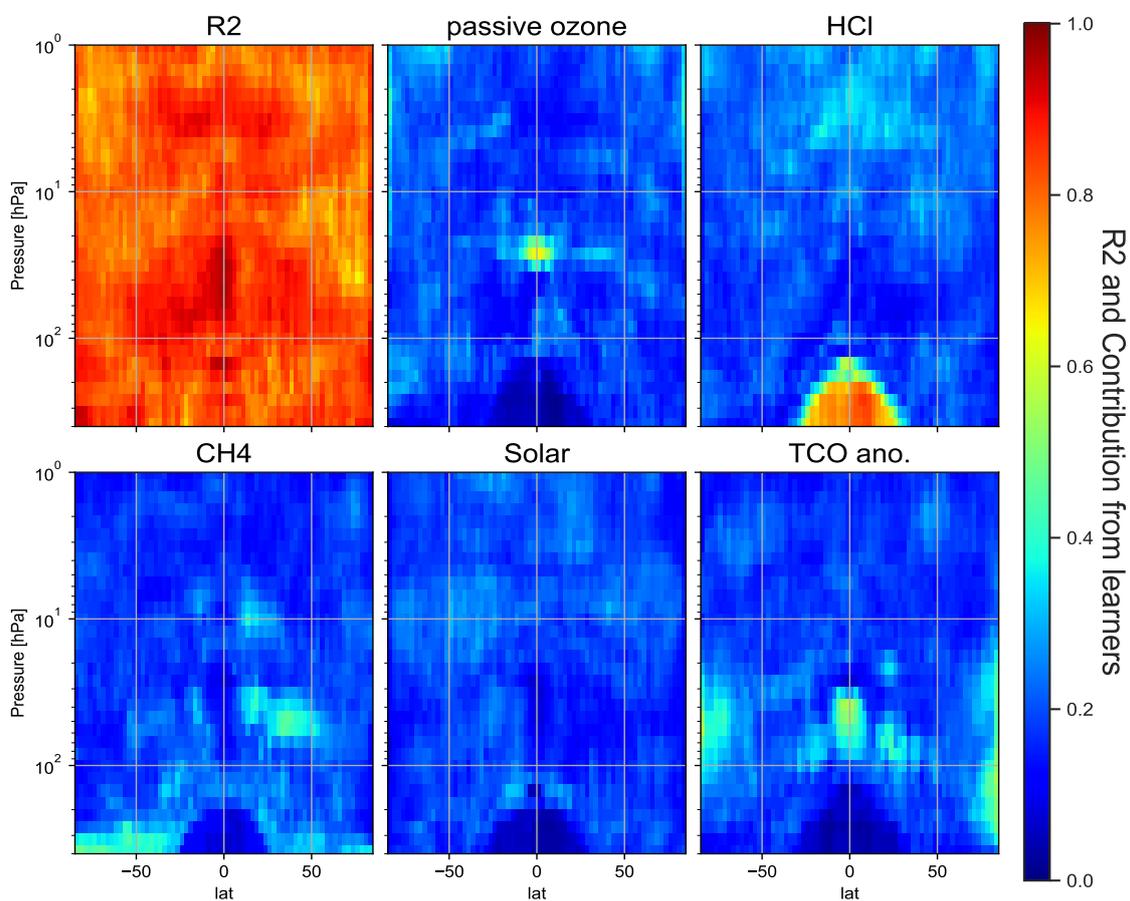
Figure 2. Latitude-pressure cross sections of the variance (R^2) and regression coefficients from passive ozone, HCl, CH$_4$, solar and total column ozone anomaly (see main text equation 1).
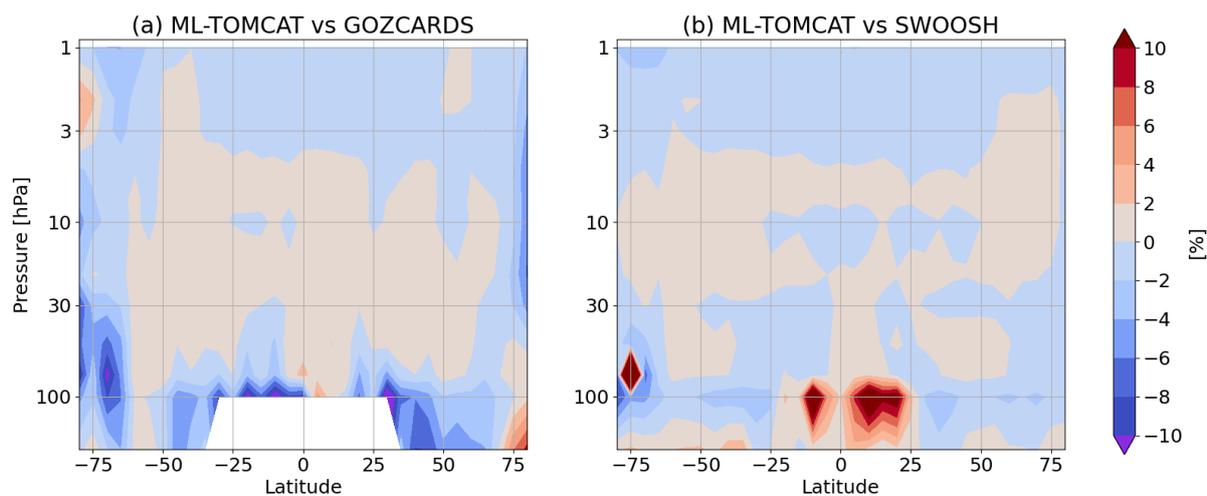
Figure 3. Relative differences (in percent) as a function of pressure and latitude between ML-TOMCAT and (a) GOZCARDS V2 (Froidevaux et al., 2019) and (b) SWOOSH (Davis et al., 2016).
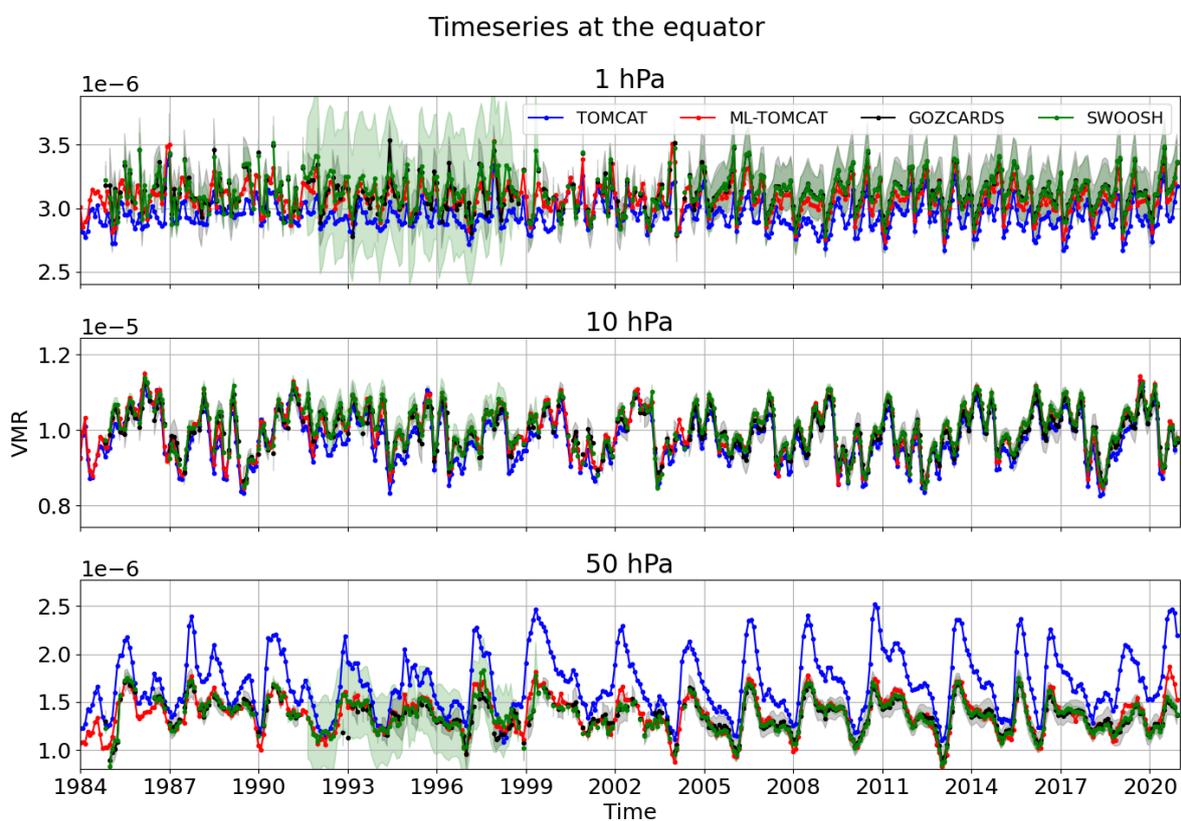
Figure 4. Comparison between TOMCAT (blue lines) and ML-TOMCAT (red lines) ozone mixing ratios over the equator (0°) at (a, top) 1 hPa, (b, middle) 10 hPa and (c, bottom) 50 hPa. Satellite-based ozone mixing ratios from GOZCARDS (Froidevaux et al., 2019) and SWOOSH (Davis et al., 2016) datasets along with their uncertainty estimates (shaded) are shown with black and green coloured lines, respectively.
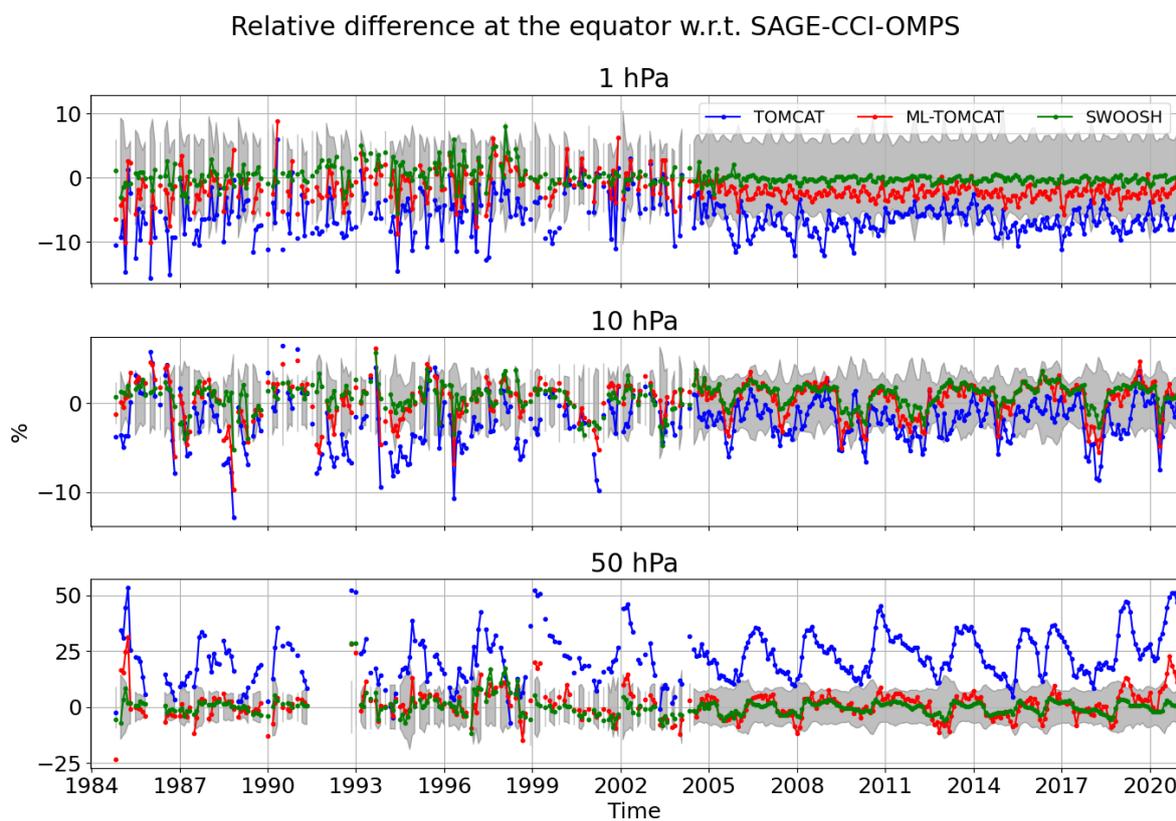
Figure 5. As Figure 4 but for the residuals, i.e. relative differences between SWOOSH (green), TOMCAT (blue) and ML-TOMCAT (red) ozone with respect to GOZCARDS ozone.
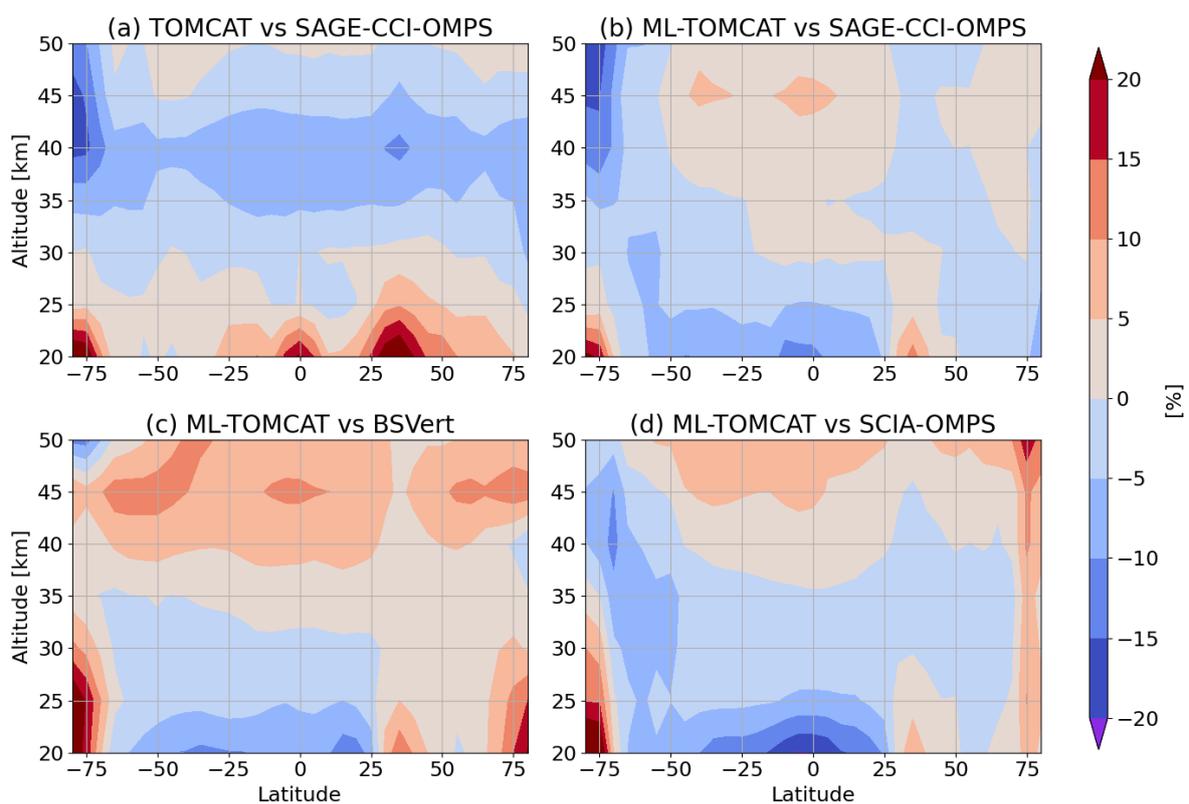
Figure 6. Relative difference (%) as a function of latitude and altitude between TOMCAT, ML-TOMCAT and the three considered datasets: SAGE-CCI-OMPS (1985-2019), BSVert (1985-2017) and SCIA-OMPS (2002-2019), averaged over the respective time series.
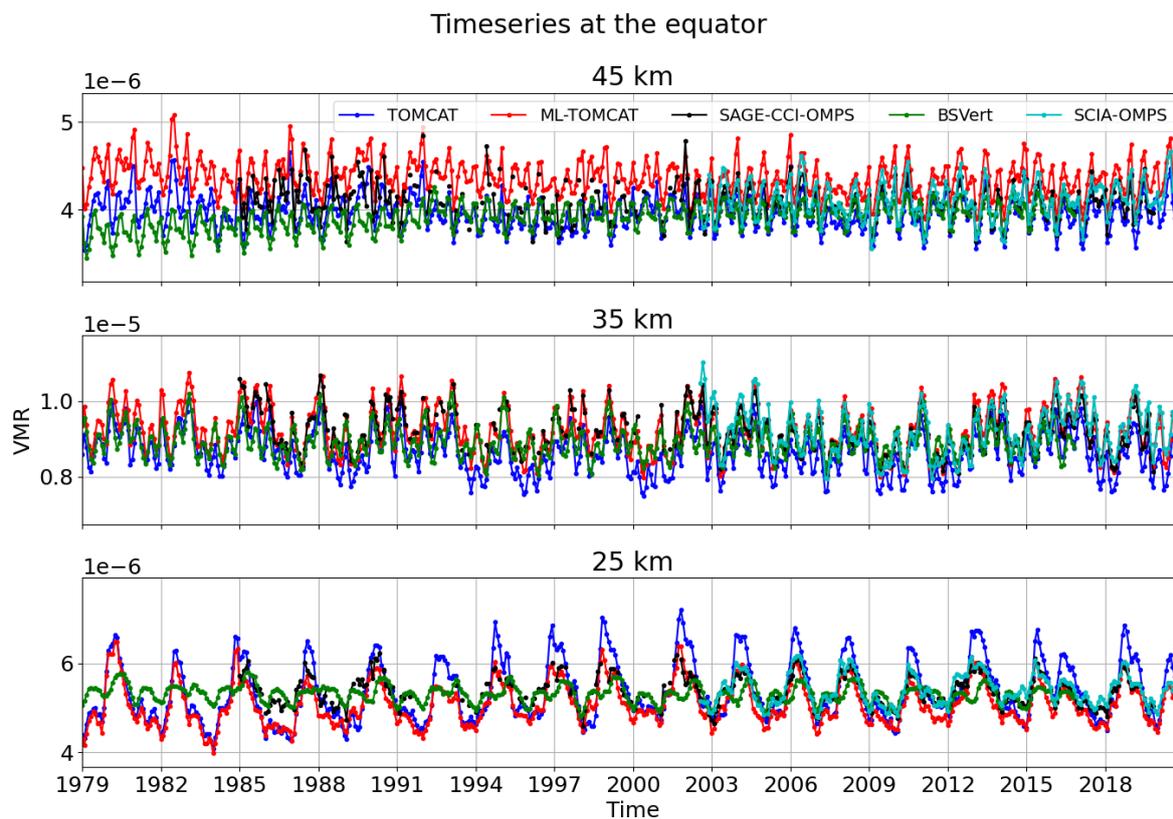
Figure 7. Comparison between TOMCAT (blue lines) and ML-TOMCAT (red lines) ozone mixing ratios over the equator (0°) at (a, top) 45 km, (b, middle) 35 km and (c, bottom) 25 km. Satellite-based ozone mixing ratios from SAGE–CCI–OMPS, BSVert (Hassler et al., 2018) and SCIA-OMPS (Arosio et al., 2019) datasets are shown with black, green and cyan coloured lines, respectively.
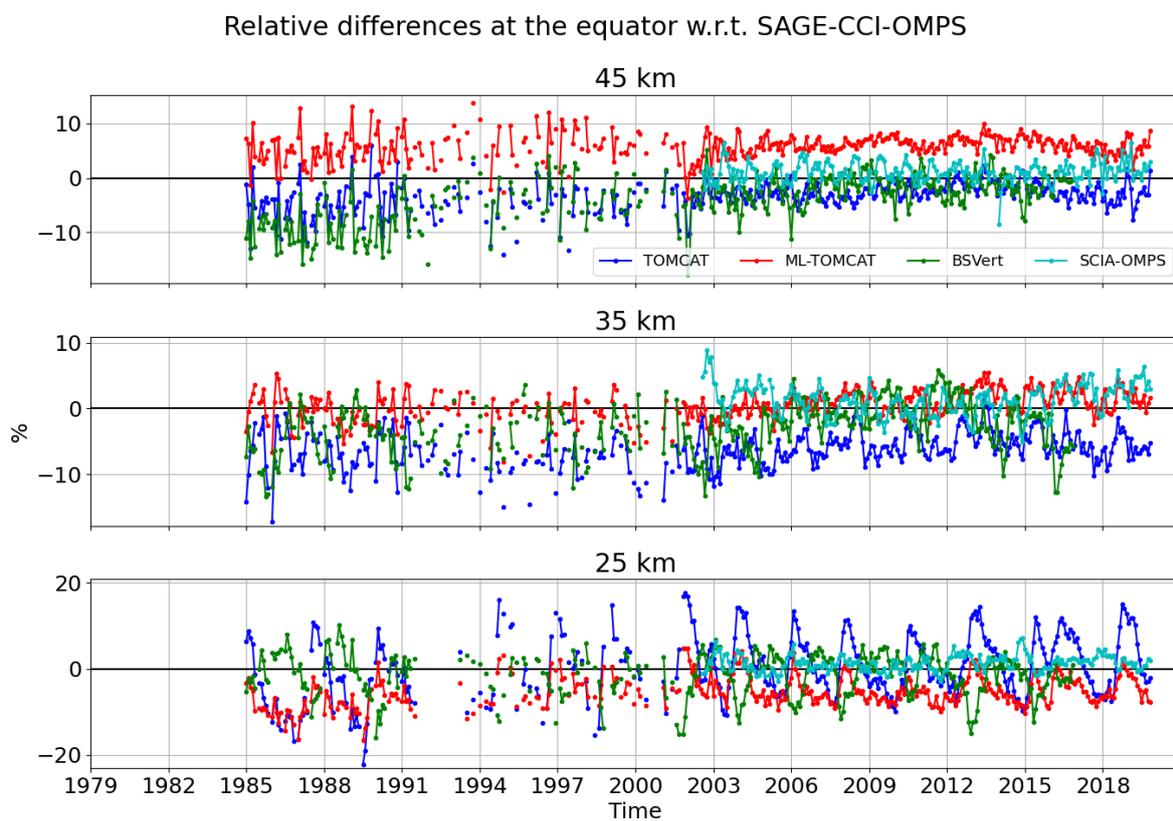
Figure 8. Same as Figure 7 but for the residuals, i.e. relative differences between TOMCAT (blue), ML-TOMCAT (red), BSVert (green) and SCIA-OMPS (cyan) ozone with respect to SAGE–CCI–OMPS.
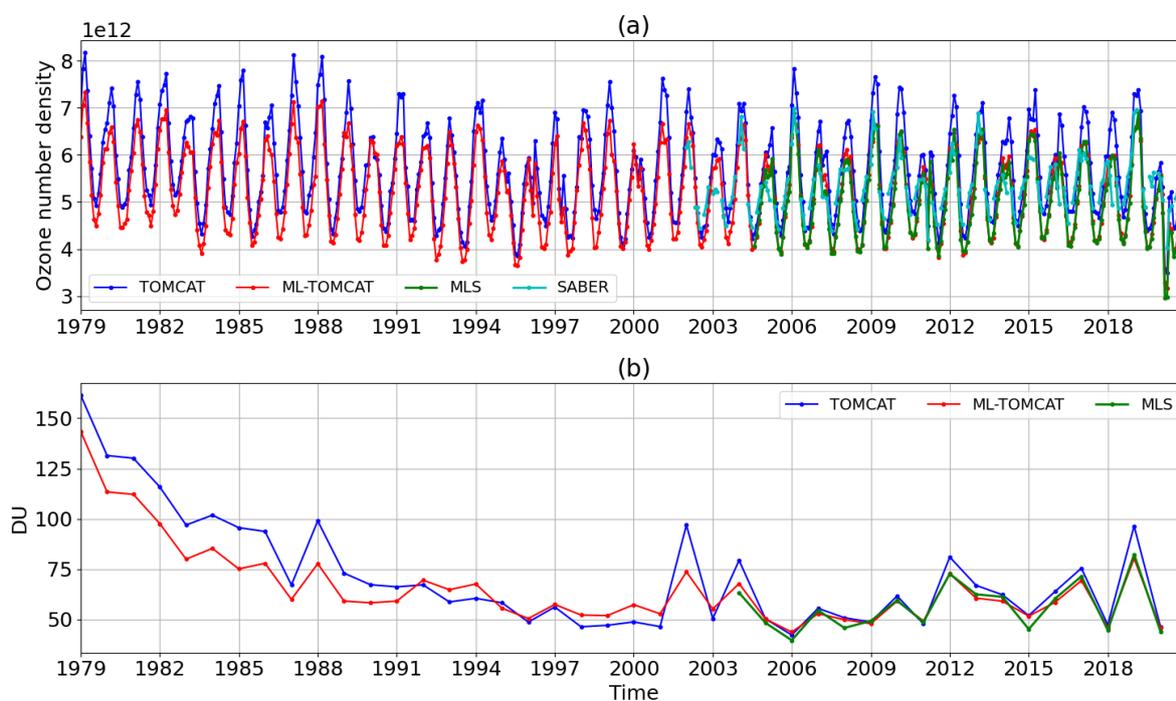
Figure 9. (a) Ozone concentration time series (molecules cm^{-3}) at 18 km over the North Polar region (latitudes poleward of 70°N). Aura-MLS and the Sounding of the Atmosphere using Broadband Emission Radiometry (SABER, Rong et al., 2008) data are superimposed on TOMCAT and ML-TOMCAT timeseries. (b) Mean ozone sub-column (DU) between 12-20 km for September and October each year over the South Polar region (latitudes poleward of 70°S.