Evaluation of the Ozone Fields in NASA's MERRA-2 Reanalysis

Krzysztof Wargan[1,2], **Gordon Labow**[2,3], **Stacey Frith**[2,3], **Steven Pawson**[1], **Nathaniel Livesey**[4], and **Gary Partyka**[1,2]

[1]Global Modeling and Assimilation Office, Code 610.1, NASA Goddard Space Flight Center, Greenbelt, MD

[2]Science Systems and Applications Inc., Lanham, MD

[3]Atmospheric Chemistry and Dynamics Laboratory, Code 614, NASA Goddard Space Flight Center, Greenbelt, MD

[4]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA

Abstract

We describe and assess the quality of the assimilated ozone product from the Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2) produced at NASA's Global Modeling and Assimilation Office (GMAO) spanning the time period from 1980 to present. MERRA-2 assimilates partial column ozone retrievals from a series of Solar Backscatter Ultraviolet (SBUV) radiometers on NASA and NOAA spacecraft between January 1980 and September 2004; starting in October 2004 retrieved ozone profiles from the Microwave Limb Sounder (MLS) and total column ozone from the Ozone Monitoring Instrument on NASA's EOS Aura satellite are assimilated. We compare the MERRA-2 ozone with independent satellite and ozonesonde data focusing on the representation of the spatial and temporal variability of stratospheric and upper tropospheric ozone and on implications of the change in the observing system from SBUV to EOS Aura. The comparisons show agreement within 10 % (standard deviation of the difference) between MERRA-2 profiles and independent satellite data in most of the stratosphere. The agreement improves after 2004 when EOS Aura data are assimilated. The standard deviation of the differences between the lower stratospheric and upper tropospheric MERRA-2 ozone and ozonesondes is 11.2 % and 24.5 %, respectively, with correlations of 0.8 and above, indicative of a realistic representation of the near-tropopause ozone variability in MERRA-2. The agreement improves significantly in the EOS Aura period, however MERRA-2 is biased low in the upper troposphere with respect to the ozonesondes. Caution is recommended when using MERRA-2 ozone for decadal changes and trend studies.

1. Introduction

Atmospheric reanalyses produce global high spatial and temporal resolution long-term records of meteorological fields and composition of earth's atmosphere by utilizing the data assimilation methodology (Cohn 1997; Kalnay 2003), whereby satellite and ground-based observations are combined with general circulation model (GCM) simulations in a

Correspondence: Krzysztof Wargan, Krzysztof.Wargan-1@nasa.gov.

statistically optimal way. The Modern Era Retrospective Analysis for Research and Applications (MERRA: Rienecker et al. 2011) was the first reanalysis generated using the Goddard Earth Observing System (GEOS) Data Assimilation System (DAS) by NASA's Global Modeling and Assimilation Office (GMAO). MERRA, first released in 2009, covered the years 1979–2015 (production ended on 29 February 2016). It was followed by the recently released MERRA-2 dataset (Bosilovich et al. 2015), which is the focus of this paper.

While most reanalyses include assimilated ozone fields, the overall lack of validation and uncertain quality of these fields has not encouraged the atmospheric ozone community to use them in scientific research. Typically, researchers prefer to utilize satellite and in-situ ozone data along with assimilated meteorological variables. To our knowledge, the only comprehensively validated reanalysis ozone fields are those from the European Centre for Medium-Range Weather Forecasts reanalyses: ERA-40 (Dethof and Hólm 2004) and ERA-Interim (Dragani 2011). On the other hand, a large body of literature evaluates ozone data assimilation results from (usually relatively short) assimilation experiments (see Lahoz et al. 2014). In addition, significant effort has been invested in the evaluation of multiyear chemical analyses under the Monitoring Atmospheric Composition and Climate (MACC) project conducted using the ECMWF Integrated Forecast system coupled with comprehensive chemistry models (e.g., Inness et al., 2013; Inness et al.; 2015l; Lefever et al. 2015). Much of this work has demonstrated an added value brought to satellite observations of ozone through data assimilation. In particular, work done at the GMAO over the past decade has shown that assimilation of retrieved ozone data from the Microwave Limb Sounder (MLS) along with total ozone observations from the Ozone Monitoring Instrument (OMI), both onboard the Earth Observing System (EOS) Aura satellite, produces realistic global distributions of ozone in the stratosphere and upper troposphere (Stajner et al. 2008; Wargan et al. 2015). Ziemke et al. (2014) compared tropospheric ozone derived from the GEOS DAS with that from trajectory mapping of the Earth Observing System (EOS) Aura data and direct profile retrieval from OMI radiances and concluded that data assimilation is the best of these three strategies at generating global ozone product. By combining available measurements with global circulation model short-term forecasts the data assimilation methodology allows the propagation of observational information by assimilated winds resulting in global 3-dimensional maps of ozone concentrations at spatial and temporal resolutions far exceeding those attainable with satellite data alone. For example, a satellite-borne instrument at a sun-synchronous orbit typically samples about 30 points along a latitude circle per day (15 if day-only observations are made), corresponding to a 12° resolution in the longitudinal direction. By contrast, the resolution of MERRA-2 is approximately 0.625° in longitude and has a three-hourly output frequency.

This paper presents a description and evaluation of the MERRA-2 ozone product against independent satellite and ozonesonde measurements. It is intended to provide guidance for researchers who may wish to use the MERRA-2 ozone product in scientific studies. We focus on the following questions:

1.   How well does MERRA-2 represent the spatio-temporal variability of the stratospheric and upper-tropospheric ozone?

> **2.**   What are the impacts of changing the ozone observing system in MERRA-2 in late 2004 when new types of data were introduced?

The first question addresses the ability of the assimilated ozone to represent the statistical characteristics of the actual ozone field, which must be established before the assimilated fields can be used in research studies. In particular we focus on results in the lower stratospheric/upper tropospheric region, where variability is enhanced and high spatio-temporal resolution satellite observations are limited. Analysis of, for example, radiative forcing near the tropopause, or stratosphere-troposphere exchange mechanisms require both consistency between the constituent and dynamical fields, and accurate representation of small-scale variability. The stability of the multi-decadal MERRA-2 ozone record and the evolution of the product's quality resulting from changes in the observing system are addressed in the second question.

While this study is not intended to be a comprehensive validation of the original MERRA reanalysis we include both MERRA and MERRA-2 ozone in a number of comparisons in order to emphasize the differences that arise (mainly) from differences in the input ozone observations between the two products.

The paper is organized as follows. Section 2 describes the MERRA-2 reanalysis, focusing on the treatment of ozone. The MERRA-2 ozone observing system is discussed in Section 3. A description of independent data used for validation is given in Section 4. The differences between MERRA and MERRA-2 relevant to ozone are summarized in Section 5. Section 6 presents the results of comparisons of the MERRA-2 ozone fields against satellite and ozonesonde observations. Section 7 summarizes the results of this study.

2. The MERRA-2 reanalysis

MERRA-2 (Bosilovich et al. 2015) is a multi-year reanalysis developed at GMAO, covering the 'satellite era' of earth observations from 1980 to present. It is produced using the Version 5.12.4 of the GEOS DAS. Gridded data are released at a 0.625° by 0.5° longitude by latitude resolution at 72 sigma-pressure hybrid layers between the surface and 0.01 hPa. The bottom 32 layers are terrain-following while remaining model layers, from 164 to .01 hPa, are constant pressure surfaces. MERRA-2 replaces the original GMAO reanalysis, MERRA (Rienecker et al. 2011). Both the general circulation model and the observing system were significantly updated since MERRA. The model updates include a transition from a regular spherical grid to the cubed sphere (Putman and Lin 2007). A re-tuning of the gravity wave parameterization affects the stratosphere, in particular resulting in a realistic quasi-biennial oscillation (QBO) generated by the model in the absence of observations (Coy et al. 2016). Other changes include upgrades to moist physics and the enforcement of dry mass conservation in the assimilation (Takacs et al. 2016). The model updates are described in detail in Molod et al. (2015). The observing system was expanded to include more recent satellite data: the Infrared Atmospheric Sounding Interferometer (IASI, starting in September 2008), the Cross-Track Infrared Sounder (on the Suomi-NPP satellite, from April 2012 onward) and Advanced Technology Microwave Sounder (on Suomi-NPP, starting in November 2011) in addition to those already used in MERRA. Radiance data from the

Stratospheric Sounding Unit instruments are used in both reanalyses but in MERRA-2 these observations are assimilated with the more advanced Community Radiative Transfer Model, the same as for all other radiance data. Both reanalyses assimilate radiance observations from the Advanced Microwave Sounding Unit and Atmospheric Infrared Sounder. A comprehensive description of the MERRA-2 observing system is given in McCarty et al. (2016). An adaptive bias correction scheme is applied to all radiance data and some aircraft temperature observations (McCarty et al. 2016). This study uses the three-hourly MERRA-2 assimilated ozone output on the native 72 levels (GMAO 2015a) and the one-hourly total ozone product (GMAO 2015b).

3. Assimilated ozone data

The ozone data sources used in MERRA-2 are listed in Table 1. Between January 1980 and September 2004 partial column and total ozone from a series of Solar Backscatter Ultraviolet (SBUV) instruments on NASA and NOAA spacecraft are assimilated. Starting 1 October 2004 these observations are turned off and replaced by total column ozone data from OMI and stratospheric ozone profiles from MLS, both on NASA's EOS Aura satellite. The MERRA-2 ozone record is thus divided into two periods, henceforth referred to as the SBUV period (January 1980- September 2004) and the Aura period (from October 2004 onward). Joint assimilation of MLS and SBUV data was considered but tests showed that it would generate unphysical features in the assimilated ozone fields in the tropical and subtropical stratosphere. Since the GEOS DAS currently ingests SBUV partial columns without accounting for averaging kernels, an apparent bias between MLS and SBUV arises leading to large analysis increments that produce satellite track-following features in the ozone fields.

The following subsections describe the ozone data sources in detail.

3.1 SBUV

SBUV is a nadir-viewing instrument that measures incoming solar irradiance and backscattered solar radiance reflected from the Earth's atmosphere directly beneath the satellite. Measurements are made in 12 narrow wavelength bands between 250 and 340 nm using a double monochromator with a 1.1 nm (full-width half maximum) triangular slit function (Fleig et al., 1990). The ratio of radiance to irradiance at various wavelengths is used to infer the amount of ozone in broad layers of the atmosphere in the column below the instrument. The total column ozone is the sum of the profile layer values. Satellite measurements of stratospheric ozone from the SBUV instrument began in November 1978 with the launch of the Nimbus-7 spacecraft. A series of SBUV/2s began operations with the launch of NOAA-9 in January 1985 and continued with the sequential launches of NOAA-11, -14, -16, -17, -18, and -19, with NOAA-19 still operational at the time of publication. The MERRA-2 processing ingests data from Nimbus-7 and NOAAs 11–17, following the time line in Table 1, before switching to data from EOS Aura. We include measurements taken at solar zenith angles less than 84 degrees. Measurements are available nearly every day, with the exception of one month of missing data in March 1991 from NOAA-11 SBUV/2. Hereafter we refer to SBUV and SBUV/2 simply as 'SBUV'.

SBUV measurements are made every 32 seconds, giving approximately 100 measurements per orbit (about 1 every 1.85 degrees in latitude; each orbit is separated by ~ 26 degrees longitude), or over 1400 measurements on a given day. The instrument field of view traces a 188 km wide swath (200 km for Nimbus-7 SBUV) along the orbital track. All the data have been processed with the Version 8.6 retrieval algorithm (V8.6; Bhartia et al. 2013), an update to the Version 8 data assimilated in MERRA. In the V8.6 algorithm, the ozone cross-sections are taken from (Daumont et al. 1992), which are superior in resolution, temperature dependence, and quality to the Bass and Paur (1985) cross-sections used in prior retrievals. A more accurate cloud height climatology has been developed using the UV rotational Raman filling technique (Vasilkov et al. 2004) from the OMI onboard EOS-Aura. The data set contains the climatological heights of the "optical centroid pressure," which reflect how deep a UV photon, on average, will penetrate into the cloud. With this more accurate cloud height climatology, the errors produced by extrapolating ozone amounts under a cloud are minimized.

Instrument noise for the majority of measurements is less than 0.5% (Deland et al., 2012). Bhartia et al. (2012) describes the V8.6 algorithm and related uncertainties in detail. The largest source of error in the SBUV profile retrievals is the smoothing error (Bhartia et al, 2012; Kramarova et al., 2012). The smoothing error describes the component of vertical ozone variability which the observation system cannot measure. Between 16 and 1 hPa, the smoothing errors for SBUV monthly zonal mean retrievals are of the order of 1%, increasing to 15–20% in the troposphere. The smoothing errors for total ozone retrievals are mostly less than 0.5%. In MERRA-2 the layer-dependent SBUV partial column error specification arises from tuning performed at the National Centers for Environmental Prediction and follows other GMAO operational analyses that use SBUV: the errors are assumed to be 1.4 Dobson units (DU, ~15%) between 1000 hPa and 631 hPa, 1.9 DU (11 % – 6 %) in the lower stratospheric layers, 100 hPa – 63 hPa and 63 hPa – 40 hPa and decreasing to 1 DU (25 % to over 100 %) in the upper stratospheric and mesospheric layers. Note that with these error specifications SBUV data have little impact in the upper stratosphere and mesosphere. The total ozone error is set to 6 DU (about 2 % of the global total ozone) in agreement with ground-based and ozonesonde comparisons done by Labow et al. (2013).

The SBUV series of instruments were calibrated using a combination of "hard calibration" from pre-launch and on-orbit instrument monitoring, and a variety of "soft calibration" techniques as described in Deland et al. (2012). In addition, the instruments are cross-calibrated to each other at the radiance level within the V8.6 retrieval algorithm. Nevertheless differences between the instruments exist and the data quality varies from one instrument to the next. In particular, measurements from the NOAA-14 SBUV/2 instrument show enhanced uncertainty relative to the other instruments in validation studies (Kramarova et al. 2013b).

3.2 OMI

The Ozone Monitoring Instrument (Levelt et al. 2006) operates on NASA's EOS Aura satellite, which was launched on 15 July 2004 into a sun-synchronous orbit with a 1:45 PM equatorial crossing time on the ascending node. The instrument is equipped with a 60-pixel

cross-track sensor array measuring backscattered solar radiation in the 270–550 nm wavelength range with a spectral resolution of ~0.5 nm at nadir. The treatment of OMI observations in MERRA-2 closely follows that of Wargan et al. (2015) and we will only summarize it briefly here. The reanalysis assimilates total column ozone data using the version-8.5 retrieval algorithm extensively evaluated by McPeters et al. (2008). Of the 60 available pixels (rows), only ozone columns from rows 3–24 are used. This is motivated by the lower quality of the data from rows affected by a mechanical issue known as the row anomaly from 2008 onward as well as large pixel sizes from rows 1 and 2. With this selection the swath width is about 1,100 km. The total column ozone from OMI is assimilated using efficiency factors provided with the data in order to account for the lower sensitivity of OMI measurements in the lower stratosphere, specifically in clouded scenes. This methodology is described in detail in Wargan et al. (2015).

The following changes to the OMI treatment were applied in MERRA-2 relative to the approach of Wargan et al. (2015) following recommendations of the OMI science team (P.K. Bhartia, personal communication): (1) Only the pixels with the total ozone quality flag set to 0 are used and (2) The efficiency factors are set to 1.0 above 125 hPa. MERRA-2 assumes a constant observation error of 5 DU everywhere. This number is within 2 % of the globally averaged total ozone and it is consistent with the level of agreement between OMI and ground-based measurements reported by McPeters et al. (2008). McPeters et al. (2015) note that OMI exhibits a latitude-dependent low bias with respect to SBUV (from 0.5 % between 60°S and 15°S to 2 % near 60°N). At present, the source of this bias is not understood and no attempt was made in MERRA-2 to remove it. It leads to an increase of the difference between SBUV data and MERRA-2 total ozone from 1 % to 2 % between the tropics and northern high latitudes, as discussed below in Section 6.1.

3.3 Aura MLS

The Microwave Limb Sounder on the EOS Aura satellite (Waters et al. 2006) measures profiles of atmospheric thermal radiation in a broad spectrum of microwave bands allowing high quality retrievals of temperature and concentrations of over a dozen chemical species in the stratosphere, including ozone. It provides about 3,500 retrieved stratospheric ozone profiles daily covering the 82°S-82°N latitude range in both day and night. The vertical resolution, determined from the full width at half maximum of the averaging kernels of the MLS ozone data, ranges from 2.5 km in the middle stratosphere to 6 km in the mesosphere (Froidevaux et al. 2008). Note that this is not the same as the spacing of the nominal levels on which the data are provided (see below). MERRA-2 uses version 2.2 ozone retrievals (Froidevaux et al. 2008) between October 2004 and May 2015 and version 4.2 (Livesey et al. 2015) afterwards. We note that MERRA-2 was already in production when version 4.2 data were released so the use of an older version was necessary in the preceding period. The decision to use version 2.2 rather than 3.3 as in Wargan et al. (2015) was motivated by the occurrence of unphysical vertical oscillations in the tropical profiles seen in version 3.3 that were present but to a much lesser degree in version 2.2. In both versions ozone is retrieved from the 230–250 GHz spectral region but the nominal vertical resolutions and ranges differ between the different versions. The recommended vertical range for version 2.2 is 215 hPa – 0.02 hPa. However, following Wargan et al. (2015) we applied mid-layer averaging to the

profiles prior to assimilation, consistent with the fact that the GEOS-5 ozone represents layer-averaged concentrations so that the lowest assimilated layer is centered on 177.8 hPa and the highest at 0.05 hPa, 21 levels total. In version 4.2 the number of MLS levels in the stratosphere increased by almost a factor of two, compared to version 2.2. Mid-layer averaging is not applied to these high-resolution data because the version 4.2 profiles are sufficiently smooth and we verified that averaging has little effect on the input data. Assimilating version 4.2 ozone at the nominal levels allows us to extend the vertical range of the input MLS profiles from 261 hPa (215 hPa since May 2016) to 0.02 hPa on 36 levels. However, preliminary comparisons indicate a high bias in the 261 hPa level compared to ozonesondes. Starting 1 May 2016 that level was turned off in MERRA-2 so that the lowest assimilated MLS level is 215 hPa from that date onward. MLS provides an almost unbroken record of observations from 2004 to present. The longest data gap occurs between 27 March and 18 April 2011 due to technical issues with the instrument.

The assumed observation errors are calculated from the square root of the sums of squares of precision (provided with the data) and 0.5 times the accuracy estimations from the MLS data quality documents (Livesey et al. 2007 for version 2.2 and Livesey et al. 2015 for version 4.2). The factor of 0.5 is applied because the accuracy estimates are quoted as a notional "2-sigma" term. The data selection is done before assimilation and follows the guidelines given in the MLS data quality documents.

This study evaluates the MERRA-2 ozone product between the years 1980 and 2013 when sufficient number of independent data (in particular from ozonesondes) are readily available to us.

3.4 Spatial coverage of the ozone data

Figure 1(a) shows the time series of the global monthly total ozone observation counts from data sources assimilated in MERRA-2. Each SBUV instrument except the sounder on Nimbus-7 provided between 30,000 and 40, 000 total ozone observations per month. The Nimbus-7 SBUV had two periods of slightly lower data counts: 1980–1983 and 1987–1990. From 1980–1983 SBUV operated on a 3-day on 1-day off cycle to reduce power load on the satellite, but as other instruments failed power constraints were reduced allowing SBUV to operate continuously after mid-1983. Nimbus-7 SBUV data after February 1987 are affected by chopper wheel synchronization errors. This caused an increase in measurement noise and a reduction in the number of usable profiles from SBUV over the remainder of the data record (Gleason and McPeters 1995). The monthly number of OMI observations (2004 onward) is about an order of magnitude greater.

Panel (b) of Figure 1 plots the monthly latitudinal coverage of each total ozone data type used. Besides a limited extent of the data during winter, the coverage of the SBUV instruments on NOAA-11 and NOAA-14 suffers from the effects of orbital drifts of these spacecraft. In particular, the equatorial crossing time of NOAA-11 changed from about 2PM in 1989 to 5PM in 1994, severely limiting SBUV coverage. By mid-1994 ozone observations were unavailable south of 30°S. Similarly, there was a loss of high-latitude coverage in 2001 caused by the drift of the NOAA-14 satellite. The MERRA-2 ozone product in the high latitudes in these years should be not be used. The OMI instrument provides steady coverage

of the sunlit atmosphere from 2004 onward with the exception of an outage between 29 May and 13 June 2016 (beyond the time period considered here).

The coverage of partial column ozone observations from SBUV (1980–2004) is the same as for the total ozone (in the SBUV period total ozone data are simply the sum of the SBUV layer values). The geographical coverage of MLS is 82°S-82°N throughout the period of the reanalysis.

4. Changes from MERRA relevant to ozone

The main difference in the treatment of ozone between MERRA and MERRA-2 is in the observing systems that the two reanalyses use. The ozone observations assimilated in MERRA were version 8 retrieved partial and total columns from SBUV. The partial columns were mapped from 21 to 12 layers and the total ozone column data were obtained by summing the layer values in each profile. By contrast, prior to October 2004 MERRA-2 uses the most recent version 8.6 of SBUV data, which is expected to provide a more realistic continuity across different satellite platforms as discussed in Section 3.1. These data are assimilated on the native 21 layers, although one should keep in mind that significant smoothing errors exist in the vertical (Kramarova et al. 2013). After 1 October 2004, MERRA-2 assimilates OMI and MLS data, which were not included in MERRA.

In data assimilation, the information from observations is applied to the background fields provided by the GCM forecast using background error covariances which control how the data information is propagated and optimally combined with the background in the horizontal and vertical directions (Cohn 1997; Lahoz et al. 2007). Following Wargan et al. (2015), the MERRA-2 data assimilation algorithm assumes that the background error standard deviation for ozone at any given grid point and time is proportional to the background concentration. Test experiments demonstrated that this leads to a more realistic representation of shallow vertical structures in the ozone profiles in the UTLS than when static errors, such as those employed for MERRA, are used.

The models in both reanalyses use the same set of monthly 2-dimensional ozone production rates and loss frequencies derived from a 2-dimensional chemistry model as described by Stajner et al. (2008) and Wargan et al. (2015). As argued in Wargan et al. (2015), such a simplified chemistry scheme is sufficient for a data-driven assimilated product in the upper troposphere and lower stratosphere where the bulk of ozone is located and where chemical time scales range from weeks to months – very long compared to the frequency of data insertion (daily). However, in the absence of day and night observations, prior to October 2004, the diurnal cycle in the upper stratospheric ozone is not represented in MERRA-2.

5. Independent ozone data

This section describes the data used for comparisons with MERRA and MERRA-2. None of the observations listed below are assimilated in either reanalysis.

5.1 TOMS

The Nimbus-7 Total Ozone Mapping Spectrometer (TOMS) instrument is a downward-viewing spectrometer, which measures Earth-backscattered UV radiances (Herman et al. 1991). Radiation in six discrete 1-nm wavelength bands (312.5, 317.5, 331.2, 339.8, 360, and 380 nm) is measured at 35 cross-track scan positions. Measurements from successive orbits overlap resulting in daily global coverage. The spatial resolution is 50 km by 50 km for the nadir view, and about 50 km by 200 km at the extreme cross-track scan positions. The TOMS instrument also measures the solar irradiance for each wavelength every day using a diffuser plate to reflect sunlight into the instrument. These solar irradiance measurements provide radiance normalization and remove some instrumental dependence. Total column ozone (TCO) data are derived from the Version 8 TOMS algorithm (Bhartia and Wellemeyer 2002; Wellemeyer et al. 2004), which was released in 2004. The algorithm uses only two wavelengths (317.5 and 331.2 nm) to derive TCO data; the other four wavelengths are used for error correction and identification of aerosols and clouds. Uncertainty estimates for Nimbus-7 TOMS total ozone are +3 % for absolute error, ±2 % (1 sigma) for random error, and +1.5% for the drift over 14 years with values somewhat larger at higher latitudes (McPeters et al. 1996).

5.2 Ozonesondes

The ozonesonde observations are as in Wargan et al. (2015): the data are from the Network for the Detection for Atmospheric Composition Change, the Southern Hemisphere Additional Ozonesondes (SHADOZ: Thompson et al. 2003a), as well as field campaigns. See Wargan et al. (2015) for a discussion of the ozonesonde locations, precision and accuracy and representativeness issues inherent in comparisons of point measurements with gridded analysis data.

In addition we compare MERRA-2 against the TCO data derived from the ozonesonde observations at the South Pole. Since the vertical range of balloon-borne observations typically does not extend above 10 hPa, the upper stratospheric contribution to the TCO provided with the ozonesonde data files is calculated assuming the constant mixing ratio above 7 hPa or at the altitude of balloon burst if that occurred below 7 hPa. The South Pole ozonesonde data was downloaded from the Earth System Research Laboratory website (http://www.esrl.noaa.gov/gmd/ozwv/ozsondes/spo.html).

5.3 UARS MLS

The Microwave Limb Sounder instrument on the Upper Atmosphere Research Satellite (UARS MLS: Barath et al. 1993) operated between September 1991 and July 1999 but the observation frequency declined significantly after 1993. In this work we use UARS MLS observations between 1991 and 1996. The spatial coverage alternated from month to month between 34°S-80°N and 80°S-34°N due to the spacecraft's monthly yaw maneuvers. We use Version 5 ozone profiles retrieved from UARS MLS measurements of the atmospheric limb emissions at 205 GHz described and validated by Livesey et al. (2003). All profiles were interpolated to a set of constant pressure levels, 16 per decade pressure. There are 29 UARS MLS levels between 60 hPa and 1 hPa (the vertical range used in our comparisons). We have applied a 2-point vertical boxcar smoother to all UARS profiles in order to reduce

unphysical oscillations resulting from the reported vertical resolution being higher than that resulting from the actual information content.

5.4 MIPAS

The Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) instrument was flown on the European Space Agency's (ESA) Envisat satellite, launched on 1 March 2002. It was a middle-infrared Fourier transform spectrometer measuring limb emissions in the 4.15 μm – 14.6 μm spectral range (Fischer et al. 2008). The instrument operated in its standard observation mode (nominal spectral resolution of 0.025 cm^{-1}) between July 2002 and March 2004. The operations were halted until the end of 2004 due to technical problems and resumed in January 2005 at a reduced spectral resolution of 0.0625 cm^{-1} and different selection of spectral microwindows, but with denser spatial sampling (Cortesi et al. 2007; Raspollini et al. 2013). The data record extends to early April 2012 when communication with Envisat was permanently lost. We use the ozone product from version 6 of the ESA retrieval algorithm described by Raspollini et al. (2013). The latter study concludes that the quality of the retrieved species is generally better for the 2005–2012 period. In particular the vertical resolution of ozone retrievals in the first period (2002–2004) ranges from 3.5 km to 5.5 km between 20 km and 50 km, but in the latter period ranges from 2.5 km to 4 km (Raspollini et al. 2013, their Figure 9). MIPAS profiles are interpolated to a 26-level vertical grid with 12 levels between 60 hPa and 1 hPa (note that this is different than for UARS MLS).

5.5 SAGE II

The Stratospheric Aerosol and Gas Experiment II (SAGE II) is a solar occultation instrument flown on the Earth Radiation Budget Satellite (ERBS) between October 1984 and August 2005 measuring the atmospheric transmission of the solar radiation in 7 channels nominally located between 1020 nm and 386 nm. The number of profiles is reduced by a factor of 2 after the year 2000. The occultation methodology provides high vertical resolution and high precision retrievals of stratospheric ozone (along with water vapor, NO_2 and aerosols) but limits the data coverage to 30 profile measurements per day, 15 at sunrise and 15 at sunset for each orbit. The latitudinal coverage varies from month to month with the extent rarely exceeding about 60°S to 60°N. The retrieved SAGE II ozone profiles have been used extensively in stratospheric ozone studies, as summarized in WMO (2014).

We use Version 7 of the retrieval algorithm described and evaluated by Damadeo et al. (2013). The profiles are interpolated onto a 61-level vertical grid between the surface and 0.2 hPa. Comparisons shown in this study use the SAGE II data at pressures lower than about 200 hPa.

6. Results of the Comparisons

This Section describes comparisons of MERRA-2 ozone fields against independent data from TOMS (total ozone); SAGE II, UARS MLS, MIPAS (stratospheric profiles) and ozonesondes (UTLS and integrated profiles). The comparisons are done as follows: for each independent observation, the closest MERRA-2 (and MERRA, if applicable) geolocation is

found. Then, for profile data, the reanalysis mixing ratio profile is interpolated linearly in log-pressure to the verifying observation levels. Unless stated otherwise, all the statistics considered here are calculated from pairs of data collocated in this way. We compute spatial and temporal means to identify biases and standard deviations and correlations to evaluate variability.

6.1. Total ozone – comparisons with TOMS and South Pole ozonesondes

Comparisons of MERRA and MERRA-2 TCO fields against TOMS data (January 1980 – May 1993) are done as follows. First, TOMS level 2 data are mapped onto a 1°×1° grid by averaging all measurements taken within each grid-box. The gridded data are then matched with the MERRA and MERRA-2 TCO in space and time (within 1 hour of the average time of measurements in the grid box) and monthly statistics are computed.

Figure 2 shows the time series of monthly TCO statistics broken down into 5 latitude bands: 90°S-60°S, 60°S-30°S, 30°S-30°N, 30°N-60°N and 60°N-90°N. The results are summarized in Table 2. Monthly mean percent differences (the left column in Figure 2, panels a, c, e, g and i) show that with the exception of the southern high latitudes (panel a) after 1990 the agreement between both reanalyses and TOMS is within 5 %, with some seasonal dependence outside of the tropics. Overall MERRA-2 is about 1 % lower than MERRA. This is consistent with the fact that the total ozone derived from the version 8.6 of SBUV algorithm is slightly lower than in version 8 used in MERRA (McPeters et al. 2013). The MERRA-2 minus TOMS total ozone difference averaged over the period of comparison is shown in Table 2. The reanalysis is slightly higher than TOMS in high latitudes (up to 1.45 % in the northern high latitudes) and lower by up to 1.83 % in the middle latitudes and the tropics. While a detailed discussion of MERRA-2 assimilation internal statistics is beyond the intended scope of this study we note that the total ozone tendencies due to dynamics in the GEOS-5 DAS are generally negative in the tropics and positive in high latitudes, suggesting that the meridional overturning (Brewer-Dobson) circulation from the assimilated meteorological fields is too rapid. This is in part countered by the tendencies arising from ozone data insertion, whose sign is the opposite of the dynamical tendencies on the average. The end result is a compromise between the tendencies due to dynamics and data assimilation, consistent with the pattern of positive (negative) bias in high (low-middle) latitudes seen in Table 2.

Starting in 1991 the agreement between the reanalyses and TOMS is degraded in the 90°S-60°S latitude band (Figure 1a). This is an expected result of the diminished coverage of the NOAA-11 SBUV in the southern high latitudes.

The right column of Figure 2 (panels b, d, f, h and j) shows the time series of monthly standard deviations of the reanalysis minus TOMS differences. The same statistics for MERRA-2 averaged over the period of comparison are summarized in Table 2. The difference standard deviations are almost identical for both reanalyses, ranging from about 2.5 % in the tropics (Figure 2f) to 8 % in high latitudes during spring (Figure 2b and j). The multiyear average for MERRA-2 is less than 5.5 % almost everywhere (Table 2), comparable with the assumed observation error for total ozone data and indicating excellent agreement of the reanalysis with TOMS. Here again, the agreement deteriorates in the

southern high latitudes after 1991 (Figure 2b). We note three spikes in the reanalysis minus TOMS standard deviation and mean time series in the tropics (Figure 2e and f), one in 1982 and two in 1991. Two of these spikes, denoted by the dashed lines in the figure, correspond to the eruptions of El Chichon and Mt. Pinatubo. High stratospheric aerosol loading following large volcanic eruptions degrades the quality of ozone retrievals from UV measurements (Bhartia et al. 2013). Shorter wavelength measurements (SBUV profile) are more affected than longer wavelength measurements (TOMS TCO) when aerosol levels are elevated above the ozone peak, leading to a different response between SBUV and TOMS (Torres et al. 1995; Torres and Bhartia 1995). Averaging over multiple scan angles further reduces the aerosol signal in TOMS (Torres et al. 1995). The third spike, denoted by the solid line, results from a full month of missing SBUV data (see Section 3.1).

When considering total ozone column, it is instructive to compare MERRA-2 with the SBUV-derived Merged Ozone Data set (MOD: Frith et al. 2014) based on version 8.6 of SBUV retrievals, the same as MERRA-2 uses. Figure 3 plots the relative difference between MERRA-2 total ozone and MOD data as a function of latitude averaged over two time periods: 1980–2003 (SBUV period) and 2005–2014 (EOS Aura period). The comparison serves as a consistency (MERRA-2 assimilates SBUV data in the former period) and continuity check between the two periods. In the SBUV period MERRA-2 is slightly lower than MOD (by up to 1 %) and the difference is approximately symmetric with respect to the equator. By contrast, the low bias in the EOS Aura period exhibits a clear latitude dependence: it is in agreement with the MERRA-2 – MOD difference in the southern hemisphere but in the northern hemisphere the absolute value of the bias increases linearly with latitude reaching ~2 % at 80°N. This is consistent with the latitude-dependent low bias in the OMI data (Section 3.2).

Figure 4 shows comparisons of total ozone from MERRA and MERRA-2 at the South Pole with values derived from ozonesonde measurements. All three datasets show pronounced annual variations (Figure 4a) with sharp minima reaching down to 100 DU during ozone hole conditions (September-October) followed by maxima in austral summers and periods of decline leading up to the next year's ozone hole. Panel (b) of Figure 4 shows the reanalysis minus ozonesonde differences relative to individual total ozone values derived from the ozonesonde profiles. Aside from large positive excursions during the austral winter-spring transitions (with limited UV data coverage) occasionally reaching over 100 %, the differences are within 20 % for both reanalyses. Statistics calculated from these comparisons show that during both the SBUV and Aura periods MERRA-2 is lower than the sondes by 3 %. The standard deviation of MERRA-2 minus ozonesonde differences is 12.47 % in the SBUV period and only 5 % in the Aura period. The corresponding standard deviation numbers for MERRA are 13.3 % and 15.9 %, respectively, indicating a much better performance of MERRA-2, especially in the period when MLS data are assimilated. We note the realism of the interannual variability in MERRA-2. In particular, the record high springtime ozone associated with the major sudden stratospheric warming in September 2002 is well represented by both MERRA and MERRA-2 (Figure 4a).

Overall, the total ozone product in MERRA-2 compares well with independent data. While small systematic season-dependent biases exist, the annual cycle, latitudinal structure and

longer-term variability are realistic and the agreement with the independent data is well within the assumed observation errors.

6.2 Comparisons with SAGE II

Comparisons of MERRA and MERRA-2 against SAGE II observations are shown in Figures 5–8 and Tables 3 and 4 for the period 1984–2005, though the number of SAGE profiles is reduced by a factor of two after 2000. Note that most of the SAGE II record falls into the SBUV period for MERRA-2.

Figure 5 shows the time series of annual relative mean ozone differences integrated between 208 hPa and 0.2 hPa (hereafter, stratospheric column) in three broad latitude bands: south of 30°S, between 30°S and 30°N, and north of 30°N. MERRA and MERRA-2 minus SAGE II differences are shown. Note that the latitudinal extent of SAGE II observations is about 60°S-60°N so that the comparisons do not cover the polar regions. The years 1992 and 1993 are omitted because of high aerosol loading affecting ozone retrievals in the lower stratosphere in years following the eruption of Mt. Pinatubo (1991). The differences are within 2 % for both reanalyses and exhibit an upward drift over the period of comparison in the southern hemisphere and in the tropics. Note the large difference between MERRA and SAGE II south of 30°S in 1994 due to limited SBUV coverage (see Figure 1). This effect is much reduced in MERRA-2. For most of the comparison period MERRA-2 is slightly higher than SAGE II in all latitude bands. Wang et al. (2002) report that version 6.1 of SAGE II ozone data exhibit a low bias with respect to ozonesondes, which is most pronounced in the tropics below 15 km (~125 hPa) but also present at the middle latitudes (see their Figures 9 and 10). While Version 7 of the SAGE II data used here is much improved in the middle to upper stratosphere compared to previous versions (Damadeo et al. 2013), the differences in the lower stratosphere are smaller and it can be expected that some of the bias between the reanalyses and SAGE II stratospheric column is due to an underestimate of the lower stratospheric ozone in SAGE II retrievals.

Figure 6 (a–c) plots the time series of annually averaged zonal mean ozone from SAGE II and both reanalyses south of 30°S at three SAGE II levels: 42.7 hPa (lower stratosphere), 10.1 hPa (middle stratosphere) and 4.3 hPa (upper stratosphere). The statistics averaged over the period of comparison (1984–2005) are given in Table 3 for MERRA-2 only. At 42.7 hPa the MERRA-2 – SAGE II differences are small (2 %) and both reanalyses follow the interannual variability seen in SAGE II observations. The same is true at 10.1 hPa with the exception of 1994 when MERRA is ~0.4 parts per million by volume (ppmv), equivalently 6 %, higher than SAGE II. The bias in MERRA-2 is about half of that value. As seen in Figure 1 the NOAA-11 SBUV coverage was very limited in that year, especially in the southern hemisphere. At 4.3 hPa both reanalyses are very close to SAGE II between 1985 and 1991. Between 1992 and 2004 MERRA is slightly closer to SAGE II measurements than MERRA-2.

Panels (d) – (f) of Figure 6 show the standard deviations of the SAGE II minus reanalysis differences at the same levels. Table 3 lists the MERRA-2 results averaged between 1984 and 2005. The results for both reanalyses are close to each other with the exception of 2005 (the first full year of the Aura data assimilation) at 42.7 hPa where the agreement of

MERRA-2 with SAGE II improves by almost a factor of 2. The MERRA-2 – SAGE II difference standard deviations range from 0.31 ppmv (4.6 %) at 4.3 hPa to 0.35 ppmv (11 %) at 42.7 hPa.

We will now describe briefly the results for the northern middle latitudes (north of 30N°, not shown). They are qualitatively similar to those for the southern hemisphere. Of note is the fact that after 1995 MERRA-2 has a larger positive bias with respect to SAGE II at 4.2 hPa but captures more of the variability within that year. A large drop in the difference standard deviation at 42.7 hPa in 2005 for MERRA-2 is seen in both hemispheres, though the number of SAGE profiles is limited in 2005. The results for the northern middle latitudes are summarized in Table 4.

The bottom row (panels g–i) of Figure 6 shows the annual SAGE II data counts for each level and hemisphere. In the middle and upper stratosphere there are close to 4,000 observations per year until the year 2000 and about 2,000 afterwards. The data counts in the lower stratosphere (at 42.7 hPa) are slightly above 2,000 before the year 2000 except in the years following the Mt. Pinatubo eruption when the numbers drop to less than 1,000.

Sunrise ozone mixing ratios from SAGE II tend to be lower by 8–10 % relative to sunset data above 35 km owing to diurnal variations in upper stratospheric ozone and a possible contribution from algorithmic biases (Kyrölä et al. 2013; Damadeo et al. 2014; Sakazaki et al. 2015). Tables 3 and 4 include statistics calculated from MERRA-2 and SAGE II for sunset-only measurements (in parentheses). While the difference standard deviations calculated for the sunset-only data are very close to the all-data values the biases are much lower, especially in the upper stratosphere. For example, at 4.2 hPa in the northern hemisphere the MERRA-2 bias drops from 1.1 % to 0.3 % if only sunset observations are used.

In Figure 7 we show the mean and standard deviation of differences between the reanalyses and SAGE II profiles relative to the SAGE II mean from all available data between January and August 2003 (the SBUV period) and for the same months in 2005 (the Aura period). The results are given for three latitude bands: south of 30°S, between 30°S and 30°N, and north of 30°N. Both reanalyses exhibit a low bias of as much as to 5 % between 5 hPa and 1 hPa, but the bias is reduced in MERRA-2 in 2005 compared to 2003. In 2003 MERRA-2 has an alternating pattern of low and high bias with respect to SAGE II in all latitude bands. By contrast, in 2005 there is almost no bias between MERRA-2 and SAGE II in the middle stratosphere (10–50 hPa) in the extratropics and only a small (up to 3 %) negative bias around 10 hPa in the tropics. The difference standard deviation ranges from 10 % – 30 % between 100 hPa and 60 hPa (30 % in the tropics near 100 hPa) to less than 10 % in the middle stratosphere. While in 2003 (the SBUV period) the standard deviation profiles are almost identical between MERRA and MERRA-2, the latter better captures the SAGE II variability in 2005 when stratospheric ozone is constrained by higher vertical resolution MLS observations. SAGE II – MERRA-2 correlation profiles (not shown) also improve in 2005 compared to 2003. South of 30°S the correlation increases from ~0.85 to ~0.95 at pressures higher than 20 hPa with the largest improvement at 40 hPa where the correlations are 0.77 and 0.94 for 2003 and 2005, respectively. In the middle stratosphere the correlations

are similar for both years (~0.95) with larger differences at pressures lower than 2 hPa. Similar improvement is seen north of 30°N. In the tropics, the correlations also increase by about 0.1 throughout the stratosphere with the 2005 values between 0.8 and 0.95, except in the upper stratosphere (the correlation is 0.5 at 1 hPa in 2005). These results confirm a good representation of vertical structures in the MERRA-2 stratospheric ozone and a pronounced further improvement in the Aura era.

6.3 Comparisons with UARS MLS

Because of the limited time span of UARS MLS measurements (days when data are available become sparse after 1993, we will only analyze the relevant statistics aggregated over the 1991 – 1996 period and refrain from discussing time-dependent statistics. Note that the years of UARS MLS operations fall entirely within the SBUV period of MERRA-2.

The results are shown in Figure 8 separately for 4 seasons: December-February (DJF), March-May (MAM), June-August (JJA), and September-November (SON). The top row shows the zonal mean MERRA-2 minus UARS MLS difference as a percent of the average UARS MLS ozone at each level between 60 hPa and 1 hPa for each of the seasons. Overall, MERRA-2 is lower than UARS MLS. In most of the stratosphere the pattern is similar in each season and consists of a low bias layer (4 % – 12 %) between 3 hPa and 1 hPa, and between 10 hPa and 30 hPa, smaller high bias between 50 hPa and 30 hPa (except in high latitudes where the difference is negative) and again low bias below 50 hPa. The latter is large (over 20 %) but one must remember that the mean ozone concentrations are small in the lower stratosphere, contributing to a large relative difference. The MERRA-2 minus UARS MLS difference between 100 hPa (not shown) and 50 hPa does not exceed ~0.2 ppmv. Also seen in Figure 8 are regions of positive differences in the high latitudes between 10 hPa and 2 hPa in the northern hemisphere during DJF, in the southern hemisphere during MAM and JJA, and in both hemispheres during SON. The largest differences of up to 25 % are seen in JJA at about 10 hPa in the southernmost latitudes.

The middle row of Figure 8 (panels e–h) show the standard deviation of the MERRA-2 – UARS MLS difference expressed as a percent of the average UARS MLS ozone. Also plotted is the zonal standard deviation of UARS MLS ozone calculated for the same periods as a measure of the ozone fields' variability (white contours). Broadly speaking, the difference standard deviations and the variability contours exhibit very similar patterns with the former being always less than the latter, indicative of a realistic representation of the stratospheric ozone variability in MERRA-2. In most of the stratosphere the difference standard deviation is much less than 10 %. Higher values of up to 25 % are seen only in regions of high variability such as the high latitudes in southern hemisphere winter and spring, where SBUV measurements are limited, and in the lower stratosphere in the tropics, where ozone concentrations are low.

Correlations between MERRA-2 and UARS MLS are shown in panels (i)–(l) of Figure 8. The zonal patterns follow that of the ozone variability with the highest correlations of 0.7 and higher in the extratropics (somewhat reduced during summer). The lowest, even negative correlations are found in the tropics, where they exhibit a vertical pattern resembling the QBO signature in ozone. Coy et al. (2016) showed that the ozone response to

the QBO in MERRA-2 does not exhibit a realistic vertical structure during the SBUV period. This is a result of a limited vertical resolution (large smoothing errors) in SBUV profile data as explained by Kramarova et al. (2012).

We want to emphasize that despite the differences between MERRA-2 and UARS MLS, the two data sets agree quite well in terms of the structure of the ozone fields. As an example, Figure 9 shows the ozone field at 10 hPa on 21 December 1993. A breaking planetary wave pulled a thin tongue of ozone-rich subtropical air from over the coast of East Asia into the high latitudes, transporting it along the edge of the polar vortex almost over the North Pole and wrapping it around the Aleutian anticyclone. This is a fairly common occurrence in the winter polar stratosphere. In this case MERRA-2 shows that the polar vortex was displaced towards Europe. The 10 hPa temperature over the pole increased rapidly by ~40°K during the final days of December (not shown), indicative of a sudden stratospheric warming. Overlaid on the MERRA-2 ozone field in Figure 9 are UARS MLS observations taken within 4.5 hours of 6Z, the time of the analysis. There is considerable agreement between the data sets: MLS samples the low mixing ratio inside the polar vortex, the tongue of high ozone from Kamchatka into the Arctic Ocean, the region of high values over Canada, and the high ozone filament between 30°N and 40°N from the North Atlantic to north of East Siberia. We note that the NOAA-11 SBUV coverage extended up to ~51°N at that time (marked by black circles in Figure 9(a)) so that all the information about the ozone distribution over the polar region in MERRA-2 is generated by the general circulation model, in particular via advection of air masses from lower latitudes where observations are available. Figure 9(b) shows a scatter plot of 10 hPa 30°N-90°N UARS MLS observations and MERRA-2 sampled at the observation locations on the same day. The mean difference is 0.08 ppmv (~1.3%) with the standard deviation of the differences of 0.48 ppmv and a correlation coefficient of 0.94. This example demonstrates the ability of the reanalysis to reproduce this dynamically driven structure accurately and corroborates the high correlations seen in the northern hemisphere during winter (Figure 8i).

6.4 Comparisons with MIPAS

Figure 10 shows zonal mean and difference standard deviation statistics between MERRA-2 and MIPAS ozone as a function of latitude and pressure (60 hPa – 1 hPa), analogous to Figure 8 for UARS MLS. The statistics are accumulated from 2003 – 2012, primarily in the Aura time period. As in the case of UARS MLS comparisons, MERRA-2 is biased low with respect to MIPAS in most of the stratosphere (Figure 10, top row). In particular, there is a low bias of up to 8 % in the upper stratosphere between 3 hPa and 1 hPa. Between 30 hPa and 10 hPa the negative bias is up to 3 % and up to 7 % below 30 hPa. The only region where MERRA-2 is higher than MIPAS is the southern high latitudes in the upper stratosphere in MAM and JJA (~1 %) and around 10 hPa in JJA (up to 5 %). Overall, the low bias pattern is consistent throughout all seasons below 10 hPa.

The middle row of Figure 10 shows the MERRA-2 minus MIPAS standard deviations (colors) and the standard deviations of the MIPAS ozone (contours) as a measure of ozone variability. The results are very similar to those of the UARS MLS comparisons in the previous section. In most of the stratosphere the difference standard deviation is less than

10 % and it is less than the variability, indicative of a good agreement between the reanalysis and MIPAS data. Higher values of 15 % - 20 % are seen in high latitudes in winter and spring when the variability is increased compared to other seasons.

The MIPAS – MERRA-2 correlations are shown in the bottom row of Figure 10 (panels i-l). Overall, these are much higher than for UARS MLS (the bottom row of Figure 8), often in excess of 0.8 in all seasons, especially in the extratropics. The weakest correlations (less than 0.4) are found in the tropics where the ozone variability is the lowest in a given season, consistent with the standard deviations shown in the middle row of Figure 10.

While the number of available MIPAS observations varies from year to year, there is enough time continuity to calculate time series of monthly statistics. These are shown in Figure 11 for MIPAS, MERRA and MERRA-2 at 9.42 hPa and 39.8 hPa in the 30°N-60°N latitude band. The zonal mean time series show distinct annual cycles at both levels with MERRA-2 biased low by ~ 0.6 ppmv at 9.42 hPa in summer during the MLS period and by ~0.1 – 0.2 ppmv at 39.8 hPa. At 9.42 hPa MERRA is in a closer agreement with MIPAS than MERRA-2, while the opposite is true at 39.4 hPa. The standard deviations of the analysis minus MIPAS differences (panels (c) and (d)) also vary seasonally, with the maximum differences in wintertime, as expected given the increased dynamical variability in winter. At 9.42 hPa the standard deviations are similar for both reanalyses, ranging between 0.25 ppmv in (~3 %) in summer and 0.6 ppmv (~10 %) during winter. At 39.4 hPa the standard deviations are the same for both reanalyses in 2003, between 0.3 ppmv (~9 %, summer) and 0.6 ppmv (15%, winter). However, in the Aura period (denoted by a yellow vertical line in the figure), while the standard deviations against MERRA continue to average ~ 0.4 ppmv, the standard deviations are much lower for MERRA-2, in the 1.8 to 3.2 ppmv range. The introduction of MLS observations led to much better agreement of the analysis ozone with MIPAS data, consistent with the larger smoothing error at 40 hPa in SBUV (Kramarova et al. 2012). Note the spike in the MERRA-2-MIPAS standard deviations in April 2011 coinciding with an over three-week period of missing MLS observations. The monthly MIPAS data counts are shown in panels (e) and (f) of Figure 11.

Comparisons in other latitude bands (not shown) reveal similar patterns: the reanalyses capture the annual cycles of ozone at 9.42 hPa and 39.8 hPa and the agreement between MERRA-2 and MIPAS improves in the Aura period, especially in the middle and high latitudes. The greatest differences between the two reanalyses are found in the lower stratosphere between 90°S and 60°S in austral winter when MERRA-2 performs much better as a result of polar ozone observations from MLS.

6.5 Summary of the comparisons against satellites

This subsection summarizes comparisons of the MERRA-2 stratospheric ozone against SAGE II, UARS MLS and MIPAS. The following are the main results of the zonal mean comparisons.

- The mean MERRA-2 total ozone agrees with TOMS within 2 %. The difference standard deviation does not exceed 6 %.

- In the upper stratosphere (3 hPa – 1 hPa) the reanalysis is biased low by 5 – 8 % with respect to all three data sets in both SBUV and MLS periods.

- In the middle stratosphere there is agreement between the SAGE II and UARS MLS comparisons in that, at least outside of high latitudes, MERRA-2 exhibits negative bias of up to ~5 % between 30 hPa and 10 hPa and a small positive bias around 40 hPa. By contrast, MERRA-2 is low compared to MIPAS throughout the 40–10 hPa layer. We note that the MIPAS comparisons are done mainly in the Aura period, during which comparisons with SAGE II show almost no bias between 50 hPa and 10 hPa (Figure 7d-f). Hubert et al. (2016) find that MIPAS overestimates ozone below 50 hPa compared to ozonesondes and lidar observations, especially in the tropics where the differences exceed 10 % (see their Figure 6) and report a much better agreement between SAGE II and independent data.

- Between 60 hPa and 50 hPa MERRA-2 is biased low compared to SAGE II, UARS MLS and MIPAS observations..

The key conclusions regarding the stratospheric ozone variability in MERRA-2 are summarized as follows.

- The difference standard deviations are within 20 % between 100 hPa and 1 hPa and are much lower (within 10 %) throughout the middle stratosphere except in regions of high ozone variability. Comparisons with the standard deviation of observed ozone concentrations (Figures 8 and 10) indicate a very realistic representation of variability in MERRA-2. This conclusion is corroborated by a case study showing excellent agreement between the morphology of the MERRA-2 ozone field at 10 hPa with UARS MLS during a complex polar transport event (Figure 9).

- The agreement between MERRA-2 and independent satellite data (SAGE II and MIPAS) improves dramatically in the Aura period (from late 2004 onward), particularly in the lower stratosphere where SBUV has reduced vertical resolution.

6.6 Comparison with ozonesondes

In this subsection we evaluate the MERRA-2 ozone in the upper troposphere and lower stratosphere against balloon-borne ozonesonde observations. The comparisons are done as follows: each ozonesonde profile is interpolated to a common vertical grid by averaging the sonde measurements within a set of layers 1 km thick; the analysis profile the closest in time and geolocation is mapped to the same vertical grid using cubic splines. In order to separate tropospheric and stratospheric ozone the comparisons are done relative to the tropopause, here defined as the 2 potential vorticity units (PVU) isosurface outside of the 10°S–10°N latitude band and as the 100 hPa pressure surface within the band. The choice of 2 PVU was motivated by the observation that the vertical gradient of ozone concentrations in the sonde data changes sharply at this pressure. The same definition of the tropopause was employed by Wargan et al. (2015). A tropopause-centered analysis was also used by Sofieva et al. (2014) to construct an ozone climatology using ozonesonde and SAGE II data. The

advantage of this approach lies in its ability to separate the tropospheric and stratospheric ozone content and emphasize the sharp cross-tropopause gradients in the ozone profiles. The potential vorticity fields are taken from the MERRA-2 output. The results for 2003 (SBUV period) and 2005 (Aura period) are shown in Figure 12. In both years all available sondes were used. There were 1632 soundings in 2003 and 2425 in 2005. We note that the results were almost unchanged when the 2005 data were randomly sub-sampled to a set of 1632 soundings. Note that the mean profiles are expressed in the units of partial pressure (millipascals, mPa). As seen in panels (a) and (d) there is very good agreement between the mean profiles from MERRA-2 and the ozonesondes in both years, however the vertical gradient in the MERRA-2 ozone within the 2 km layer below the tropopause is sharper (and in a better agreement with the sondes) in 2005. In the troposphere MERRA-2 is biased low by up to ~0.2–0.3 mPa (5 – 12 %) in 2005. In 2003 there is a small high bias in the uppermost troposphere and small low bias below the 5 km mark. In the stratosphere there are patterns of small alternating biases not exceeding ~0.5 mPa (about 5 %). The standard deviations of the sonde data shown as the magenta lines in panels (a) and (b) have a very similar structure in both years. We note the enhanced variability within the 10 km layer above the tropopause caused by vigorous advection associated with Rossby wave breaking in the middle latitude surf zone during winter and spring (McIntyre and Palmer 1984). The standard deviation of the MERRA-2 minus sonde differences (panels a, b, d, and e) is smaller than the standard deviation of the sondes, especially above the tropopause, indicative of a good representation of the ozone variability in that layer. This is particularly evident in 2005 but it is also true in 2003. The ozonesonde – MERRA-2 correlations shown in panels (c) and (f) are 0.7 – 0.8 below the tropopause in both years. In the layer of maximum variability (within 10 km above the tropopause) the correlations are 0.8 and above 0.9 in 2003 and 2005, respectively. These statistics are calculated from all available ozonesondes and all seasons and therefore include latitudinal and seasonal as well as small-scale, short-term variability. In order to test the robustness of the results we repeated the comparisons for sonde locations between 0°E-60°E, 45°N-60°N in the spring (March-May) 2003 and 2005. The results are shown in Figure 13. The lower-stratospheric ozone variability is larger here than in the global case (Figure 12) as is the gradient across the tropopause. Both are captured very well by MERRA-2 in the Aura period (panels d and e) but in the SBUV period the agreement is less good than in the global average (compare panels a and b of Figures 12 and 13). The 'kink' in the ozonesonde profiles at 2 km and 5 km above the tropopause is a result of frequent isentropic intrusions of ozone-poor tropical air into the mid-latitudinal lower stratosphere often associated with a double tropopause (Sofieva et al. 2014). The feature is reproduced by MERRA-2 in both years but the agreement with the sondes is closer in 2005. In 2003 the ozonesonde-analysis correlations are low at the tropopause and the upper troposphere but they reach approximately 0.6 in the layer between 5 km and 10 km above the tropopause and 0.8 above. In 2005 the stratospheric correlations are higher, similar to the global case but they are much lower in the troposphere except in the 2 km layer below the tropopause. The maximum ozone variability relative to the mean occurs at the dynamical tropopause itself. This is also where the reanalysis minus ozonesonde correlations reach a minimum as seen in Figures 12 and 14. The latter feature is also present but less pronounced when thresholds for the dynamical tropopause of 4 PVU and 6 PVU are used.

Following Wargan et al. (2015) we perform separate comparisons of ozone integrated within two layers: between 500 hPa and the tropopause (hereafter, upper troposphere or UT) and between the tropopause and 50 hPa (lower stratosphere or LS). The results are expressed in Dobson units. This analysis is motivated by the importance of an accurate representation (including an assessment of errors) of the mid-latitude lower stratospheric ozone and the separation of the stratospheric and upper-tropospheric air masses for studies of tracer transport and radiative forcing. Monthly statistics calculated from MERRA and MERRA-2 are compared against ozonesondes between 30°N and 60°N for the period 1991 to 2012. The results are shown in Figures 14 and 15 for the upper troposphere and lower stratosphere, respectively. The MERRA-2 statistics are also summarized in Table 5. Note that the number of sondes used per month increases in time from only a few before 1993 to about 60 between 1997–2003 and then to 100 and above starting 2004 (Figure 14a). In particular, the use of frequent soundings from field campaigns results in several spikes in data counts. We first discuss the results in the UT shown in Figure 14. Both reanalyses reproduce the month-to-month variations including the annual cycle seen in the sonde data throughout the period of comparison. In the SBUV period the reanalyses are close to each other with a small positive bias of no more than 3 DU with respect to the sondes. For MERRA-2 the bias is within 10 % of the sonde mean (Table 5). In the Aura period MERRA-2 becomes systematically lower than the sondes by 13.6 % on the average. The bias has a seasonal dependence and varies between -1 DU and -4 DU. The existence of a low bias in the UT in the Aura period is consistent with the results of Wargan et al. (2015) where it was attributed to the absence of NO_x chemistry in the model and a low sensitivity of OMI observations to the ozone below 500 hPa. The latter affects the UT through model transport. Before the introduction of OMI and MLS observations in 2004 (and in MERRA) the UT ozone is constrained by the partial column information from SBUV, which although largely a priori information at these altitudes, prevents it from deviating significantly from a climatological average. The standard deviation of the analysis – sonde differences (Figure 14c), similar for MERRA and MERRA-2, is generally between 2 DU and 4 DU. It is lower for MERRA-2 in the Aura period by about 1 DU, indicative of a better representation of the UT ozone variability after 2004. We performed a 95 % one-tailed F-test to assess the significance of the ozonesonde-analysis standard deviations between the two reanalyses: if MERRA-2 is closer to the sondes (the black line in Figure 14c is above zero) then the statistic $\sigma(MERA2 - sonades)/\sigma(MERA- sondes)$ (where the symbol σ denotes the standard deviation) is calculated and compared with the appropriate threshold given by the F-distribution in order to determine if the difference is significant. If MERRA is closer to the sondes then the same is applied to the reciprocal of the quotient above. Significant differences are marked by blue (MERRA is closer to the sondes) and red (MERRA-2 is closer) squares overplotted on the zero line. The sonde-MERRA-2 correlation also improves from 0.81 to 0.9 between the two periods (Figure 14d).

In the LS the agreement between the reanalyses and sonde data is very close (Figure 15a). For MERRA-2 the average bias is 3.8 % and 1.2 % in the SBUV and Aura periods, respectively. The difference standard deviations are 11.2 % and 8.1% and the analysis-sonde correlations are 0.96 and 0.98 (Figure 15b and c). In the Aura period MERRA-2 is significantly closer to the sondes in terms of standard deviation of the analysis-sonde

differences as indicated by the difference line and the results of the F-test (the red square marks in Figure 15b).

We reiterate that the structure of the tropospheric ozone profiles in MERRA-2 is mainly controlled by transport by assimilated winds and parameterized convection. In the absence of an explicit representation of NO_x chemistry and high-resolution observations in the troposphere the quality of the ozone profiles below the tropopause is unavoidably degraded compared to the stratosphere.

Figures 16 and 17 show comparisons between the UT and LS ozone from two reanalyses and ozone sondes as a function of latitude for the SBUV (1991–2003) and Aura (2005–2012) periods, respectively. The statistics are calculated within 5° wide latitude bins with the condition that at least 10 soundings are available in a given bin. In the SBUV period the mean agreement between the reanalyses and the ozonesondes in the LS is within 10 % everywhere. It is within 20 % in the UT except between 30°S and the Equator, where both reanalyses overestimate ozone by ~50% with respect to the sonde data. This large overestimate results from a 5 DU positive bias in MERRA and MERRA-2 at three stations located in the West Pacific ozone minimum: Samoa, Fiji and Tahiti. The minimum, a part of the wave-one pattern in the tropical ozone, is associated with strong convection and resulting lofting of ozone-poor air leading to a decrease of the tropospheric ozone column (Thompson et al. 2003b). In the absence of accurate boundary layer ozone chemistry in the model this feature is not correctly represented in the reanalyses. In the EOS Aura period the agreement between MERRA-2 and ozonesondes in that region is much closer (Figure 17e). Standard deviations of the differences (panels c and f) are between 10 % and 20 % in the LS and between 15 % and 35 % in the UT. MERRA-2 exhibits a somewhat worse agreement than MERRA in the northern high latitudes. In the Aura period (Figure 17) the LS bias is within 10 % except in the tropics where the ozone columns are small (~ 20 DU) and where MERRA-2 overestimates it by 20 %. This is consistent with the result seen in the profile comparisons with SAGE II (Figure 7e) and the findings of Froidevaux et al. (2008) who showed that version 2.2 of MLS ozone exhibit a positive bias of up to 50 % in the tropical LS with respect to SAGE II data. A 20 % underestimate in the MERRA-2 UT in the northern hemisphere (Figure 17e) has a similar latitudinal structure to that seen for the reanalysis total ozone comparison with the MOD data in Figure 3. We attribute this underestimate to the latitude-dependent bias in the OMI data (McPeters et al. 2015). As expected, MERRA-2 performs better than MERRA in terms of the difference standard deviations in the LS at almost all latitudes (Figure 17c). A smaller improvement is also seen in the UT in that period (Figure 17f).

An example of the upper tropospheric (500 hPa to the tropopause) ozone distribution from MERRA-2 are shown in Figure 18 for 1 July 2013. The plot illustrates the degree of spatial variability resolved by the reanalysis and the realism of the large-scale patterns: the summertime maximum in the northern hemisphere, a wave 1 pattern in the tropics with the minimum over the Western Pacific and enhancements over the southern tropical Atlantic, and high values along the subtropical jet stream over the southern Indian Ocean. Similar seasonal features in tropospheric ozone fields have been reported by Ziemke et al. (2011)

who used an earlier version of MLS and OMI data to construct a monthly tropospheric ozone climatology.

6.7 Summary of the ozonesonde comparisons

The main results of the comparisons with ozonesondes are summarized as follows

- In the lower stratosphere the agreement in terms of standard deviation in the tropopause-50 hPa layer is within 20 % globally and about 10 % in the northern middle latitudes. In the Aura period the agreement improves to less than 10 % in the extratropics.

- The vertical structure of the LS ozone and the cross-tropopause gradient is much better represented in the Aura period (Figures 12 and 13)

- In the upper troposphere the standard deviation of the reanalysis minus ozonesonde differences is within 25 % and improves slightly in the Aura period but this improvement comes at the expense of a low bias of 13.6 %.

7. Conclusions

In this study we described and evaluated the ozone product in GMAO's MERRA-2 reanalysis aiming to address the following two questions

1. How well does MERRA-2 represent the spatio-temporal variability of the stratospheric and upper-tropospheric ozone?

2. What are the impacts of changing the ozone observing system in MERRA-2 in late 2004 when new types of data were introduced?

By combining observations from multiple sources with 6-hourly forecasts the data assimilation methodology enables the production of high-frequency global gridded ozone fields consistent with dynamical variability ranging from hourly to interannual scales, thus providing an added value compared with ozone observations alone. The main goal of this paper is to assess the usefulness of the MERRA-2 ozone for scientific applications by highlighting the strengths as well as deficiencies of the reanalysis ozone.

The treatment of ozone data in MERRA-2 represents advancement over the original MERRA reanalysis. In particular, the new reanalysis employs an improved version of SBUV retrievals (version 8.6) between 1980 and 2004 and introduces EOS Aura data in the later period with the OMI instrument providing over 300,000 total ozone observations monthly and MLS constraining the stratosphere with observations at a relatively high-resolution of about 2.5 km in the vertical and near-global day and night coverage.

The main findings are summarized as follows

- MERRA-2 total ozone agrees with TOMS data (1980–1993) very well, with less than 2 % bias and less than 6 % difference standard deviation, close to the assumed observation error of 5 %.

- The bias in the stratospheric profiles with respect to SAGE II is reduced in 2005 when MLS observations were assimilated compared to the SBUV period. The difference standard deviations with MIPAS are significantly smaller in the Aura period than in the SBUV period in the lower stratosphere. This is an expected consequence of the MLS profile data assimilation.

- There is good representation of the variability of stratospheric ozone in MERRA-2. The difference standard deviations between the reanalysis and independent limb satellite data range from 11 % for SAGE II in the lower stratosphere to less than 5 % at 4.3 hPa. Comparisons against UARS MLS (the SBUV period) and MIPAS (mostly the Aura period) produced less that 10 % differences in most of the stratosphere except in regions of high variability such as winter high latitudes. The difference standard deviations do not exceed the standard deviation of the observed ozone concentrations (Figures 8 and 10), indicating the ability of the analysis to reproduce day-to-day ozone variability. As illustrated by a case study (Section 6.3. Figure 9) the reanalysis captures the complex dynamically driven morphology of the ozone field at 10 hPa during a cross-pole transport of middle-latitude high ozone in excellent agreement with UARS MLS data.

- In the northern middle latitudes the UT ozone is 9.5 % high in the SBUV period and 13.6 % low in the Aura period compared to ozonesondes. After 2004 the bias has some seasonal dependence but does not change much from year to year. By contrast, in the LS the bias is reduced in the Aura period but it is also small in the SBUV period (3.8 %). Overall, MERRA-2 reproduces the interannual to weekly (approximate frequency of ozonesonde observations) variability of the UTLS ozone as well as the sharp gradient of ozone concentrations across the tropopause.

The introduction of EOS Aura data leads to an improved representation of ozone in MERRA-2 but it also generates a discontinuity that should be taken into account in applications of the reanalysis to climate studies, in particular in trend analyses. The notable discontinuities in the ozone observing system are summarized in Table 6.

The results of this study provide a sound justification for using the MERRA-2 stratospheric and upper tropospheric ozone in scientific research, in particular for studies requiring high frequency, highly resolved global ozone maps with realistic variability consistent with dynamics. They highlight the value of NASA's research data from the EOS Aura mission to the reanalyses of long-term ozone records. While discontinuities have to be taken into account, MERRA-2 should be a viable dataset for use in monitoring the long-term changes in total and mid-stratospheric ozone as atmospheric halogen levels decrease. This includes the study of inter-annual changes (such as El Nino and the QBO), and potential impacts of extreme meteorological events (such as hurricanes) on ozone. For long-term time series analysis in the UTLS region extreme caution is recommended. Because of the larger SBUV smoothing error, signals in the UTLS that vary in altitude, such as the QBO discussed in section 6.3, will be represented differently in the SBUV and MLS time periods. Analyses requiring high vertical resolution, such as the study of inter-annual changes in mid-latitude

stratosphere-troposphere exchange in relation to dynamical conditions, are recommended for the MLS-period only. In future reanalyses we will work to incorporate the SBUV averaging kernels to better represent the true vertical resolution of the input data. Despite these cautions, the results discussed here demonstrate progress towards representing ozone in a realistic fashion in Earth System reanalyses.

Acknowledgments

References

Bass, AM., Paur, RJ. The ultraviolet cross-sections of ozone, I, The measurements. In: Zerefos, CS., Ghazi, A., editors. Atmospheric Ozone. D. Reidel; Norwell, Mass: 1985.

Barath FT, Chavez MC, Cofield RE, Flower DA, Frerking MA, Gram MB, Harris WM, Holden JR, Jarnot RF, Kloezeman WG, Klose GJ, Lau GK, Loo MS, Maddison BJ, Mattauch RJ, McKlnney RP, Peckham GE, Pickett HM, Siebes G, Soltis FS, Suttie RA, Tarsala JA, Waters JW, Wilson WJ. The Upper Atmosphere Research Satellite microwave limb sounder instrument. J Geophys Res. 1993; 98(D6):10751–10762. DOI: 10.1029/93JD00798

Bhartia PK, McPeters RD, Mateer CL, Flynn LE, Wellemeyer C. Algorithm for the estimation of vertical ozone profiles from the backscattered ultraviolet technique. J Geophys Res-Atmos. 1996; 101:18793–18806.

Bhartia, PK., Wellemeyer, C. TOMS-V8 total O3 algorithm. In: Bhartia, PK., editor. OMI Algorithm Theoretical Basis Document, vol. II, OMI Ozone Products. NASA Goddard Space Flight Cent; Greenbelt, Md: 2002. p. 15-31.(Available at http://eospso.gsfc.nasa.gov/eos_homepage/for_scientists/atbd/index.php)

Bhartia PK, McPeters RD, Flynn LE, Taylor S, Kramarova NA, Frith S, Fisher B, DeLand M. Solar Backscatter UV (SBUV) total ozone and profile algorithm. Atmos Meas Tech. 2013; 6:2533–2548. DOI: 10.5194/amt-6-2533-2013

Bosilovich M, Akella S, Coy L, Cullather R, Draper C, Gelaro R, Kovach R, Liu Q, Molod A, Norris P, Wargan K, Chao W, Reichle R, Takacs L, Vikhliaev Y, Bloom S, Collow A, Firth S, Labow G, Partyka G, Pawson S, Reale O, Schubert SD, Suarez M. MERRA-2: Initial Evaluation of the Climate. NASA Tech Rep Series on Global Modeling and Data Assimilation. 2015; 43 NASA/TM–2015- 104606.

Cohn SE. An Introduction to Estimation Theory. J Met Soc of Japan. 1997; 75(1B):257–288.

Cortesi U, Lambert JC, De Clercq C, Bianchini G, Blumenstock T, Bracher A, Castelli E, Catoire V, Chance KV, De Mazière M, Demoulin P, Godin-Beekmann S, Jones N, Jucks K, Keim C, Kerzenmacher T, Kuellmann H, Kuttippurath J, Iarlori M, Liu GY, Liu Y, McDermid IS, Meijer YJ, Mencaraglia F, Mikuteit S, Oelhaf H, Piccolo C, Pirre M, Raspollini P, Ravegnani F, Reburn WJ, Redaelli G, Remedios JJ, Sembhi H, Smale D, Steck T, Taddei A, Varotsos C, Vigouroux C, Waterfall A, Wetzel G, Wood S. Geophysical validation of MIPAS-ENVISAT operational ozone data. Atmos Chem Phys. 2007; 7:4807–4867. DOI: 10.5194/acp-7-4807-2007

Coy L, Wargan K, McCarty WR, Pawson S, Molod A. Structure and Dynamics of the Quasi-Biennial Oscillation. J Climate. 2015; 29:5339–5354. DOI: 10.1175/JCLI-D-15-0809.1

Damadeo RP, Zawodny JM, Thomason LW, Iyer N. SAGE version 7.0 algorithm: application to SAGE II. Atmos Meas Tech. 2013; 6:3539–3561. DOI: 10.5194/amt-6-3539-2013

Damadeo RP, Zawodny JM, Thomason LW. Reevaluation of stratospheric ozone trends from SAGE II data using a simultaneous temporal and spatial analysis. Atmos Chem Phys. 2014; 14:13455–13470. DOI: 10.5194/acp-14-13455-2014

Daumont M, Brion J, Charbonnier J, Malicet J. Ozone UV spectroscopy, I:Absorption cross-sections at room temperature. J Atmos Chem. 1992; 15:145–155.

DeLand MT, Taylor SL, Huang LK, Fisher BL. Calibration of the SBUV version 8.6 ozone data product. Atmos Meas Tech. 2012; 5:2951–2967. DOI: 10.5194/amt-5-2951-2012

Dethof A, Hólm EV. Ozone assimilation in the ERA-40 reanalysis project. QJR Meteorol Soc. 2004; 130:2851–2872. DOI: 10.1256/qj.03.196

Dragani R. On the quality of the ERA-Interim ozone reanalyses: comparisons with satellite data. Q J R Meteorol Soc. 2011; 137:1312–1326. DOI: 10.1002/qj.821

Fischer H, Birk M, Blom C, Carli B, Carlotti M, von Clarmann T, Delbouille L, Dudhia A, Ehhalt D, Endemann M, Flaud JM, Gessner R, Kleinert A, Koopman R, Langen J, López-Puertas M, Mosner P, Nett H, Oelhaf H, Perron G, Remedios J, Ridolfi M, Stiller G, Zander R. MIPAS: an instrument for atmospheric and climate research. Atmos Chem Phys. 2008; 8:2151–2188. DOI: 10.5194/acp-8-2151-2008

Fleig, AJ., McPeters, RD., Bhartia, PK., Schlesinger, BM., Cebula, RP., Klenck, KF., Taylor, SL., Heath, DF. Ref Publ. NASA; Greenbelt, Md: 1990. Nimbus 7 solar backscatter ultraviolet (SBUV) ozone products user's guide; p. 1234

Frith SM, Kramarova NA, Stolarski RS, McPeters RD, Bhartia PK, Labow GJ. Recent changes in total column ozone based on the SBUV Version 8.6 Merged Ozone Data Set. J Geophys Res Atmos. 2014; 119:9735–9751. DOI: 10.1002/2014JD021889

Froidevaux L, et al. Validation of Aura Microwave Limb Sounder stratospheric ozone measurements. J Geophys Res. 2008; 113:D15S20.doi: 10.1029/2007JD008771

Global Modeling and Assimilation Office (GMAO). MERRA-2 inst3_3d_asm_Nv: 3D IAU State, Meteo- rology Instantaneous 3-hourly (0.625 3 0.5L72), version 5.12.4. Goddard Space Flight Center Distributed Active Archive Center; 2015a. [accessed June 2016]

Global Modeling and Assimilation Office (GMAO). MERRA-2 inst1_2d_asm_Nx: 2d,3-Hourly,Instantaneous,Single-Level,Assimilation,Single-Level Diagnostics V5.12.4, version 5.12.4. Greenbelt, MD, USA: Goddard Earth Sciences Data and Information Services Center (GES DISC); 2015b. [Accessed June 2016]

Herman JR, Hudson R, McPeters R, Stolarski R, Ahmad Z, Gu X-Y, Taylor S, Wellemeyer C. A new self-calibration method applied to TOMS/SBUV backscattered ultraviolet data to determine long-term global ozone change. J Geophys Res. 1991; 96:7531–7545.

Hilsenrath E, Cebula RP, DeLand MT, Laamann K, Taylor S, Wellemeyer C, Bhartia PK. Calibration of the NOAA-11 Solar Backscatter UltraViolet (SBUV/2) ozone data set from 1989 to 1993 using in-flight calibration data and SSBUV. J Geophys Res. 1995; 100(D1):1351–1366. DOI: 10.1029/94JD02611

Huang LK, Cebula RP, Hilsenrath E. New procedure for interpolating NIST FEL lamp irradiances. Metrologia. 1998; 35(4):381–386. DOI: 10.1088/0026-1394/35/4/27

Hubert D, Lambert J-C, Verhoelst T, Granville J, Keppens A, Baray J-L, Bourassa AE, Cortesi U, Degenstein DA, Froidevaux L, Godin-Beekmann S, Hoppel KW, Johnson BJ, Kyrölä E, Leblanc T, Lichtenberg G, Marchand M, McElroy CT, Murtagh D, Nakane H, Portafaix T, Querel R, Russell JM III, Salvador J, Smit HGJ, Stebel K, Steinbrecht W, Strawbridge KB, Stübi R, Swart DPJ, Taha G, Tarasick DW, Thompson AM, Urban J, van Gijsel JAE, Van Malderen R, von der Gathen P, Walker KA, Wolfram E, Zawodny JM. Ground-based assessment of the bias and long-term stability of 14 limb and occultation ozone profile data records. Atmos Meas Tech. 2016; 9:2497–2534. DOI: 10.5194/amt-9-2497-2016\

Inness A, Baier F, Benedetti A, Bouarar I, Chabrillat S, Clark H, Clerbaux C, Coheur P, Engelen RJ, Errera Q, Flemming J, George M, Granier C, Hadji-Lazaro J, Huijnen V, Hurtmans D, Jones L, Kaiser JW, Kapsomenakis J, Lefever K, Leitão J, Razinger M, Richter A, Schultz MG, Simmons AJ, Suttie M, Stein O, Thépaut J-Nn, Thouret V, Vrekoussis M, Zerefos C. the MACC team. The MACC reanalysis: an 8 yr data set of atmospheric composition. Atmos Chem Phys. 2013; 13:4073–4109. DOI: 10.5194/acp-13-4073-2013

Inness A, Blechschmidt A-M, Bouarar I, Chabrillat S, Crepulja M, Engelen RJ, Eskes H, Flemming J, Gaudel A, Hendrick F, Huijnen V, Jones L, Kapsomenakis J, Katragkou E, Keppens A, Langerock B, de Mazière M, Melas D, Parrington M, Peuch VH, Razinger M, Richter A, Schultz MG, Suttie M, Thouret V, Vrekoussis M, Wagner A, Zerefos C. Data assimilation of satellite-retrieved ozone, carbon monoxide and nitrogen dioxide with ECMWF's Composition-IFS. Atmos Chem Phys. 2015; 15:5275–5303. DOI: 10.5194/acp-15-5275-2015

Janz S, Hilsenrath E, Butler J, Heath DF, Cebula RP. Uncertainties in radiance calibrations of backscatter ultraviolet (BUV) instruments. Metrologia. 1995; 32(6):637–641. DOI: 10.1088/0026-1394/32/6/48

Kalnay, E. Atmospheric Modeling, Data Assimilation and Predictability. Cambridge Univ. Press; 2003. p. 1063

Kramarova N, Bhartia PK, Frith S, McPeters R, Stolarski R. Interpreting SBUV smoothing errors: An example using the quasi-biennial oscillation. Atmos Meas Tech Discuss. 2013; 6:2721–2749. DOI: 10.5194/amtd-6-2721-2013

Kyrölä E, Laine M, Sofieva V, Tamminen J, Päivärinta S-M, Tukiainen S, Zawodny J, Thomason L. Combined SAGE II–GOMOS ozone profile data set for 1984–2011 and trend analysis of the vertical distribution of ozone. Atmos Chem Phys. 2013; 13:10645–10658. DOI: 10.5194/acp-13-10645-2013

Labow GJ, McPeters RD, Bhartia PK, Kramarova N. A comparison of 40 years of SBUV measurement of column ozone with data from the dosben /brewer network. J Geophys Res Atoms. 2013; 118:7370–7378. DOI: 10.1002/jgrd.50503

Lahoz WA, Errera Q, Swinbank R, Fonteyn D. Data assimilation of stratospheric constituents: a review. Atmos Chem Phys. 2007; 7:5745–5773. DOI: 10.5194/acp-7-5745-2007

Lahoz WA, Schneider P. Data Assimilation: Making sense of Earth observations. Front Environ Sci. 2014 May 28.doi: 10.3389/fenvs.2014.00016

Levelt PF, van den Oord GHJ, Dobber MR, Mälkki A, Visser H, Vries JD, Stammes P, Lundell JOV, Saari H. The Ozone Monitoring Instrument. IEEE Trans Geosci Remote Sens. 2006; 44:1093–1101. DOI: 10.1109/TGRS.2006.872333

Lefever K, van der RA, Baier F, Christophe Y, Errera Q, Eskes H, Flemming J, Inness A, Jones L, Lambert J-C, Langerock B, Schultz MG, Stein O, Wagner A, Chabrillat S. Copernicus stratospheric ozone service, 2009–2012: validation, system intercomparison and roles of input data sets. Atmos Chem Phys. 2015; 15:2269–2293. DOI: 10.5194/acp-15-2269-2015

Livesey NJ, Read WG, Froidevaux L, Waters JW, Santee ML, Pumphrey HC, Wu DL, Shippony Z, Jarnot RF. The UARS Microwave Limb Sounder version 5 data set: Theory, characterization, and validation. J Geophys Res. 2003; 108:4378.doi: 10.1029/2002JD002273,D13

Livesey, NJ., Read, WG., Lambert, A., Cofield, RE., Cuddy, DT., Froidevaux, L., Fuller, RA., Jarnot, RF., Jiang, JH., Jiang, YB., Knosp, BW., Kovalenko, LJ., Pickett, HM., Pumphrey, HC., Santee, ML., Schwartz, MJ., Stek, PC., Wagner, PA., Waters, JW., Wu, DL. Version 2.2x Level 2 Data Quality and Description Document. 2007. JPL D-33509. Available from https://mls.jpl.nasa.gov/data/v2-2_data_quality_document.pdf

Livesey, NJ., Read, WG., Wagner, PA., Froidevaux, L., Lambert, A., Manney, G., Valle, LFM., Pumphrey, HC., Santee, ML., Schwartz, MJ., Wang, S., Fuller, RA., Jarnot, RF., Knosp, BW., Martinez, E. Version 4.2x Level 2 Data Quality and Description Document. 2015. JPL D-33509 Rev. A. Available from https://mls.jpl.nasa.gov/data/v4-2_data_quality_document.pdf

McCarty W, Coy L, Gelaro R, Merkova D, Smith EB, Sienkiewicz M, Wargan K. MERRA-2 Input Observations: Summary and Assessment. NASA/TM. 2016 in preparation.

McIntyre ME, Palmer TN. The 'surf zone' in the stratosphere. J Atmos Terr Phys. 1984; 46:825–849.

McPeters RD, Labow GJ. An assessment of the accuracy of 14.5 years of Nimbus 7 TOMS Version 7 ozone data by comparison with the Dobson Network. Geophys Res Lett. 1996; 23:3695–3698.

McPeters RD, et al. Validation of the Aura Ozone Monitoring Instrument total column ozone product. J Geophys Res. 2008; 113:D15S14.doi: 10.1029/2007JD008802

McPeters RD, Frith S, Labow GJ. OMI total column ozone: extending the long-term data record. Atmos Meas Tech. 2015; 8:4845–4850. DOI: 10.5194/amt-8-4845-2015

Molod A, Takacs L, Suarez M, Bacmeister J. Development of the GEOS-5 atmospheric general circulation model: Evolution from MERRA to MERRA2. Geosci Model Dev. 2015; 8:1339–1356.

Putman WM, Lin S-J. Finite-volume transport on various cubed-sphere grids. Journal of Computational Physics. 2007; 227(1):55–78. DOI: 10.1016/j.jcp.2007.07.022

Raspollini P, Carli B, Carlotti M, Ceccherini C, Dehn A, Dinelli BM, Dudhia A, Flaud J-M, López-Puertas M, Niro F, Remedios JJ, Ridolfi M, Sembhi H, Sgheri L, von Clarmann T. Ten years of MIPAS measurements with ESA Level 2 processor V6 – Part 1: Retrieval algorithm and diagnostics of the products. Atmos Meas Tech. 2013; 6:2419–2439. DOI: 10.5194/amt-6-2419-2013

Sakazaki T, Shiotani M, Suzuki M, Kinnison D, Zawodny JM, McHug M, Walker KA. Sunset–sunrise difference in solar occultation ozone measurements (SAGE II, HALOE, and ACE–FTS) and its relationship to tidal vertical winds. Atmos Chem Phys. 2015; 15:829–843. DOI: 10.5194/acp-15-829-2015

Sofieva VF, Tamminen J, Kyrölä E, Mielonen T, Veefkind P, Hassler B, Bodeker GE. A novel tropopause-related climatology of ozone profiles. Atmos Chem Phys. 2014; 14:283–299. DOI: 10.5194/acp-14-283-2014

Stajner I, et al. Assimilated ozone from EOS-Aura: Evaluation of the tropopause region and tropospheric columns. J Geophys Res. 2008; 113:D16S32.doi: 10.1029/2007JD008863

Takacs LL, Suárez MJ, Todling R. Maintaining atmospheric mass and water balance in reanalyses. Q J Roy Meteor Soc. 2016; 142:1565–1573. DOI: 10.1002/qj.2763

Thompson AM, et al. Southern Hemisphere Additional Ozonesondes (SHADOZ) 1998–2000 tropical ozone climatology 1. Comparison with Total Ozone Mapping Spectrometer (TOMS) and ground-based measurements. J Geophys Res. 2003a; 108:8238.doi: 10.1029/2001JD000967,D2

Thompson AM, et al. Southern Hemisphere Additional Ozonesondes (SHADOZ) 1998–2000 tropical ozone climatology 2. Tropospheric variability and the zonal wave-one. J Geophy Res. 2003b; 108:8241.doi: 10.1029/2002JD002241,2003

Torres O, Bhartia PK. Effect of Stratospheric Aerosol on Ozone Profile from BUV Measurements. Geophys Res Lett. 1995; 22:235–238.

Torres O, Herman JR, Bhartia PK, Ahmad Z. Properties of Mount-Pinatubo Aerosols as Derived from Nimbus-7 Total Ozone Mapping Spectrometer Measurements. J Geophys Res-Atmos. 1995; 100:14043–14055.

Vasilkov AP, Joiner J, Yang K, Bhartia PK. Improving total column ozone retrievals by using cloud pressures derived from Raman scattering in the UV. Geophysical Research Letters. 2004; 31:L20109.doi: 10.1029/2004GL020603

Wang HJ, Cunnold DM, Thomason LW, Zawodny JM, Bodeker GE. Assessment of SAGE version 6.1 ozone data quality. J Geophys Res. 2002; 107(D23):4691.doi: 10.1029/2002JD002418

Wargan K, Pawson S, Olsen MA, Witte JC, Douglass AR, Ziemke JR, Strahan SE, Nielsen JE. The global structure of upper troposphere- lower stratosphere ozone in GEOS-5: A multiyear assimilation of EOS Aura data. J Geophys Res Atmos. 2015; 120:2013–2036. DOI: 10.1002/2014JD022493

Waters JW, et al. The Earth Observing System Microwave Limb Sounder (EOS MLS) on the Aura satellite. IEEE Trans Geosci Remote Sens. 2006; 44:1075–1092.

Wellemeyer, CG., Bhartia, PK., Taylor, SL., Qin, W., Ahn, C. Version 8 Total Ozone Mapping Spectrometer (TOMS) Algorithm. Eur Comm; paper presented at Quadrennial Ozone Symposium; Kos, Greece: 2004.

World Meteorological Organization. WMO. Global Ozone Research and Monitoring Project – Report no. 55. 2014. Scientific Assessment of Ozone Depletion: 2014.

Ziemke JR, Chandra S, Labow GJ, Bhartia PK, Froidevaux L, Witte JC. A global climatology of tropospheric and stratospheric ozone derived from Aura OMI and MLS measurements. Atmos Chem Phys. 2011; 11:9237–9251. DOI: 10.5194/acp-11-9237-2011

Ziemke JR, et al. Assessment and applications of NASA ozone data products derived from Aura OMI/MLS satellite measurements in context of the GMI chemical transport model. J Geophys Res Atmos. 2014; :119.doi: 10.1002/2013JD020914
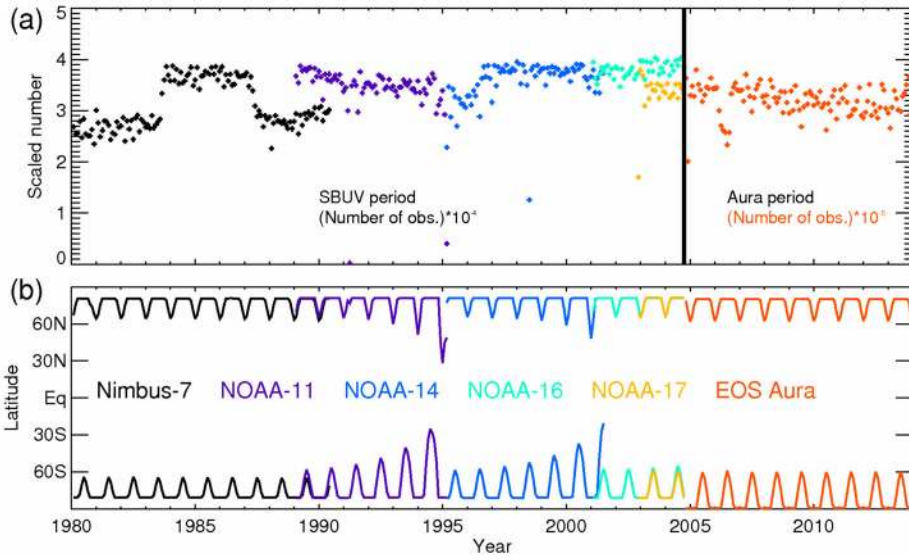
Figure 1.
The total column ozone observations assimilated in MERRA-2. (a) the monthly averaged number of total ozone observations per day from SBUV and OMI instruments as a function of time. The SBUV and Aura periods are separated by a thick vertical divider. Note that the daily numbers for SBUV and OMI are given in the units of 10,000 and 100,000, respectively. (b) the monthly latitudinal coverage of the total ozone observations. Colors denote different instruments as indicated in the legend in panel b.
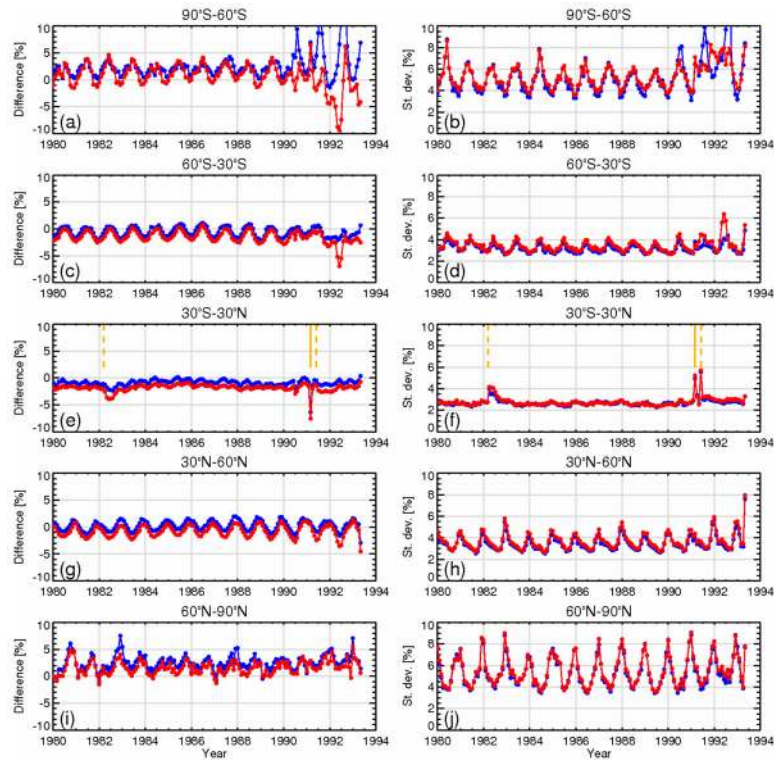
Figure 2.
Comparisons of the monthly statistics total ozone from MERRA (blue) and MERRA-2 (red) with TOMS data in 5 latitude bands: (a, b) 90°S-60°S, (c, d) 60°S-30°S, (e, f) 30°S-30°N, (g, h) 30°N-60°N, and (i, j) 60°N-90°N. The left column (panels a, c, e, g and i) shows time series of reanalysis minus TOMS differences expressed as percentages of TOMS monthly averages. The right panels (b, d, f, h, and j) plot the corresponding standard deviations of reanalysis minus TOMS differences relative to TOMS. The times of El Chichón (March-April 1982) and Mt. Pinatubo (June 1991) eruptions are marked by yellow dashed lines in the central panels (e and f). The only full month of missing ozone data (March 1991) is indicated by solid yellow lines in the same panels.
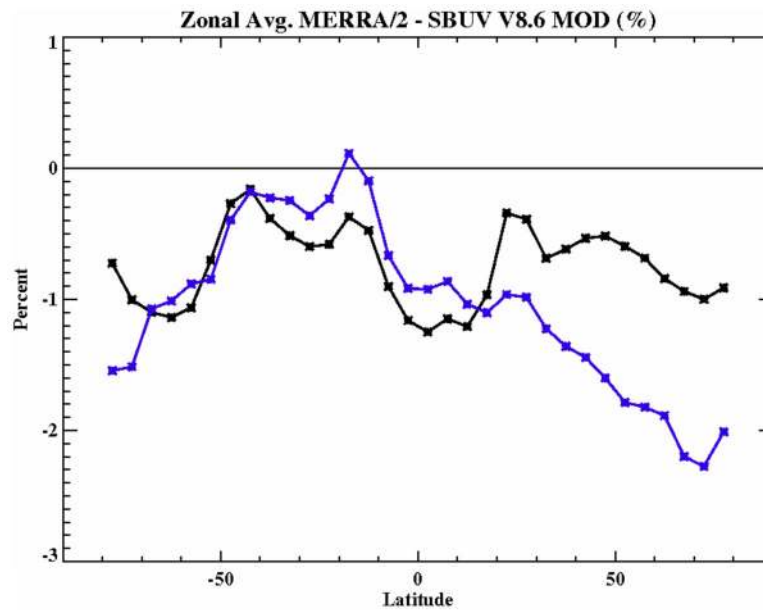
Figure 3.
Comparisons of the MERRA-2 total ozone against the Merged Ozone Data Set (MOD). The plot shows the reanalysis minus MOD difference as a percentage of MOD for two periods: 1980–2003 (black) and 2005–2014 (blue).
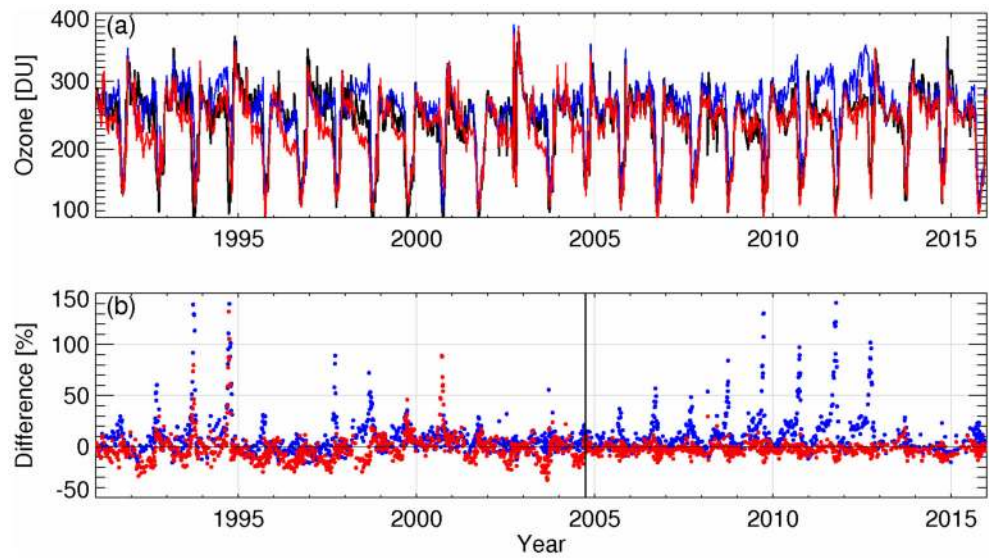
Figure 4.
(a) The 1991–2015 time series of total ozone in Dobson units at the South Pole derived from individual ozonesonde measurements (black) and collocated reanalysis: MERRA (blue) and MERRA-2 (red). (b) reanalysis minus ozonesonde differences divided by sonde total ozone and expressed in percent: MERRA (blue) and MERRA-2 (red). The SBUV and Aura periods are separated by the black vertical line.
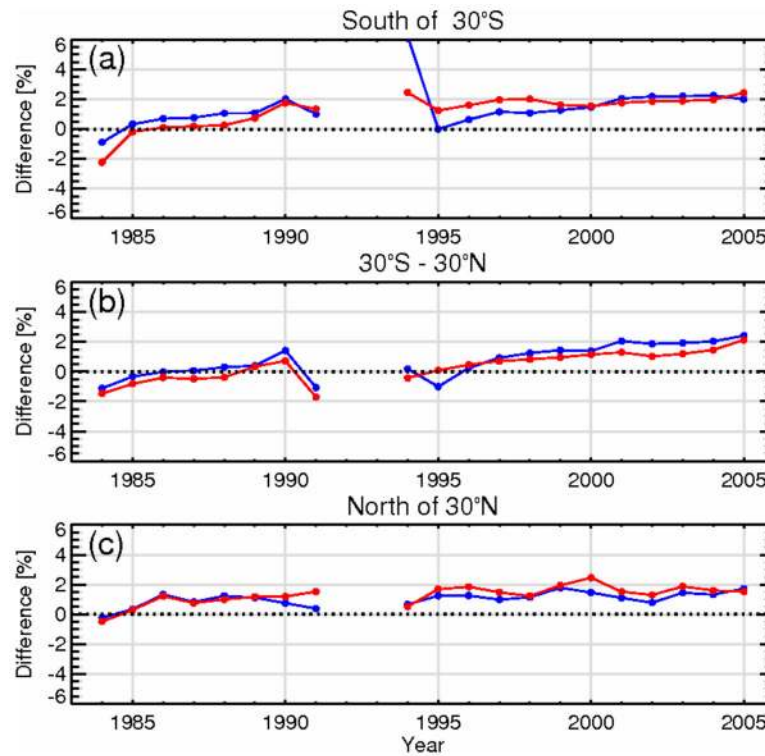
Figure 5.
Time series of annually averaged ozone differences integrated between 208 hPa and 0.2 hPa: MERRA minus SAGE II (blue) and MERRA-2 minus SAGE II (red). The differences are expressed as percent of the SAGE II annual means. The averaging is done in three latitude bands: south of 30°S (a), between 30°S and 30°N (b) and north of 30°N (c). The years 1992 and 1993 are omitted because of a lower quality of SAGE II data in the lower stratosphere following the Mt. Pinatubo eruption.
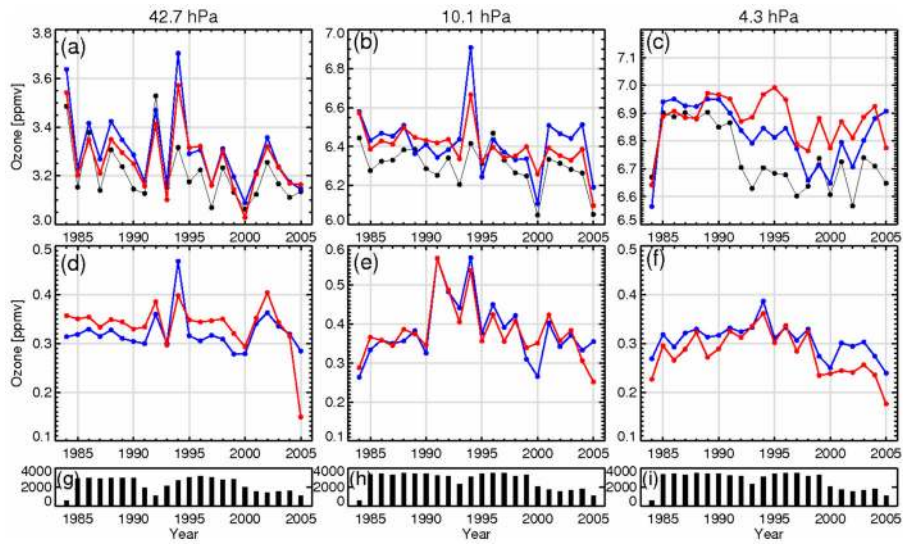
Figure 6.
Time series of annually averaged statistics from SAGE II, MERRA and MERRA-2 south of 30°S at 42.7 hPa (a, d and g), 10.1 hPa (b, e and h) and 4.3 hPa (c, f and i). The top row (a–c): the annual mean mixing ratio from SAGE II (black), MERRA (blue) and MERRA-2 (red). The middle row (d–f): the standard deviation of the MERRA minus SAGE II (blue) and MERRA-2 minus SAGE II (red) differences. The total annual SAGE II data counts for each level are shown in the bottom row (g–i). The units are parts per million by volume
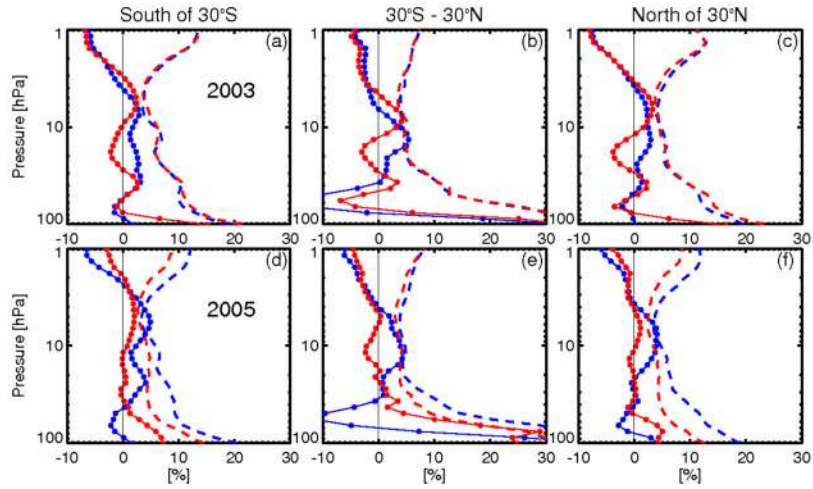
Figure 7.
Profiles of MERRA – SAGE II (blue) and MERRA-2 – SAGE II (red) comparisons relative to the SAGE II average calculated for all available data in January-August 2003 (a–c) and 2005 (d–f) and expressed in percent. Statistics for three latitude bands are shown: south of 30°S (a and d), between 30°S and 30°N (b and e) and north of 30°N (c and f). The solid lines represent the mean differences between the reanalysis and SAGE II. The standard deviations of the differences are shown as dashed lines. The SAGE II observation levels are indicated by dots superimposed on the solid lines. All statistics are shown as a percentage of the SAGE II mean over 2003 (top row) or 2005 (bottom).
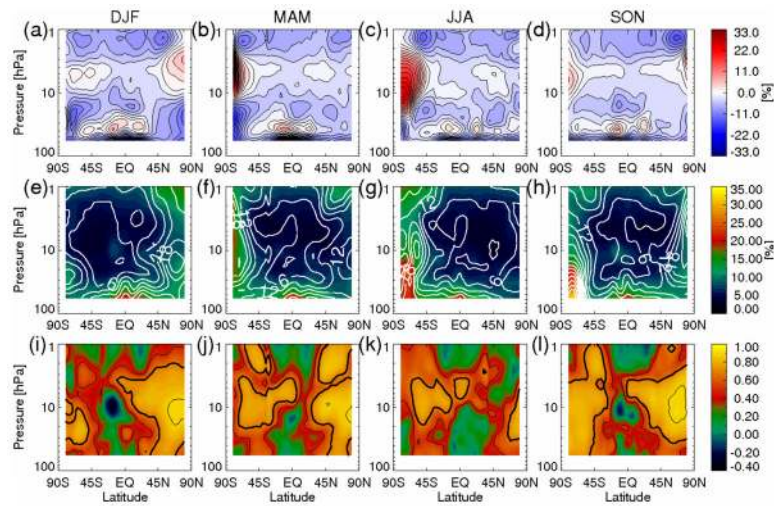
Figure 8.
Top row: the zonal mean difference between MERRA-2 and UARS MLS ozone computed from all UARS MLS data between 1991 and 1996 for December – February (a), March-May (b), June-August (c) and September-November (d). The difference is expressed as a percent of the zonal mean UARS MLS ozone. Middle row: The standard deviation of the difference between the two data sets, relative to the zonal mean UARS MLS observations (colors) and relative standard deviation of UARS MLS ozone (white contours) for December – February (e), March-May (f), June-August (g) and September-November (h). Bottom row: correlations between MERRA-2 and UARS MLS for December – February (i), March-May (j), June-August (k) and September-November (l). The black contours are 0.5, 0.7 and 0.9, the 0.7 contour is thickened. The comparisons are done in the 57 hPa – 1 hPa range.
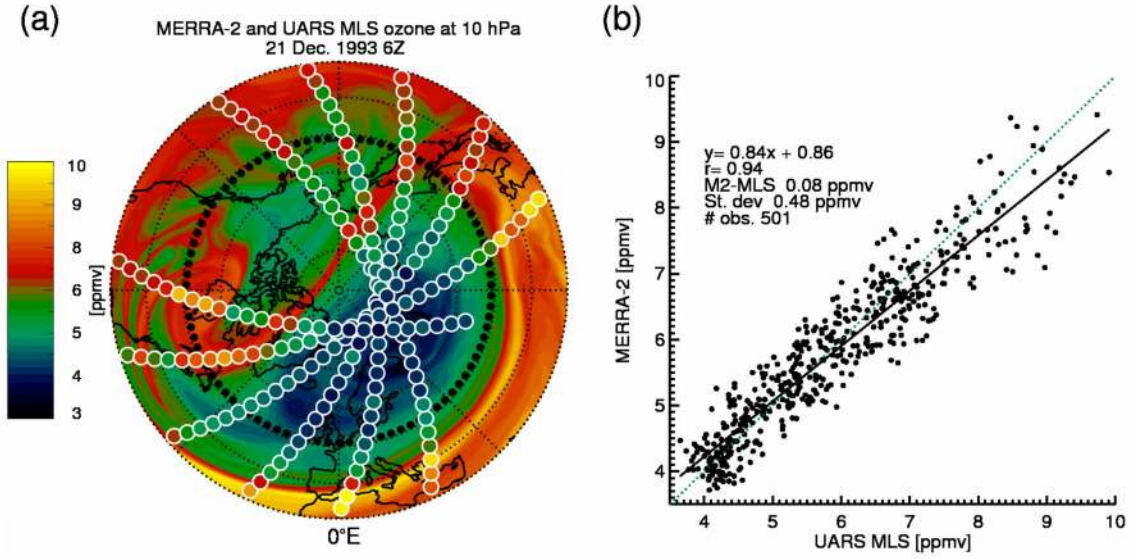
Figure 9.
(a) Color background: MERRA-2 ozone between 30°N and 90°N at 10 hPa on 21 December 1993 at 6Z. UARS MLS observations made between 1:30Z and 10:30Z are shown as filled circles color-coded by the retrieved mixing ratios. The color contour spacing is 0.2 ppmv. The highest latitude observed by SBUV around the same day (51°N) is marked by black dots. (b) a scatter plot of all UARS MLS observations at 10 hPa on 21 December and collocated MERRA-2 ozone the closest in time to the MLS data (no more than 1.5 h apart). The diagonal is shown as a dotted line and the solid black line is the linear fit. The units are parts per million by volume.
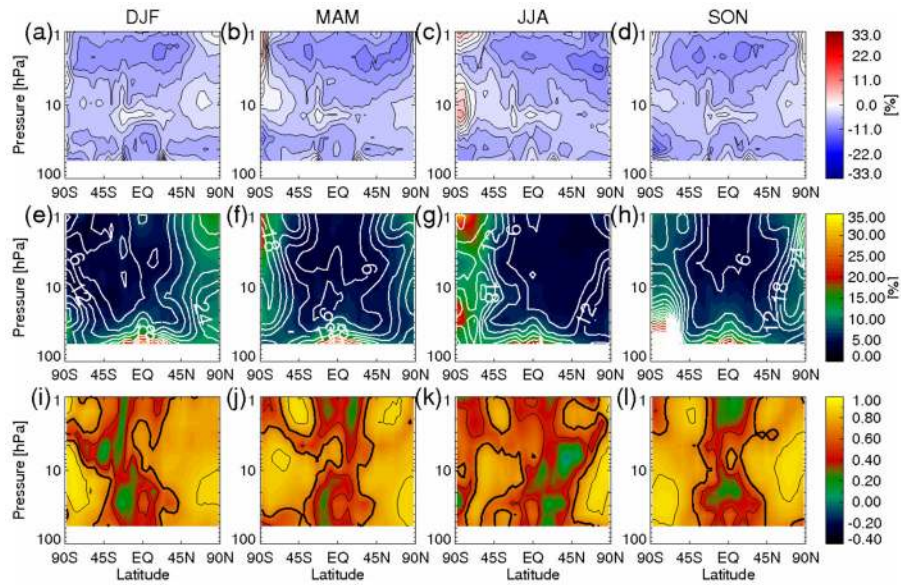
Figure 10.
Top row: the zonal mean difference between MERRA-2 and MIPAS ozone computed from all MIPAS data between 2003 and 2012 for December – February (a), March-May (b), June-August (c) and September-November (d). The difference is expressed as a percent of the zonal mean MIPAS ozone. Middle row: The standard deviation of the difference between the two data sets, relative to the zonal mean MIPAS observations (colors) and relative standard deviation of MIPAS ozone (white contours) for December – February (e), March-May (f), June-August (g) and September-November (h). Bottom row: Correlations between MERRA-2 and MIPAS data for December – February (i), March-May (j), June-August (k) and September-November (l). The black contours are 0.5, 0.7 and 0.9, the 0.7 contour is thickened. The comparisons are done in the 57 hPa – 1 hPa range.
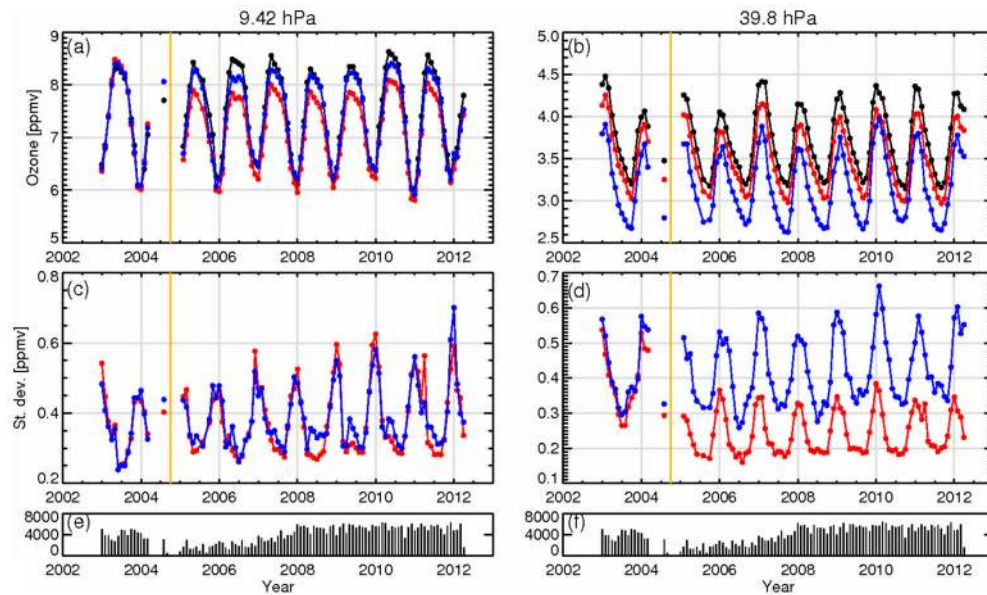
Figure 11.

Time series of MIPAS, MERRA, and MERRA-2 monthly ozone statistics. 30°N – 60°N. Top row: the monthly average from MIPAS (black), MERRA (blue) and MERRA-2 (red) at 9.42 hPa (a) and 39.8 hPa (b). Middle row: the reanalysis-MIPAS difference standard deviation at 9.42 hPa (c) and 39.8 hPa (d). The units are parts per million by volume. Bottom row: The corresponding monthly MIPAS observation counts at 9.42 hPa (e) and 39.8 hPa (f). The yellow lines in panels a–d mark the transition from the SBUV to EOS Aura period in MERRA-2.
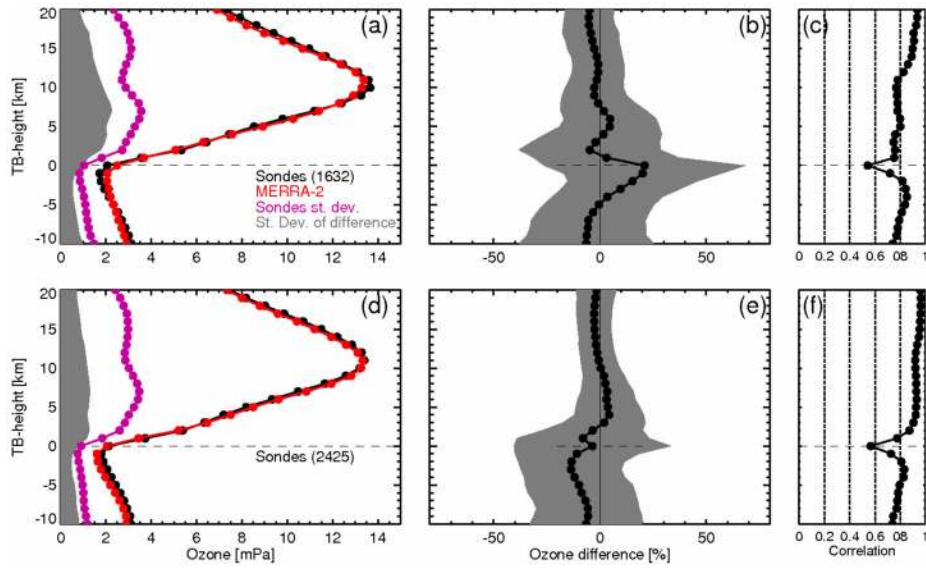
Figure 12.
Statistical comparisons of MERRA-2 ozone against ozonesondes available in 2003 (a–c) and 2005 (d–f). All data are interpolated to the tropopause-based vertical coordinate. (a and d) the average ozone partial pressure from ozone sondes (black) and MERRA-2 (red), the standard deviation of ozonesonde measurements (magenta), and the standard deviation of MERRA-2 minus ozonesonde differences (shaded). (b and e) The mean MERRA-2 minus ozonesonde differences (black line) and the standard deviation of the difference (shaded), all expressed as a percent of sonde ozone. (c and f) the MERRA-2 – ozonesonde correlations. The filled circles are the levels to which all data are interpolated and the dashed line is the dynamical tropopause. The units in a and d are millipascals.
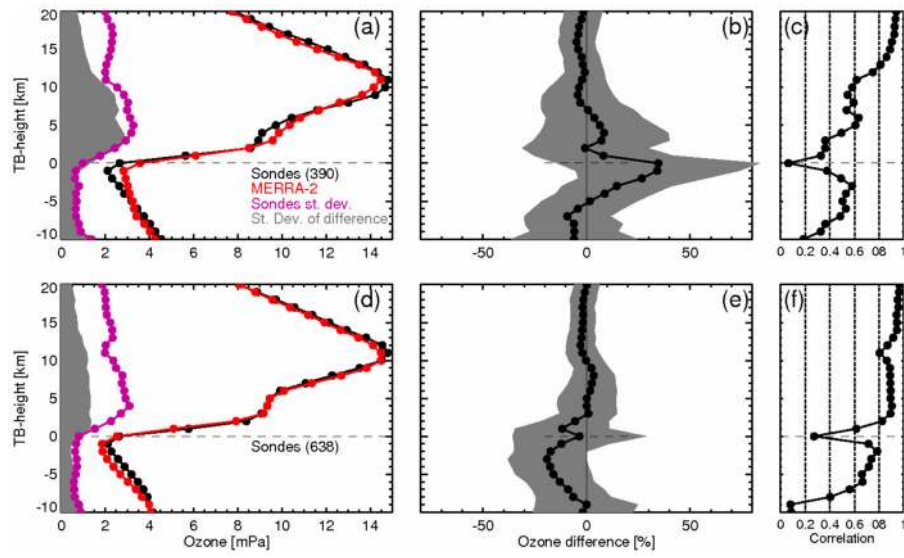
Figure 13.
The same as Figure 13 but with soundings limited to 0°E-60°E, 45°N-60°N, March-May.
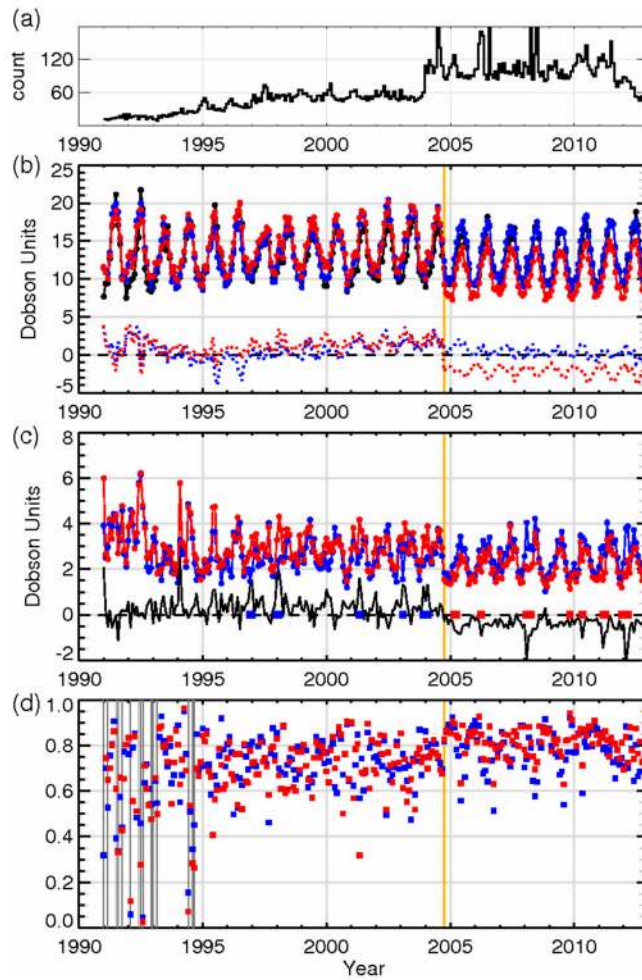
Figure 14.
Time series of various statistics comparing ozonesondes between 30°N and 60°N (black), MERRA (blue) and MERRA-2 (red) integrated in the upper troposphere (500 hPa to the dynamical tropopause). (a) number of sondes in each month. (b) the monthly averaged ozone (solid lines with dots) and the reanalysis minus sonde differences in Dobson units (dotted lines). (c) the standard deviation of the reanalysis minus ozonesonde differences (blue and red for MERRA and MERRA-2, respectively) and their difference (MERRA-2 standard deviation, minus MERRA standard deviation,, black) in Dobson units. The red (blue) squares on the zero line indicate months when MERRA-2 (MERRA) is closer to the sondes than MERRA (MERRA-2) at the 95% significance level. (d) monthly reanalysis-ozonesonde correlations. The grey bars indicate the months for which the correlations are not statistically significant at 95%. The vertical yellow lines indicate October 2004 when Aura data were introduced in MERRA-2.
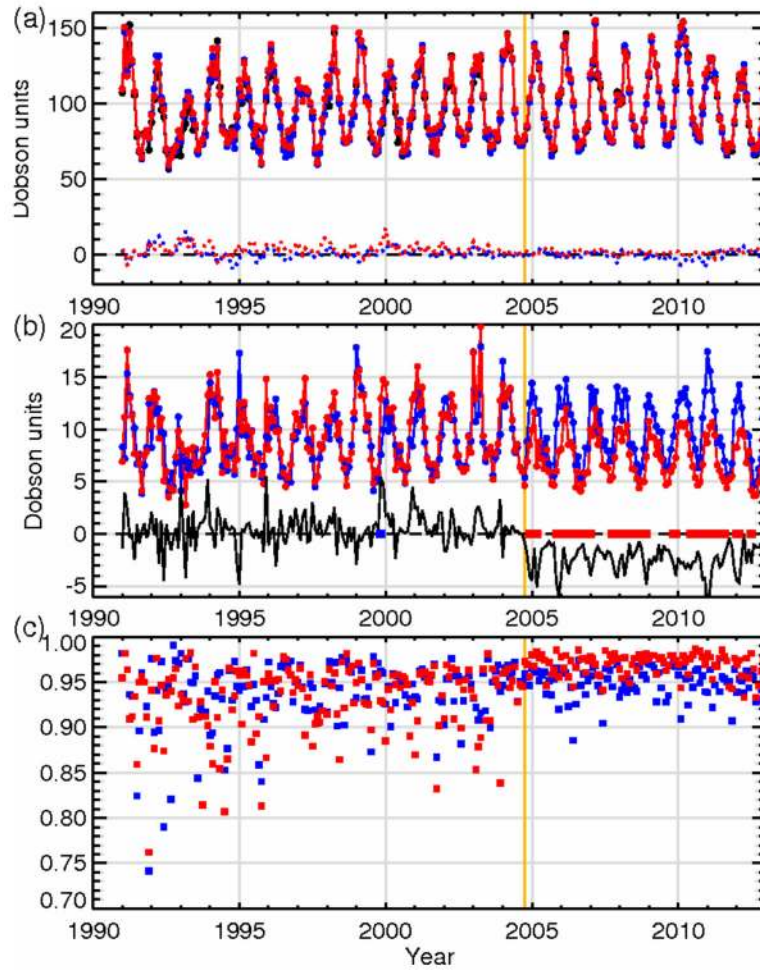
Figure 15.
Time series of various statistics comparing ozonesondes between 30°N and 60°N (black), MERRA (blue) and MERRA-2 (red) integrated in the lower stratosphere (between the dynamical tropopause and 50 hPa). (a) the monthly averaged ozone (solid lines with dots) and the reanalysis minus sonde differences in Dobson units (dotted lines). (b) the standard deviation of the reanalysis minus ozonesonde differences (blue and red for MERRA and MERRA-2, respectively) and their difference (MERRA-2 standard deviation. minus MERRA standard deviation, black) in Dobson units. The red (blue) squares on the zero line indicate months when MERRA-2 (MERRA) is closer to the sondes than MERRA (MERRA-2) at the 95% significance level. (c) monthly reanalysis-ozonesonde correlations. The vertical yellow lines indicate October 2004 when Aura data were introduced in MERRA-2.
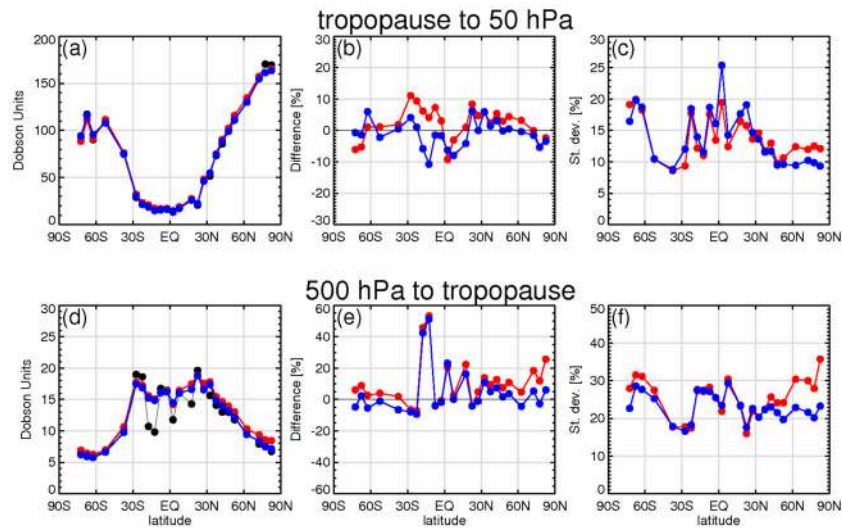
Figure 16.
Comparisons of LS (a-c) and UT (d–f) ozone from MERRA (blue) and MERRA-2 (red) with ozonesondes (black) as a function of latitude. All available sondes between 1991 and 2003 are used. The data are binned into 36 latitude bins, each 5° wide. Shown are the average in Dobson units (a and d), the reanalysis minus sonde difference relative to the sonde mean in percent (b and e) and the standard deviation of the difference also relative to the sonde average in percent (c and f).
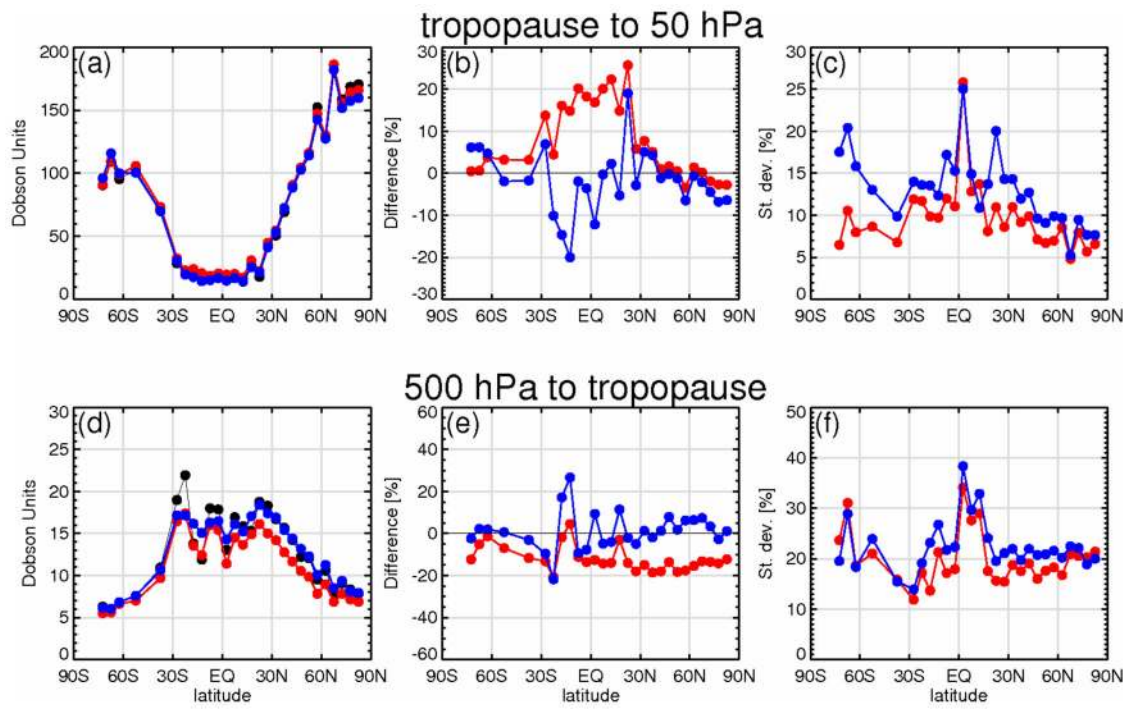
Figure 17.
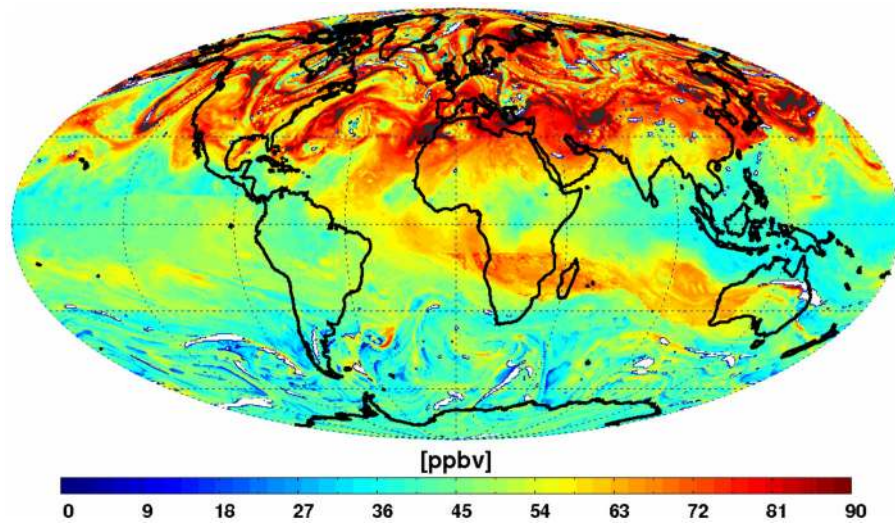As in Figure 16 but for the 2005–2012 period.

Figure 18.
Mean ozone mixing ratio in parts per billion by volume between 500 hPa and the tropopause from MERRA-2 on 1 July 2013. The areas where the tropopause lies below the 500-hPa surface are marked white.

Table 1

Ozone data assimilated in MERRA-2

Instrument and platform	Time period	Type of measurement	Coverage
SBUV Nimbus-7	1 Jan 1980 – 31 May 1990	Partial and total columns	Sunlit atmosphere, surface to TOA
SBUV/2 NOAA-11	1 Jan 1989 – 4 Feb 1995	Partial and total columns	Sunlit atmosphere, surface to TOA
SBUV/2 NOAA-14	5 Feb 1995 – 30 Jun 2001	Partial and total columns	Sunlit atmosphere, surface to TOA
SBUV/2 NOAA-16	1 Jan 2001 – 30 Sep 2004	Partial and total columns	Sunlit atmosphere, surface to TOA
SBUV/2 NOAA-17	17 Nov 2002 – 30 Sep 2004	Partial and total columns	Sunlit atmosphere, surface to TOA
OMI EOS Aura	1 Oct 2004 - present	Total column	Sunlit atmosphere
MLS EOS Aura	1 Oct 2004 - present	Mixing ratio profiles	82°S – 82°N mainly stratosphere

Table 2

Statistical comparisons of the MERRA-2 total ozone against TOMS data. The percent values are relative to TOMS averages.

	Bias		Error standard deviation	
	[DU]	[%]	[DU]	[%]
90°S-60°S	1.80	0.63	15.15	5.32
60°S-30°S	−4.49	−1.43	10.84	3.46
30°S-30°N	−4.92	−1.83	7.44	2.77
30°N-60°N	−2.60	−0.77	12.58	3.72
60°N-90°N	5.09	1.45	19.35	5.53

Table 3

Statistical comparisons of the MERRA-2 stratospheric ozone against SAGE II data south of 30°S. Results obtained using sunset only data are shown in parentheses. The percent values are relative to SAGE II averages.

	Bias		Error standard deviation	
	[ppmv]	**[%]**	**[ppmv]**	**[%]**
4.3 hPa	0.14 (0.06)	2 (0.9)	0.31 (0.3)	4.6 (4.3)
10.1 hPa	0.09 (0.07)	1.5 (1.1)	0.4 (0.42)	6.4 (6.5)
42.7 hPa	0.06 (0.07)	1.9 (2)	0.35 (0.35)	11 (11)

Table 4

Statistical comparisons of the MERRA-2 stratospheric ozone against SAGE II data north of 30°N. Results obtained using sunset only data are shown in parentheses. The percent values are relative to SAGE II averages.

	Bias		Error standard deviation	
	[ppmv]	[%]	[ppmv]	[%]
4.3 hPa	0.07 (0.02)	1.1 (0.3)	0.32 (0.31)	4.6 (4.5)
10.1 hPa	0.08 (0.06)	1.3 (1.0)	0.37 (0.37)	5.7 (5.8)
42.7 hPa	0.05 (0.05)	1.6 (1.6)	0.37 (0.37)	11.3 (11.3)

Table 5

Statistical comparisons of the MERRA-2 UT and LS ozone against ozonesonde measurements between 30°N and 60°N. The percent values are relative to ozonesondes averages.

	Bias				Error standard deviation				Correlation	
	SBUV period		Aura period		SBUV period		Aura period		SBUV period	Aura period
	[DU]	[%]	[DU]	[%]	[DU]	[%]	[DU]	[%]		
Upper troposphere	1.22	9.5	−1.77	−13.6	3.15	24.5	2.74	21.1	0.81	0.9
Lower stratosphere	3.65	3.8	1.23	1.2	10.62	11.2	8.27	8.1	0.96	0.98

NASA Author Manuscript

Table 6

Discontinuities in the ozone observing system.

Time /period	Source of discontinuity
1–31 March 1991	Missing NOAA-11 SBUV observations; no ozone data are assimilated
Late 1994	SBUV coverage limited to latitudes north of 30°S
1 October 2004	Introduction of EOS Aura data; SBUV data turned off
27 March – 18 April 2011	MLS data outage
1 June 2015	Transition from Version 2.2 to Version 4.2 of MLS retrievals
1 May 2016	The 261 hPa MLS level turned off
29 May – 13 June 2016	OMI data outage (not discussed in this study but included here for completeness)

NASA Author Manuscript

NASA Author Manuscript