



Assessment of Odin-OSIRIS ozone measurements from 2001 to the present using MLS, GOMOS, and ozonesondes

C. Adams^{1,*}, A. E. Bourassa¹, V. Sofieva², L. Froidevaux³, C. A. McLinden^{1,4}, D. Hubert⁵, J.-C. Lambert⁵, C. E. Sioris¹, and D. A. Degenstein¹

¹Institute of Space and Atmospheric Studies, University of Saskatchewan, Saskatoon, Saskatchewan, Canada

²Finnish Meteorological Institute, Helsinki, Finland

³Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA

⁴Environment Canada, Downsview, Ontario, Canada

⁵Belgian Institute for Space Aeronomy, Brussels, Belgium

* now at: Alberta Environment and Sustainable Resource Development, Edmonton, Alberta, Canada

Correspondence to: C. Adams (cristenlfadams@gmail.com)

Received: 1 March 2013 – Published in Atmos. Meas. Tech. Discuss.: 17 April 2013

Revised: 11 October 2013 – Accepted: 24 November 2013 – Published: 8 January 2014

Abstract. The Optical Spectrograph and InfraRed Imaging System (OSIRIS) was launched aboard the Odin satellite in 2001 and is continuing to take limb-scattered sunlight measurements of the atmosphere. This work aims to characterize and assess the stability of the OSIRIS 11 yr v5.0x ozone data set. Three validation data sets were used: the v2.2 Microwave Limb Sounder (MLS) and v6 Global Ozone Monitoring by Occultation of Stars (GOMOS) satellite data records, and ozonesonde measurements. Global mean percent differences between coincident OSIRIS and validation measurements are within 5 % at all altitudes above 18.5 km for MLS, above 21.5 km for GOMOS, and above 17.5 km for ozonesondes. Below 17.5 km, OSIRIS measurements agree with ozonesondes within 5 % and are well-correlated ($R > 0.75$) with them. For low OSIRIS optics temperatures ($< 16^\circ\text{C}$), OSIRIS ozone measurements have a negative bias of 1–6 % compared with the validation data sets for 25.5–40.5 km. Biases between OSIRIS ascending and descending node measurements were investigated and found to be related to aerosol retrievals below 27.5 km. Above 30 km, agreement between OSIRIS and the validation data sets was related to the OSIRIS retrieved albedo, which measures apparent upwelling, with a positive bias in OSIRIS data with large albedos. In order to assess the long-term stability of OSIRIS measurements, global average drifts relative to the validation data sets were calculated and were found to be $< 3\%$ per decade for comparisons with MLS for 19.5–36.5 km, GOMOS for

18.5–54.5 km, and ozonesondes for 12.5–22.5 km. Above 36.5 km, the relative drift for OSIRIS versus MLS ranged from ~ 0 to 6 % per decade, depending on the data set used to convert MLS data to the OSIRIS altitude versus number density grid. Overall, this work demonstrates that the OSIRIS 11 yr ozone data set from 2001 to the present is suitable for trend studies.

1 Introduction

In order to assess trends in stratospheric ozone, consistency within long-term data records is essential. Small drifts in a time series, caused by, for example, instrument degradation or changes to a satellite orbit, can have a large effect on trend calculations. The SI²N (SPARC – Stratospheric Processes and their Role in Climate, IO₃C – International Ozone Commission, IGACO-O3 – Integrated Global Atmospheric Chemistry Observations, NDACC – Network for the Detection of Atmospheric Composition Change) initiative aims to compile short-term satellite, long-term satellite, and ground-based ozone measurements in a consistent manner in order to assess current and past changes in the vertical distribution of ozone (SI²N, 2012). The European Space Agency (ESA) Ozone Climate Change Initiative (Ozone_cci), which is a major contributor to SI²N, aims to create a merged ozone data set from satellite measurements meeting quality

requirements of climate change research. Trends in stratospheric ozone since the 1980s vary across latitude and altitude, but range from ~ 0 to 8 % per decade according to satellite data records (e.g. WMO, 2010). Therefore, for trend studies, the Ozone_cci suggests that instruments be stable within 3 % per decade (Ozone_cci, 2011).

The Optical Spectrograph and InfraRed Imaging System (OSIRIS) has been taking limb-scattered measurements of the atmosphere from 2001 to the present, yielding an 11 yr ozone number density data set. While excellent agreement was found between OSIRIS and the Stratospheric Aerosol and Gas Experiment II (SAGE II) for 2001–2005 (Adams et al., 2013), the long-term stability of the OSIRIS ozone data set has not yet been demonstrated. Furthermore, small biases were identified between OSIRIS and SAGE II, but they could not be fully characterized due to the small number of coincident measurements.

This work aims to characterize the OSIRIS ozone measurements fully using Microwave Limb Sounder (MLS) and Global Ozone Monitoring by Occultation of Stars (GOMOS) satellite, and ozonesonde data sets, all of which span at least 8 yr of the OSIRIS measurement period. The GOMOS data set was selected for validation because of its high precision and good vertical resolution (Tamminen et al., 2010), while the MLS data set has excellent spatiotemporal coverage (Waters et al., 2006) and ozonesondes have very high vertical resolution. Detailed comparisons between these instruments can be used for future projects, such as the merging of ozone data sets, including both OSIRIS and GOMOS, in the SPIN (ESA SPARC Initiative), Ozone_cci, and SI²N collaborations.

This paper is organized as follows. In Sect. 2, the OSIRIS instrument and ozone retrievals are presented, and the small changes that have occurred during the 11 yr mission are described. The three validation data sets are presented in Sect. 3, and the validation methodology is given in Sect. 4. In Sect. 5, the OSIRIS data set is characterized, with a focus on long-term consistency and the investigation of small biases within the data set. Finally, a summary of results is given in Sect. 6.

2 The OSIRIS ozone data set from 2001 to the present

The Canadian-made OSIRIS instrument, aboard the Swedish satellite Odin, was launched into a Sun-synchronous orbit on 20 February 2001 (Murtagh et al., 2002; Llewellyn et al., 2004). Odin has a polar orbit with a 96 min period with a northward equatorial crossing at $\sim 18:00$ LT (ascending node) and a southward equatorial crossing at $\sim 06:00$ LT (descending node). Ozone measurements are only taken in the summer hemisphere, with coverage in both hemispheres in the spring and autumn. The optical spectrograph measures limb scattered sunlight at 280–810 nm, with a ~ 1 nm spectral resolution. A review of the first decade of OSIRIS measurements is given by McLinden et al. (2012).

The OSIRIS SaskMART v5.0x ozone data are retrieved using the multiplicative algebraic reconstruction technique (MART) retrieval algorithm (Roth et al., 2007; Degenstein et al., 2009), which uses radiative transfer from the SASK-TRAN model (Bourassa et al., 2008b). Ozone number density profiles are retrieved from the cloud tops to 60 km (down to a minimum of 10 km in the absence of clouds) using a combination of ozone absorption information in the UV and visible parts of the spectrum. Aerosol and NO₂ are also retrieved with ozone to reduce biases (Bourassa et al., 2007, 2008a, 2011). SaskMART v5.0x ozone has an estimated precision of 3–4 % in the middle stratosphere (Bourassa et al., 2012b) and a vertical resolution of ~ 2 km at low altitudes and in the middle stratosphere, with decreasing resolution toward higher altitudes, reaching ~ 3 km at 50 km altitude. OSIRIS ozone data were filtered for outliers according to the techniques described by Adams et al. (2013).

From 2001 to the present, few changes have been observed in OSIRIS measurements. The absolute calibration of the OSIRIS spectra, which is used in the albedo calculations for the ozone retrievals, has shown very little variation over the course of the Odin mission. The calibration is performed twice per year, and most of the pixel responses have changed by less than 5 % over the mission. The dark current has increased over time but remains at insignificant levels for stratospheric ozone retrievals. Furthermore, the configuration of the Odin satellite is small and simple, which ultimately favours stability, as there are not many moving parts or operational modes.

The time series of OSIRIS optics temperatures (the temperature associated with the OSIRIS instrument and often used as a diagnostic) and latitudinal coverage of descending and ascending node measurements are shown in Fig. 1. Low optics temperatures are observed when Earth comes between Odin and the Sun for part of the orbit annually in May–July. These low temperatures are associated with a negative bias of 5–12 % in OSIRIS measurements when compared with SAGE II (Adams et al., 2013). This may be caused by misalignment between OSIRIS and the star tracker due to thermal deformation, leading to pointing errors (McLinden et al., 2007). Furthermore, the quality of ozone retrievals may be affected by defocusing, reduced spectral resolution, and wavelength shifts at low temperatures (Llewellyn et al., 2004). The optics temperature, both during and outside of this period of eclipse, decreased at a rate of ~ 0.5 °C yr⁻¹ from 2001 to 2012. Furthermore, due to the precession of the local time of the equatorial crossing, the coverage of descending versus ascending node measurements has changed over the 11 yr mission. Since biases have been observed between descending and ascending node ozone profiles (Adams et al., 2013), this precession could cause systematic changes in the OSIRIS ozone data set over time. Therefore the long-term validation of OSIRIS ozone verifies whether changes to the optics temperature and equatorial crossing time have introduced significant drifts within the time series.

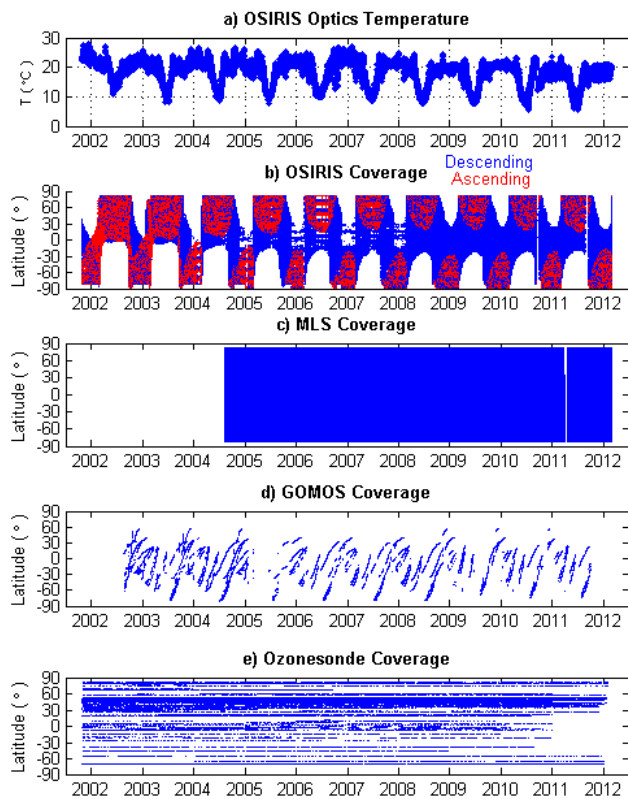


Fig. 1. Time series of (a) OSIRIS optics temperature and latitudes sampled by (b) OSIRIS, (c) MLS, (d) GOMOS, and (e) ozonesondes since the beginning of the OSIRIS ozone measurement record. OSIRIS descending node measurements are given in blue, and ascending node measurements are given in red.

3 Validation data sets

The MLS and GOMOS satellite data sets and balloon-borne ozonesonde measurements were selected for the long-term validation of OSIRIS. The latitudinal and temporal coverage of these validation data sets are shown in Fig. 1. All three data sets span much of the OSIRIS time series, and sample a large range of latitudes. The three validation data sets complement one another. MLS yields excellent spatiotemporal coverage at 2.5–3 km vertical resolution, while GOMOS provides higher vertical resolution of ~ 2 km and single-profile precision. GOMOS measures in the same wavelength region (UV-visible) as OSIRIS, but uses the stellar occultations instead of scattered sunlight, while MLS measures in a different spectral region than OSIRIS. Ozonesonde data offer very high vertical resolution of 150 m in the troposphere and lower stratosphere. Since all three validation data sets use independent measurement techniques, the confidence in the conclusions about OSIRIS data quality and biases is increased compared with earlier papers. The validation instruments and ozone retrievals are described below.

The Earth Observing System MLS (Waters et al., 2006) is on board the Aura satellite, which was launched by NASA on 15 July 2004 and has been operational, taking measurements of ozone and other trace gases, from 13 August 2004 to the present. MLS measures thermal radiance emissions from the limb at spectral regions centred near 118, 190, 240 and 640 GHz, and 2.5 THz with heterodyne radiometers. Aura is in a Sun-synchronous orbit at a 705 km altitude with 98° inclination, a 98.8 min period, and an ascending equatorial crossing time of $\sim 13:45$ LT. This results in near-global coverage from 82° S to 82° N with 3500 limb scans daily. The MLS v2.2 ozone product (Froidevaux et al., 2008), which is retrieved in the 240 GHz spectral region, was used for this study. For MLS, the v3.3 ozone data set with higher vertical resolution is also available. However, for the present study, the v2.2 data set was used because it does not have the larger systematic oscillations in upper troposphere/lower stratosphere profiles (mainly at low latitudes) found in v3.3 data; the two versions are otherwise extremely similar, with systematic mean biases of less than 1–2%. In the stratosphere, profiles are retrieved on a pressure grid, with a resolution of six levels per decade or ~ 2.5 –3 km. Uncertainties in MLS v2.2 ozone are estimated at 5% for much of the stratosphere and reach 10% (and sometimes 20%) in the lower stratosphere (Froidevaux et al., 2008). MLS v2.2 ozone agrees with ozonesondes within 8% between 150 and 3 hPa, averaging over the globe, with a positive bias of ~ 20 –30% at 215 hPa (Jiang et al., 2007). MLS ozone data were filtered using the recommendations of Froidevaux et al. (2008).

GOMOS (Kyrölä et al., 2004; Bertaux et al., 2010) is a spectrometer on board the Envisat satellite, which operated from 2002 to 2012. GOMOS took stellar occultation measurements in three wavelength bands (248–690 nm, 755–774 nm, and 926–924 nm). From these measurements, vertical profiles of ozone, NO_2 , NO_3 , H_2O , O_2 and aerosol are retrieved. The GOMOS v6 ozone data were used in this study. The vertical resolution of the GOMOS ozone profiles is 2 km at altitude levels below 30 km and 3 km at altitude levels above 40 km, with a smooth transition in resolution for 30–40 km (Kyrölä et al., 2010). Since the stellar occultation method is self-calibrating, GOMOS data are useful for long-term studies. GOMOS v5 night-time measurements were validated against ozonesonde, microwave radiometer, and lidar measurements and were found to agree within 2% for 20–40 km (van Gijssel et al., 2010). The main changes in v6 processor are improved instrument calibration and accurate characterization of modelling errors implemented in the “full covariance matrix” inversion (Sofieva et al., 2010). Standard screening for outliers was performed (GOMOS, 2012). Furthermore, in this study, only measurements of sufficiently bright (visual magnitude < 1.9) and hot stars ($T > 7000$ K) in full dark illumination conditions were used, in order to achieve the best data quality.

Balloon-borne ozonesondes sample the ozone partial pressure in situ from the ground up to a burst height around

Table 1. Coincidence criteria and smoothing techniques.

	Time (\pm h)	Distance (\pm km)	Latitude (\pm $^\circ$)	Smoothing
OSIRIS-MLS	6*	500	1	OSIRIS smoothed to the resolution of MLS using least squares fitting technique followed by smoothing with averaging kernels
OSIRIS-GOMOS	24	1000	1	–
OSIRIS-Ozone sondes	24	1000	1	Ozonesondes smoothed to the resolution of OSIRIS using triangular filter

* For each OSIRIS measurement, only the MLS measurement closest in time was selected as a coincident pair.

33–35 km, with a vertical resolution of ~ 150 m. A pump sucks ambient air into an electrochemical cell, within which a sensing solution is oxidized by the available ozone. The produced current is directly proportional to the ozone partial pressure, and can be converted to number density using the pressure–temperature data from the attached pressure, temperature, relative humidity (PTU) radiosonde. Ozone soundings are carried out regularly at numerous stations around the world, mostly using electrochemical concentration cell (ECC) sondes. The bias in ozone partial pressure is estimated to be about 3–5 %, with a precision around 5–7 % (Smit et al., 2007; Deshler et al., 2008). Ozonesonde measurements were obtained from the World Ozone and Ultraviolet Radiation Data Centre (WOUDC, 2012). Ozonesonde data were filtered using the criteria of Hassler et al. (2008) and were not used above 33 km or above 5 hPa because of reduced data quality at higher altitudes.

4 Methodology

The coincidence criteria and smoothing techniques used in this study are summarized for the various measurement pairs in Table 1. Coincident measurement pairs were selected based on criteria in geographic distance, latitude, and time differences. The coincidence criteria were selected to be as narrow as possible while covering the largest possible range of OSIRIS measurement (e.g. both orbital nodes, various optics temperatures, solar zenith angles, etc.) for a complete bias characterization. All comparisons were performed in number densities using the OSIRIS altitude grid, which is evenly spaced at 1 km intervals. For OSIRIS versus MLS, a distance criterion of ± 500 km, a time criterion of ± 6 h, and a latitude criterion of $\pm 1^\circ$ were employed. For each OSIRIS measurement only the single MLS measurement nearest in time was selected as its coincident pair. The resolution of MLS (~ 2.5 – 3 km) is slightly lower than the resolution of OSIRIS (~ 2.2 km below 30 km and increasing to ~ 3 km at 50 km altitudes). Furthermore, MLS data are retrieved on a pressure versus volume mixing ratio (VMR) grid, while OSIRIS profiles are on an altitude versus number density

grid. The OSIRIS data were transformed to a pressure versus VMR grid using European Centre for Medium-range Weather Forecasting (ECMWF) analysis data and were then interpolated to the MLS grid, using a least-squares fitting technique (e.g. Livesey et al., 2011). The averaging kernels and MLS a priori profiles were then used to present OSIRIS and MLS profiles in the same vertical resolution, using the technique of Rodgers and Connor (2003). Because of sharply peaked averaging kernels, this smoothing had a minor impact on the results. Finally, the MLS and OSIRIS pressure versus VMR profiles were converted back to an altitude versus number density grid using the ECMWF analysis data, and interpolated to the OSIRIS retrieval altitude grid. The impact of the data set used to convert MLS pressure versus VMR to altitude versus number density is discussed in Sect. 5.3. In addition to this method recommended by the MLS team, we have tested also a simpler approach: conversion of MLS data to number density and interpolation to the OSIRIS grid, and found practically the same results (as expected).

For comparisons with GOMOS and ozonesondes, a broader set of coincidence criteria of ± 1000 km, ± 24 h, and $\pm 1^\circ$ latitude were employed because the measurement frequency of these data sets is lower than for MLS. In comparisons between OSIRIS and SAGE II, these coincidence criteria yielded similar overall results to narrower coincidence criteria (± 500 km, ± 1 h, $\pm 1^\circ$ latitude), with slightly weaker correlation due to mismatching of air masses (Adams et al., 2013). For GOMOS and OSIRIS coincidences, smoothing was not applied because they have similar vertical resolution. Ozonesonde data were smoothed to the resolution of OSIRIS using a triangular filter, with a half-width of 2 km. Note that a Gaussian filter was also tested and yielded very similar results.

In order to assess overall agreement between data sets, the mean percent difference PD between the OSIRIS (M^{os}) and validation (M^{val}) ozone number density measurements is defined as

$$\text{PD}(z) = 100 \% \times \frac{1}{N(z)} \sum_{i=1}^N \frac{(M_i^{\text{os}}(z) - M_i^{\text{val}}(z))}{M_i^{\text{val}}(z)}, \quad (1)$$

where N is the number of coincident measurements. The standard deviation (σ) and the standard error (σ/\sqrt{N}) in the percent differences between coincident measurements were also calculated. For each altitude level, the R correlation coefficient between all available coincident measurements at the given altitude was calculated. Correlation was calculated between individual measurements (not, for example, daily zonal means).

In order to assess the long-term consistency of OSIRIS ozone data, time series of percent differences between OSIRIS and the correlative data sets were analysed. Hubert et al. (2014) tested techniques for evaluating drifts between satellite instruments and ozonesonde and lidar networks. They calculated drifts for various temporal averaging settings (e.g. no averaging, daily, monthly), regression techniques, and regression models (e.g. the inclusion of a seasonal term). The methods used in the present study were based on the preferred settings of Hubert et al. (2014), and are described below. Linear regressions of daily-averaged mean percent differences versus time were performed using a bi-square weighted robust fitting technique (e.g. Holland and Welsch, 1977). The robust fitting technique is preferred to classical least squares fits as it is less sensitive to outliers. The 2σ fitting error of the drift was approximately corrected for autocorrelation of the noise (Weatherhead et al., 1998), using the formula

$$\sigma' = \sigma \sqrt{\frac{1 + \phi}{1 - \phi}}, \quad (2)$$

where ϕ is the autocorrelation within the percent difference time series for a lag of one measurement, σ the fitting error, and σ' the autocorrelation-corrected error. This yields a more conservative estimate of σ . Error estimates were also calculated using the bootstrapping technique (Efron, 1979) over the fitting residuals, and were found to be very similar to the standard fitting error. For comparisons with MLS and GOMOS, the drift was calculated within 10° latitude bins. For ozonesondes, calculations were performed at individual ozonesonde stations to avoid artificial drifts if, for example, a station started or stopped taking measurements part way through the time series. Drifts were only considered for latitude bins or ozonesonde stations with at least 8 yr of correlative measurements, and at least 10 measurements per year.

5 Results

5.1 Overall agreement

Comparison statistics for all OSIRIS coincidences with MLS, GOMOS, and ozonesondes are shown in Fig. 2. For comparisons with MLS there are >200 000 coincidences, and for comparisons with GOMOS and ozonesondes there are >10 000 coincidences. At most altitudes, the standard deviations from both instruments are much larger than the

reported measurement errors (not shown) and are similar to one another, indicating that OSIRIS and the validation data sets have sampled similar large-scale seasonal and latitudinal structures in the ozone field.

Mean percent differences between OSIRIS and the validation data sets are within 5 % of zero at all altitude layers above 18.5 km for MLS, above 21.5 km for GOMOS, and at all altitudes for ozonesondes. For 14.5–16.5 km, OSIRIS has a negative bias compared with MLS of 10–11 %. Below 18.5 km, the variability in the GOMOS data is much larger than OSIRIS, suggesting that comparisons at these altitudes are unreliable. Above 20.5 km, OSIRIS agrees with MLS within 2 %, except for positive biases of ~ 2 –4 % for 22.5–24.5 km and 39.5–45.5 km. A positive bias for 22.5–24.5 km is also observed in comparisons with ozonesondes and was noted in comparisons with SAGE II (Adams et al., 2013). Possible reasons for this bias are explored further in Sect. 5.2.2. OSIRIS agrees with GOMOS within 2 % for 22.5–53.5 km. The small positive bias for 22.5–24.5 km is not observed in these comparisons because there are few coincidences between OSIRIS and GOMOS at high latitudes, where this bias is the largest.

Correlation coefficients indicate that OSIRIS ozone data are well correlated with coincident MLS, GOMOS, and ozonesonde data sets. For comparisons with MLS, R greater than 0.8 is observed for 11.5–44.5 km, with R values exceeding 0.9 at many altitude levels. For comparisons with ozonesondes, R is greater than 0.8 above 10.5 km. Correlations between OSIRIS and GOMOS coincidences, which are primarily sampled at lower latitudes, are slightly weaker. For all three validation data sets, the correlation decreases at ~ 24.5 km. This dip in correlation is reduced if the coincidence criteria are narrowed (not shown here), indicating that this is caused by the mismatching of air masses. The correlation at 24.5 km was found to be weakened primarily by lower latitude measurements, where it coincides with the peak in the ozone profile, suggesting possible mismatching of the altitude of the peak in the ozone profile.

Above 45 km, correlation is expected to weaken as coincident measurement pairs taken at different local times sample different parts of the diurnal cycle of ozone. Furthermore, OSIRIS measurements near twilight are affected by the diurnal effect (e.g. Natarajan et al., 2005; McLinden et al., 2006), as various local times are sampled along the line of sight. With the additional coincidence criterion of a difference in solar zenith angle $< 2^\circ$ within the same twilight, R improved from 0.77 to 0.87 at 54.5 km for OSIRIS versus MLS coincidences. Furthermore, R for coincident OSIRIS and GOMOS measurements improved from 0.73 to 0.77 at 54.5 km when OSIRIS measurements were restricted to solar zenith angles greater than 85° , to match the GOMOS nighttime measurements better. This suggests that discrepancies at high altitudes can be partly explained by the diurnal variation of ozone. In order to assess biases at these altitudes further,

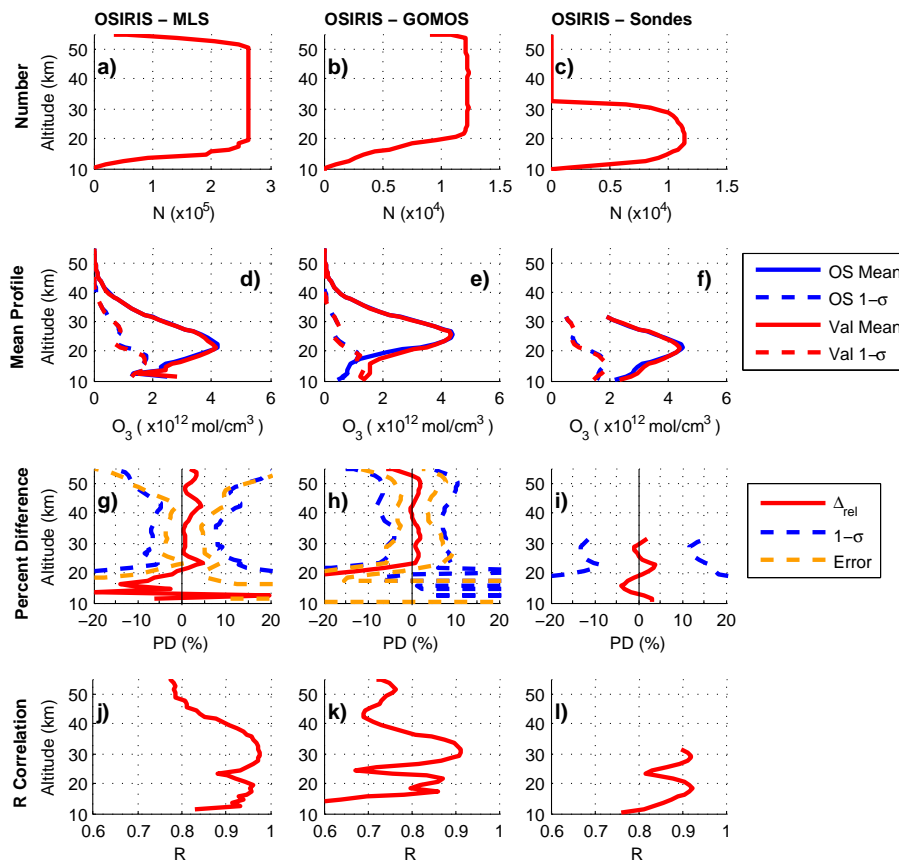


Fig. 2. Summary of global mean comparison results for (a, d, g and j) OSIRIS versus MLS, (b, e, h and k) OSIRIS versus GOMOS, and (c, f, i and l) OSIRIS versus ozonesondes over the full time period of coincident measurements. (a–c) Number of coincidences. (d–f) Mean (solid lines) and standard deviation (dashed lines) of coincident OSIRIS (blue) and validation data set (red) ozone profiles. (g–i) Mean of percent differences (red), standard deviation (blue), and combined retrieval error added in quadrature (yellow) for OSIRIS minus validation data sets. (j–l) R correlation coefficient for OSIRIS versus validation data sets. Data were not weighted (by, for example, latitude or season) prior to calculation of mean ozone profiles and mean percent differences.

full characterization of the diurnal variation of ozone and the diurnal effect would be necessary.

Figure 3 shows comparison results for the coincidences between OSIRIS and the three validation data sets, binned by latitude and altitude. At most latitudes, agreement between OSIRIS and the validation data sets is within 5% above 20 km. In some latitude bins, the positive bias at 22.5–24.5 km exceeds 5%. The largest discrepancies between OSIRIS and the validation data sets are observed at low altitudes, as expected due to the difficulties in measuring ozone from satellite platforms where number densities are small and the atmosphere becomes optically thick. At low latitudes (40° S–40° N), between the tropopause and 22.5 km, OSIRIS has a negative bias against all three validation data sets, with values of ~5–20% for comparisons with MLS and ozonesondes. At higher latitudes (latitudes north of 40° N and south of 40° S), OSIRIS has a negative bias of ~5% compared with MLS and ozonesondes for 10.5–17.5 km. Note that larger positive and negative percent

differences are observed at Southern Hemisphere high latitudes for comparisons with MLS, and are associated with large standard deviations. These values do not improve when only Southern Hemisphere summer months are considered, suggesting that this is not due to polar stratospheric clouds.

Figure 4 shows comparison results for OSIRIS versus ozonesondes, when data are re-gridded in altitude relative to the World Meteorological Organization (WMO) thermal tropopause, calculated from ECMWF analysis data. Note that the lower limit of OSIRIS retrievals is 10 km, so air masses below the tropopause are primarily measured at low latitudes. At the tropopause, OSIRIS data have a negative bias of $-10 \pm 1\%$, where the given error is the standard error. Furthermore, R is less than 0.8 from the tropopause to 4 km above the tropopause, indicating weaker correlation between OSIRIS and the ozonesondes in this region. The reduced agreement at the tropopause may be caused by mismatching of air masses due to the broad coincidence criteria and/or by imperfect smoothing of the ozonesondes to the vertical

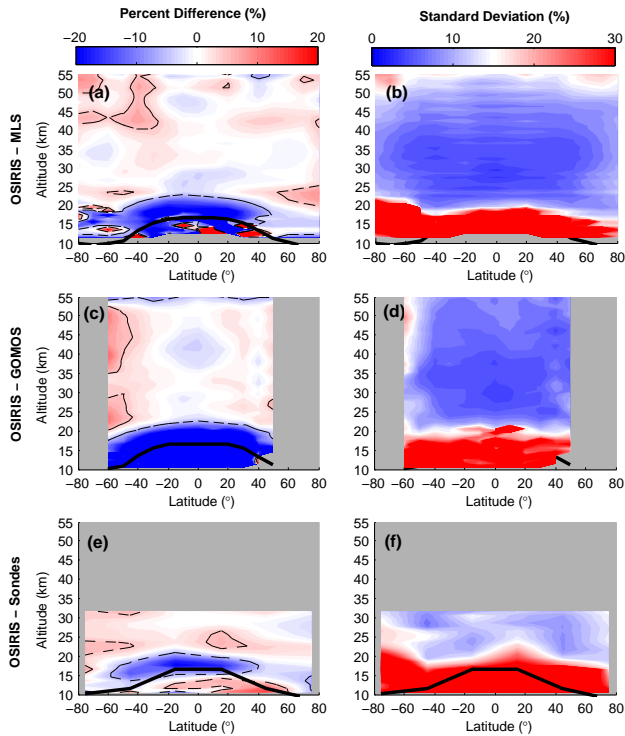


Fig. 3. Comparison results binned by latitude for (a and b) OSIRIS minus MLS, (c and d) OSIRIS minus GOMOS, and (e and f) OSIRIS minus ozonesondes. (a, c and e) Mean percent difference, with $\pm 5\%$ contour lines indicated with dashed lines. (b, d and f) Standard deviation in mean percent difference. Comparison results were calculated in 10° latitude bins for comparisons with MLS and GOMOS and in 30° latitude bins for comparisons with ozonesondes. The thick black lines indicate the average World Meteorological Organization (WMO) thermal tropopause height of the coincident measurements, calculated from ECMWF analysis data. The grey shading indicates regions for which there are fewer than 10 coincidences.

resolution of OSIRIS. For tropospheric air, 2–5 km beneath the tropopause, strong correlation $R > 0.9$ is observed between OSIRIS and ozonesondes. When this was investigated further, it was found that the large correlation coefficients were influenced by large ozone values ($> 1 \times 10^{12} \text{ mol cm}^{-3}$) recorded by both OSIRIS and ozonesondes below the thermal tropopause. This suggests that the correlation below the tropopause was influenced by stratospheric air masses, due perhaps in part to poor estimates of the tropopause height at the measurement locations. As an approximate method to remove this effect, the correlation below the tropopause was re-calculated with the requirement that ozone measured by both the ozonesonde and by OSIRIS be $< 1 \times 10^{12} \text{ mol cm}^{-3}$. The R coefficients for these comparisons (shown in blue in Fig. 4) are lower ($R \sim 0.5\text{--}0.6$), but still indicate that OSIRIS and ozonesonde data are correlated below the tropopause. This suggests that OSIRIS captures tropospheric variations, despite mean percent differences of up to 15 %.

5.2 Investigation of OSIRIS biases

In order to understand the OSIRIS measurements fully, the agreement between OSIRIS and the validation data sets was characterized for various OSIRIS retrieval parameters including optics temperature, aerosol extinction, albedo, solar zenith angle, and measurement node. Several systematic biases were identified and are discussed in the sub-sections below.

5.2.1 OSIRIS optics temperature

OSIRIS ozone measurements have previously been shown to have a dependence on the optics temperature, with lower ozone values for lower optics temperatures in the middle and upper stratosphere (Adams et al., 2013). One possible cause for this type of bias is pointing error due to the thermal deformation of the spacecraft (McLinden et al., 2007), leading to a shift of the profile in altitude. Another possible cause for this bias is the reduced spectral resolution and wavelength shifts that occur when optics temperatures are low because the instrument was focused at a higher temperature (Llewellyn et al., 2004). During these periods, ozone retrievals could be biased low at UV wavelengths because the ozone cross section is not correctly smoothed for the instrument's spectral resolution. UV wavelengths are used for retrievals in the middle and upper stratosphere, so this could also explain a negative bias in ozone in the upper and middle stratosphere for low optics temperatures.

Figure 5 shows mean percent differences between OSIRIS and the validation data sets binned according to the OSIRIS optics temperature. For optics temperatures $< 16^\circ\text{C}$ OSIRIS ozone data have negative biases for 25.5–40.5 km of 1–6 %, while for optics temperatures $> 16^\circ\text{C}$ positive biases of $\sim 1\text{--}3\%$ are observed. This qualitative change in the biases between OSIRIS and the validation instruments with temperature is consistent with SAGE II comparisons, in which a much larger negative bias of 5–12 % was observed for OSIRIS optics temperatures $< 16^\circ\text{C}$, but no positive bias was noted for larger optics temperatures (Adams et al., 2013). The reason for the smaller magnitude in the bias in the present study is not known. However, there are many factors that affect the comparisons, including, for example, latitudinal distributions of coincidences. Note that the larger negative biases observed below 22.5 km at lower latitudes (Fig. 3) are not apparent here for comparisons with MLS and ozonesondes because most of these coincidences are at higher latitudes.

In order to investigate whether pointing errors contribute to the negative bias with low optics temperatures, profile comparisons were performed in various latitude bands. An example of this is shown in Fig. 6 for OSIRIS versus MLS at both high and low latitudes. In both latitude ranges, OSIRIS has a more negative bias for low optics temperatures in the middle and upper stratosphere and a more positive bias at low

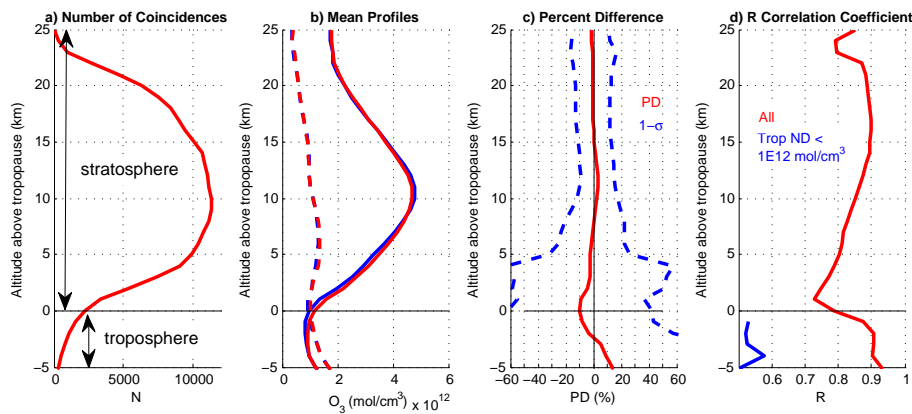


Fig. 4. As for the right column of Fig. 2, for OSIRIS and ozonesonde measurements re-gridded relative to the WMO thermal tropopause calculated with ECMWF analysis data. The black horizontal zero line indicates the tropopause. In (b) OSIRIS (blue) and ozonesonde (red) mean (complete lines) and standard deviation (dashed lines) in profiles are shown. In (c) OSIRIS mean percent difference (red) and standard deviation (dashed blue lines) are shown. In (d), the R correlation coefficients for all data (red) and for the troposphere when only data with $< 1 \times 10^{12}$ mol cm^{-3} of ozone at each layer are included.

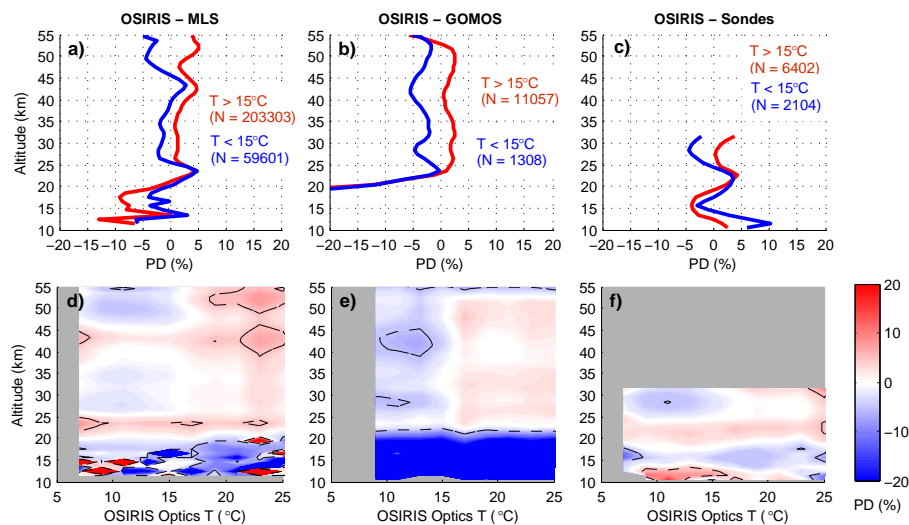


Fig. 5. Mean percent differences for (a and d) OSIRIS minus MLS, (b and e) OSIRIS minus GOMOS and (c and f) OSIRIS minus ozonesondes at various OSIRIS optics temperatures. (a–c) Profiles for all coincident measurements with OSIRIS optics temperatures $> 15^\circ\text{C}$ (red) and $< 15^\circ\text{C}$ (blue). The number of coincidences at 30.5 km is given by N in the text boxes. (d–f) Contours of mean percent differences (colour scale) were calculated for 2°C OSIRIS optics temperature bins (x axis) at various altitudes (y axis). The black dashed lines indicate $\pm 5\%$ mean percent difference. The grey shading indicates bins for which there are fewer than 10 coincidences.

altitudes. This is consistent with pointing errors. If OSIRIS measures an amount of ozone at a given registered altitude, but is actually pointing higher in the atmosphere due to a pointing error, there would be a negative bias in the ozone profile in the middle and upper stratosphere because the ozone number density decreases with altitude. Similarly, at lower altitudes, below the ozone number density maximum, OSIRIS would have a positive bias because ozone number densities decrease with altitude. In Fig. 6, the OSIRIS low optics temperature bias switches sign relative to the bias at regular temperatures at 24.5 km for low latitudes and 18.5 km for high latitudes. This corresponds to the approximate ozone number density maxima for the two latitude

bands. Similar figures produced at other latitude bands and with the other validation data sets yielded similar results (not shown here). This supports the explanation that the negative bias in middle-stratosphere ozone measurements for low optics temperature is caused in part by pointing errors. This does not rule out a contribution from the lower spectral resolution under low optics temperature, which also likely contributes to the low bias in OSIRIS measurements in the middle and upper stratosphere. Improvements to the retrieval software to account for spectral resolution and pointing errors are currently being tested and will be implemented for future versions of the OSIRIS ozone data set.

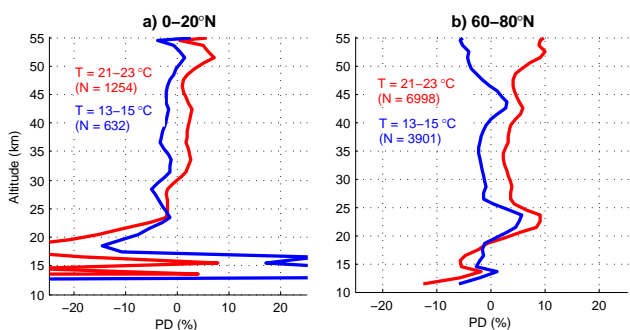


Fig. 6. Percent different versus altitude for OSIRIS minus MLS for coincidences with OSIRIS optics temperatures of 21–23 °C (red) and 13–15 °C (blue) for (a) 0–20° N and (b) 60–80° N. The number of coincidences at 30.5 km is given by N in the text.

5.2.2 OSIRIS ascending versus descending measurement nodes

Mean percent differences between OSIRIS and the validation data sets, divided into ascending and descending node measurements, are shown in Fig. 7. In the tropical troposphere, there are hemispheric asymmetries in comparisons of ascending and descending nodes with MLS and ozonesondes. Biases are more positive (negative) in the Northern Hemisphere than in the Southern Hemisphere for the descending (ascending) node measurements. The positive bias in OSIRIS measurements at 22.5–24.5 km is largest in the Southern Hemisphere for descending node measurements (Fig. 7d, f) and in the Northern Hemisphere for ascending node measurements (Fig. 7g, i). For 25.5–40.5 km agreement is within 5 % at most latitudes for both ascending and descending nodes. The hemispheric asymmetry is qualitatively consistent with comparisons with SAGE II (Adams et al., 2013).

Note that the ascending versus descending node biases are separate from the optics temperature biases (discussed in Sect. 5.2.1). The bias with optics temperature is similar to Fig. 5 when comparisons are performed for only one OSIRIS measurement node (not shown here). Similarly, the ascending versus descending node biases are similar to Fig. 7 when a restricted optics temperature range is used (not shown here).

Below ~ 30 km, some of the information in OSIRIS ozone retrievals is gained from the Chappuis band (544 nm, 602 nm, and 679 nm) (Degenstein et al., 2009). At these wavelengths, limb-scattered sunlight is sensitive to aerosols. At 600 nm, the peak sensitivity is at ~ 23 km (Fig. 1 of Bourassa et al., 2007). In order to investigate the connection between retrieved aerosol and ozone, OSIRIS descending and ascending nodes measurements, matched using coincidence criteria of ± 24 h, ± 1000 km, and $\pm 1^\circ$ latitude, were used. The mean relative differences in ozone from these collocated measurements are given in latitude–altitude bins by Adams et al. (2013, Fig. 8) and are consistent with Fig. 7. Latitudinal

structure was observed in the differences between ascending and descending node measurements even for these coincident OSIRIS versus OSIRIS measurements. Therefore, the observed latitudinal structure in comparisons between ascending and descending nodes and the validation data sets is not caused solely by the different seasonal coverage of the ascending and descending nodes.

Figure 8 shows the mean percent difference between OSIRIS aerosol extinctions in the descending minus ascending nodes at various latitudes. Note that this is the aerosol used in the ozone retrievals, not the OSIRIS level 2 aerosol extinction product, which is screened prior to distribution. OSIRIS aerosol has been validated previously with SAGE II and SAGE III and was found to agree within 15 % for the lower stratosphere (Bourassa et al., 2012a, 2007). Latitudinal and hemispheric biases in the aerosol are observed and are qualitatively similar to the ascending versus descending node ozone biases.

OSIRIS aerosol and ozone retrievals depend on the viewing geometry of the measurement, which varies systematically with latitude and season. Therefore, it is very difficult to isolate relationships between OSIRIS ozone and aerosol retrievals. However, when mean percent differences in ascending minus descending node ozone were binned by latitude, and the difference in retrieved aerosol extinction at 22.5 km, as shown in Fig. 9, the relationship between the difference in ozone and the difference in aerosol was observed. While 22.5 km is not the peak of the aerosol layer at all latitudes, it is the altitude at which OSIRIS ozone is sensitive to aerosol. For latitudes north of 30° N, and south of 30° S, the measurement node with the larger aerosol extinction observes more ozone. The magnitude of this bias increases toward higher latitudes. Mean percent differences for OSIRIS minus MLS at 22.5 km, binned by latitude and the OSIRIS aerosol extinction are also shown. At high latitudes, OSIRIS measures more ozone than MLS when OSIRIS aerosol extinctions are large for both ascending and descending nodes. While this aerosol-dependent bias is the clearest at 22.5 km, near the peak sensitivity of limb-scattered measurements to aerosol, some systematic dependence on aerosol extinction was observed for ~ 12.5 –27.5 km (not shown). This may explain the positive bias in OSIRIS measurements at 22.5 km.

Aerosol can be used as a dynamical tracer (e.g. Harvey et al., 1999) and is therefore expected to be related to ozone for natural reasons. For example, aerosol extinctions and ozone amounts are both expected to be low within the polar vortex. In order to test whether the retrieved ozone at 22.5 km is related to aerosol for this reason, the analysis was repeated with a variety of filters described below. In order to test for mismatched air masses, the coincidence criteria were narrowed to ± 1 h and ± 500 km. Furthermore a stratospheric temperature criterion of ± 1 K was applied. The analysis was also repeated for individual seasons, to determine whether, for example, the presence of the polar vortex affected results. In all cases, results were consistent with Fig. 9, suggesting

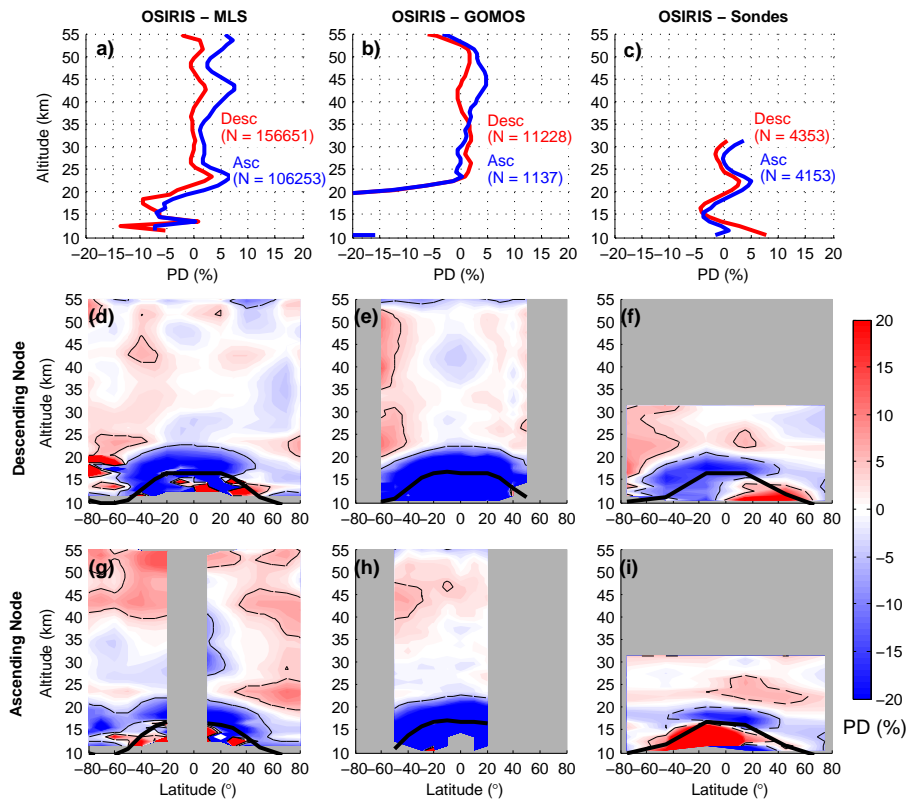


Fig. 7. Mean percent differences for (a, d and g) OSIRIS minus MLS, (b, e and h) OSIRIS minus GOMOS and (c, f and i) OSIRIS minus ozonesondes for OSIRIS ascending versus descending nodes. (a–c) Profiles for all coincident measurements for OSIRIS ascending node (red) and descending node (blue). The number of coincidences at 30.5 km is given by N in the text boxes. Contours of mean percent differences (colour scale) for the OSIRIS (d–f) descending node and (g–i) ascending node were calculated for 10° latitude bins (x axis) at various altitudes (y axis). The black dashed lines indicate $\pm 5\%$ mean percent difference. The grey shading indicates bins for which there are fewer than 10 coincidences.

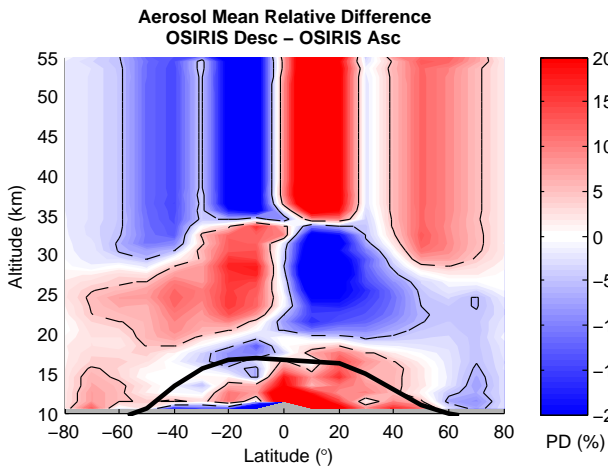


Fig. 8. As for left panel of Fig. 3, for OSIRIS aerosol extinction measured in the descending node minus aerosol extinction measured in the ascending node.

that biases observed between ascending and descending node ozone are related to similar biases in the retrieved aerosol extinction. Improvements to the aerosol retrieval software are currently being tested and may lead to the reduction of biases in ozone between OSIRIS ascending and descending node measurements in future versions of the data product.

5.2.3 Albedo

Above ~ 30 km, agreement between OSIRIS and the validation data sets was found to be related to the OSIRIS albedo, which measures apparent upwelling and is obtained by fitting the absolute value of the 740 nm modelled limb radiance at 40 km by adjusting the albedo with a forward model (Bourassa et al., 2007). The strongest bias was observed at ~ 42.5 km, and is shown in Fig. 10 for comparisons with MLS and GOMOS. At all latitudes, OSIRIS ozone measurements are larger for higher albedo. At these altitudes, UV wavelengths are used in the ozone retrievals, so very little limb-scattered sunlight would originate from the lower altitudes, suggesting that this bias is not caused by errors in the radiative transfer due to poor estimates of albedo. Therefore,

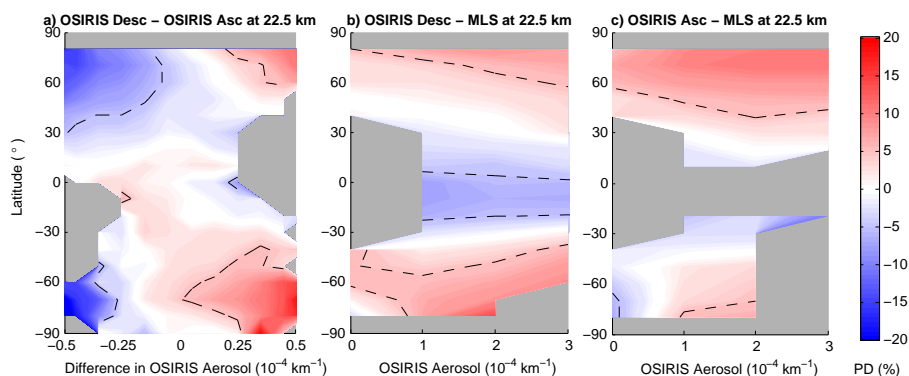


Fig. 9. (a) Percent difference (colour scale) in ozone for coincident OSIRIS descending node minus ascending node measurements binned by the difference in OSIRIS descending node minus ascending node aerosol extinction at 22.5 km (x axis) and latitude (y axis). Mean percent difference (colour scale) in ozone for OSIRIS (b) descending node and (c) ascending node minus MLS binned according the OSIRIS aerosol extinction at 22.5 km (x axis) and latitude (y axis). The dashed contour lines indicate mean percent differences of $\pm 5\%$. Comparison results were calculated in 10° latitude bins. The grey shading indicates bins for which there are fewer than 10 coincidences.

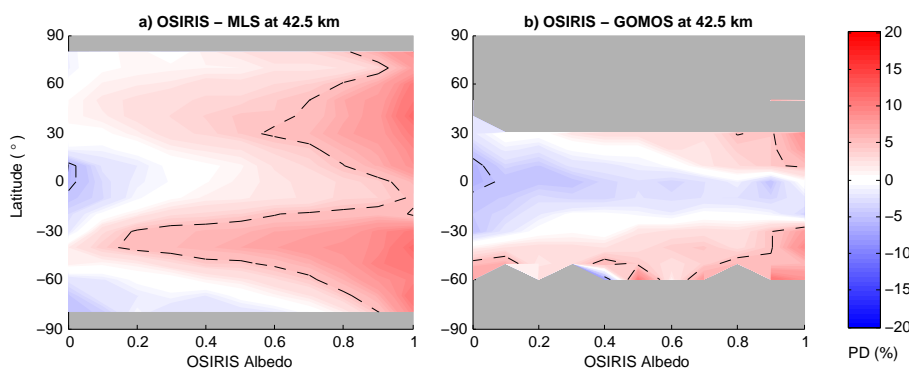


Fig. 10. Mean percent differences for (a) OSIRIS minus MLS and (b) OSIRIS minus GOMOS at 42.5 km in albedo (x axis) and 10° latitude (y axis) bins. The black dashed lines indicate $\pm 5\%$ mean percent difference. The grey shading indicates bins for which there are fewer than 10 coincidences.

this may reflect the relationship between albedo and another OSIRIS measurement parameter. In order to check for this, Fig. 10 was reproduced and found to be consistent under a variety of additional conditions, including various narrow ranges of OSIRIS optics temperature, solar zenith angle, solar scattering angle, and measurement season. This observed bias may also point to a feature in the spectra in this altitude range, since albedo is retrieved using radiances at 40 km. The reason for the relationship between ozone and albedo at these altitudes remains under investigation.

5.3 Drift analysis

MLS ozone profiles were previously found to be stable relative to lidar to within $\pm 5\%$ per decade for 20–40 km at most lidar stations considered (Nair et al., 2013). GOMOS data are expected to be stable because the stellar occultation technique is self-calibrating. In addition, only occultations of sufficiently bright and hot stars are used in our analyses. Thus the potential problems related to weak and cool stars,

as discussed for example in Keckhut et al. (2010), are removed. Ozonesondes are widely used for the assessment of long-term satellite stability (e.g. SPARC/IOC/GAW, 1998).

Figure 11 shows drift analysis results in 10° latitude bins for OSIRIS minus MLS and OSIRIS minus GOMOS. Since ozone trends are small, a few percent per decade, stable instruments are needed to detect them. The Ozone_cci recommends that data sets with drifts with magnitudes $< 3\%$ per decade be used for trend studies (Ozone_cci, 2011). Above 20.5 km, drifts are $< 2\%$ per decade at most latitudes and altitudes for comparisons with MLS and GOMOS, suggesting good stability in the OSIRIS measurements. Some latitudinal structure is observed in the drifts, with positive drifts exceeding 3% per decade in the tropics at some altitudes in both the MLS and GOMOS comparisons. Furthermore, at low latitudes near the tropopause, drifts of up to 10% per decade are observed, with errors of $\sim 5\text{--}10\%$ per decade. These larger drifts are mostly positive, but the latitudinal structure of the sign of the drifts is not exactly consistent between the MLS

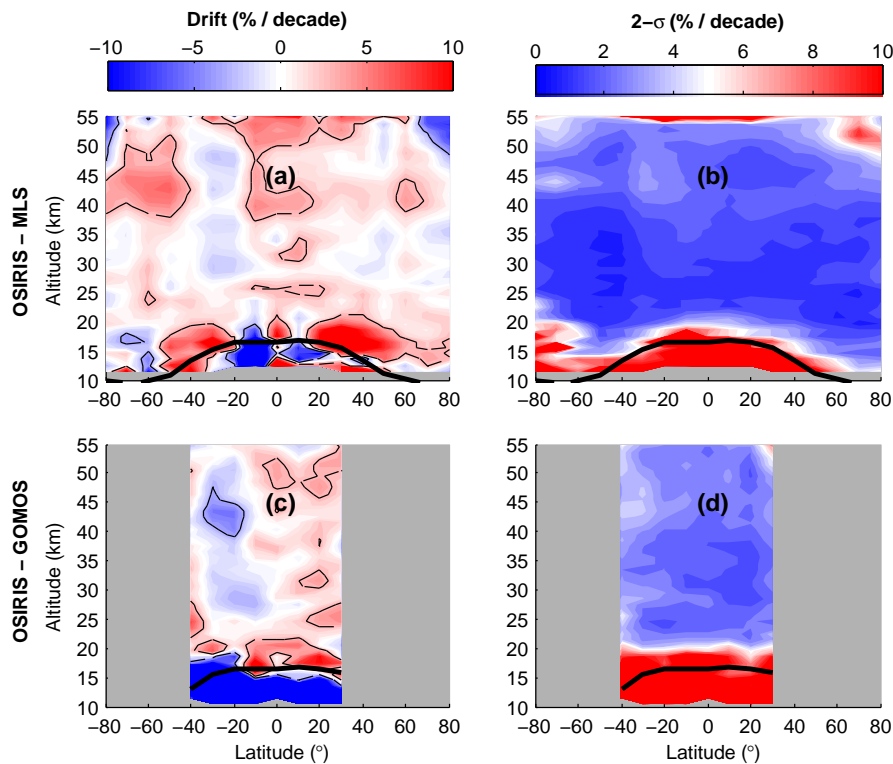


Fig. 11. Drift analysis results for (a and b) OSIRIS minus MLS and (c and d) OSIRIS minus GOMOS in 10° latitude bins. (a and c) Drift in percent per decade, with drifts of $\pm 3\%$ indicated by the black dashed contour lines. (b and d) 2σ error in drift estimate. The thick black lines indicate the average WMO thermal tropopause height of the coincident measurements, calculated from ECMWF analysis data. The grey shading indicates regions for which there are fewer than 10 coincidences.

and GOMOS data sets. This suggests that perhaps the apparent drifts at lower altitudes are affected by sampling of highly variable percent differences.

The mean of the drifts within the latitude bins and ozonesonde stations is shown in Fig. 12, with the associated standard error. Absolute values of global drifts are $< 3\%$ per decade above 18.5 km relative to MLS and above 19.5 km relative to GOMOS. Drifts relative to ozonesondes are within 3% per decade for 13.5–22.5 km, with a positive drift of 3–5 % per decade for 23.5–29.5 km. For comparisons with MLS, drifts plus or minus the standard error are within $\pm 3\%$ per decade at all altitudes. At altitudes below ~ 25 km, small positive drifts are observed in the OSIRIS data (~ 1 – 3% per decade) and are significant from zero according to error estimates.

When assessing the time series of percent differences between two instruments, inconsistencies in parameters used to enable the comparisons may bias the result. In particular, erroneous trends in temperature may manifest as trends in ozone, as temperature affects neutral density and layer thickness. This has been observed for temperatures from the National Centers for Environmental Prediction (NCEP) reanalysis (e.g. McLinden and Fioletov, 2011). For comparisons with GOMOS, this is not an issue, as GOMOS, like OSIRIS, retrieves number density on an altitude grid. Furthermore, for

ozonesondes, temperature is recorded simultaneously, and can be used for unit conversion.

Drift calculations with MLS data are potentially problematic because the ECMWF analysis data were used to perform the unit and coordinate conversion. Similar to NCEP (McLinden and Fioletov, 2011), ECMWF temperatures possess a warming trend between 2000 and 2012 above about 35 km (not shown here). Therefore the unit conversion was re-calculated using two alternate sources of pressure and temperature: MLS-measured temperature and geopotential height values, and a trended temperature–altitude–pressure climatology. The MLS temperatures and geopotential heights were screened using the recommendations of Schwartz et al. (2008). Neutral densities and geometric heights were then calculated from the temperature and geopotential height profiles. The climatological neutral densities were calculated from a temperature–altitude climatology (Nagatani and Rosenfield, 1993), with the linear trend in temperature from Fig. 19 of Randel et al. (2009) superimposed. At each altitude in the climatology, the corresponding pressure was calculated using the hydrostatic equation, with a surface pressure of 1000 hPa. The climatological approach will lead to some errors but is useful to ensure that the experimental long-term temperature trend is being used.

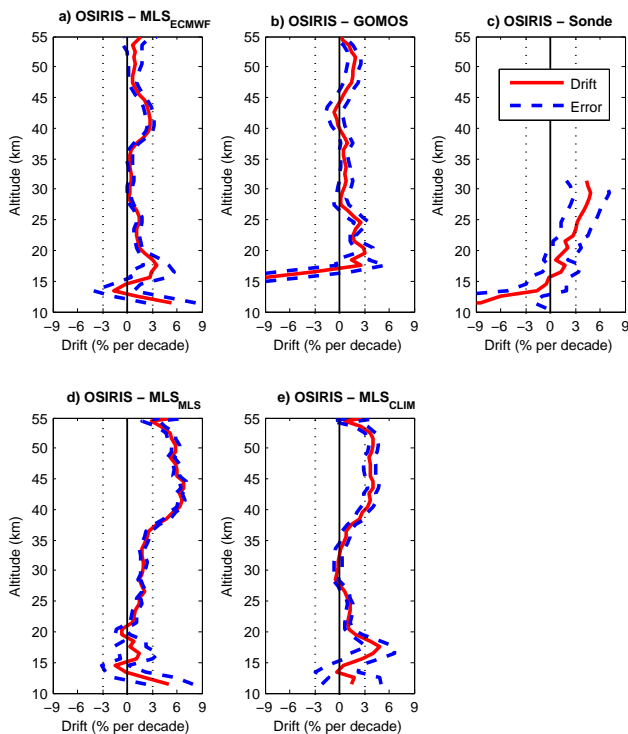


Fig. 12. Global average results from drift analysis for (a) OSIRIS minus MLS converted with ECMWF altitudes and temperatures, (b) OSIRIS minus GOMOS, (c) OSIRIS minus ozonesondes, (d) OSIRIS minus MLS converted using MLS-measured altitudes and temperatures, and (e) OSIRIS minus MLS converted using climatological altitudes and temperatures with measured temperature trends superimposed on the data set. The global average drift (solid red line) and the standard error (dashed blue line) across latitude bins and sonde station are shown. The thin black dotted lines indicate $\pm 3\%$ per decade.

These two neutral density conversions were found to have minimal impact on the overall comparison results, particularly below 40 km and therefore have not been presented in this work. However, they did affect the drift calculations, particularly above 35 km, as shown in Fig. 12. In the 19.5–36.5 km range, all three conversion approaches for MLS data lead to relative drifts within the $\pm 3\%$ per decade threshold, but with drift estimates differing by as much as 2 % per decade. Above 36.5 km, positive drifts of 3–6 % per decade are calculated when the MLS data and climatological conversion approaches are used, compared to 0–3 % per decade with ECMWF data. These discrepancies are consistent with the altitude range in which the warming trend was observed in ECMWF temperatures. Therefore, the assessment of drifts between data sets measured in different units is problematic, especially in the upper stratosphere where reanalyses often produce unreliable temperature trends (Gaffen et al., 2000).

To summarize, the OSIRIS data set meets the Ozone_cci recommendation for stability relative to MLS and GOMOS for 19.5–36.5 km. Above 36.5 km, drifts between OSIRIS

and GOMOS meet the standards, but drifts relative to MLS vary from 0 to 6 % per decade, depending on the conversion of MLS profiles to the OSIRIS grid. Therefore, this work demonstrates that gradual changes in the local time of the OSIRIS ascending node and optics temperature (Fig. 1) have not caused significant drifts in the OSIRIS ozone data set, with the most robust results for 19.5–36.5 km.

6 Conclusions

The 2002–present OSIRIS v5.0x ozone data set was characterized using MLS v2.2, GOMOS v6, and ozonesonde measurements. Biases in the OSIRIS data set were investigated in detail by binning mean percent differences according to various measurement parameters. For optics temperatures $< 16^\circ\text{C}$, OSIRIS ozone data have negative biases for 25.5–40.5 km of 1–6 %, while for optics temperatures $> 16^\circ\text{C}$ positive biases of ~ 1 –3 % are observed. This is qualitatively consistent with, but smaller than the biases observed in previous comparisons with SAGE II v7.0 (Adams et al., 2013). The way that this bias varies with latitude and altitude is consistent with a pointing error when optics temperatures are low. Reduced spectral resolution during low optics temperatures also likely contributes. Biases between ascending and descending node measurements were also observed and are consistent with comparisons with SAGE II (Adams et al., 2013). Below 30 km, these biases were found to be associated with biases in aerosol extinction, particularly near 22.5 km. This may therefore explain the 2–4 % positive bias in OSIRIS measurements at this altitude. Above 30 km, ozone biases varied systematically with albedo. The reasons for this are still under investigation. The information on OSIRIS biases gained through this study will be used in order to improve the retrieval algorithm for future versions of the OSIRIS ozone data set.

While biases were identified, good overall agreement between OSIRIS and the validation data sets was observed. Overall, OSIRIS agreed with all the three validation data sets to within 5 % above 21.5 km. Furthermore, the absolute value of the global average drift between OSIRIS and the validation data sets was $< 3\%$ per decade for comparisons with MLS for 19.5–36.5 km, with GOMOS above 19.5 km, and with ozonesondes for 13.5–22.5 km. For comparisons with MLS, the calculated drift was found to be dependent on the choice in data sets for conversion to the OSIRIS altitude versus neutral density grid, particularly above 35 km, where trends in ECMWF temperatures do not match the measured trends. Overall this work demonstrates that the 11 yr OSIRIS ozone data set has remained consistent with other ozone measurements throughout the Odin mission, and can therefore be used in merged ozone data products and trend studies.

Acknowledgements. We thank Nick Lloyd, Chris Roth, and Landon Rieger for answering many questions about the OSIRIS data set. This work was supported by the Natural Sciences and Engineering Research Council (Canada) and the Canadian Space Agency. Odin is a Swedish-led satellite project funded jointly by Sweden (SNSB), Canada (CSA), France (CNES), and Finland (Tekes). The work at the Finnish Meteorological Institute was supported by the ESA Ozone_cci project and the Academy of Finland (projects MIDAT and ASTREX). Work at the Jet Propulsion Laboratory, California Institute of Technology, was performed under contract with the National Aeronautics and Space Administration. The correlative data from balloon-based ozonesonde used in this publication were obtained from the World Ozone and Ultraviolet Radiation Data Centre (WOUDC) archives (see <http://www.woudc.org>). We thank warmly several members of the NDACC community and ozonesonde working group for fruitful discussions.

Edited by: P. K. Bhartia

References

- Adams, C., Bourassa, A. E., Bathgate, A. F., McLinden, C. A., Lloyd, N. D., Roth, C. Z., Llewellyn, E. J., Zawodny, J. M., Flittner, D. E., Manney, G. L., Daffer, W. H., and Degenstein, D. A.: Characterization of Odin-OSIRIS ozone profiles with the SAGE II dataset, *Atmos. Meas. Tech.*, 6, 1447–1459, doi:10.5194/amt-6-1447-2013, 2013.
- Bertaux, J. L., Kyrölä, E., Fussen, D., Hauchecorne, A., Dalaudier, F., Sofieva, V., Tamminen, J., Vanhellefont, F., Fanton d'Andon, O., Barrot, G., Mangin, A., Blanot, L., Lebrun, J. C., Pérot, K., Fehr, T., Saavedra, L., Leppelmeier, G. W., and Fraisse, R.: Global ozone monitoring by occultation of stars: an overview of GOMOS measurements on ENVISAT, *Atmos. Chem. Phys.*, 10, 12091–12148, doi:10.5194/acp-10-12091-2010, 2010.
- Bourassa, A. E., Degenstein, D. A., Gattinger, R. L., and Llewellyn, E. J.: Stratospheric aerosol retrieval with optical spectrograph and infrared imaging system limb scatter measurements, *J. Geophys. Res.*, 112, D10217, 1–15, doi:10.1029/2006JD008079, 2007.
- Bourassa, A. E., Degenstein, D. A., and Llewellyn, E. J.: Retrieval of stratospheric aerosol size information from OSIRIS limb scattered sunlight spectra, *Atmos. Chem. Phys.*, 8, 6375–6380, doi:10.5194/acp-8-6375-2008, 2008a.
- Bourassa, A. E., Degenstein, D. A., and Llewellyn, E. J.: SASK-TRAN: A spherical geometry radiative transfer code for efficient estimation of limb scattered sunlight, *J. Quant. Spectrosc. Ra.*, 109, 52–73, doi:10.1016/j.jqsrt.2007.07.007, 2008b.
- Bourassa, A. E., McLinden, C. A., Sioris, C. E., Brohede, S., Bathgate, A. F., Llewellyn, E. J., and Degenstein, D. A.: Fast NO₂ retrievals from Odin-OSIRIS limb scatter measurements, *Atmos. Meas. Tech.*, 4, 965–972, doi:10.5194/amt-4-965-2011, 2011.
- Bourassa, A. E., Rieger, L. A., Lloyd, N. D., and Degenstein, D. A.: Odin-OSIRIS stratospheric aerosol data product and SAGE III intercomparison, *Atmos. Chem. Phys.*, 12, 605–614, doi:10.5194/acp-12-605-2012, 2012a.
- Bourassa, A. E., McLinden, C. A., Bathgate, A. F., Elash, B. J., and Degenstein, D. A.: Precision estimate for Odin-OSIRIS limb scatter retrievals, *J. Geophys. Res.*, 117, D04303, doi:10.1029/2011JD016976, 2012b.
- Degenstein, D. A., Bourassa, A. E., Roth, C. Z., and Llewellyn, E. J.: Limb scatter ozone retrieval from 10 to 60 km using a multiplicative algebraic reconstruction technique, *Atmos. Chem. Phys.*, 9, 6521–6529, doi:10.5194/acp-9-6521-2009, 2009.
- Deshler, T., Mercer, J. L., Smit, H. G. J., Stubi, R., Levrat, G., Johnson, B. J., Oltmans, S. J., Kivi, R., Thompson, A. M., Witte, J., Davies, J., Schmidlin, F. J., Brothers, G., and Sasaki, T.: Atmospheric comparison of electrochemical cell ozonesondes from different manufacturers, and with different cathode solution strengths: The Balloon Experiment on Standards for Ozonesondes, *J. Geophys. Res.*, 113, D04307, doi:10.1029/2007JD008975, 2008.
- Efron, B.: Bootstrap Methods: Another Look at the Jackknife, *Ann. Stat.*, 7, 1–26, 1979.
- Froidevaux, L., Jiang, Y. B., Lambert, A., Livesey, N. J., Read, W. G., Waters, J. W., Browell, E. V., Hair, J. W., Avery, M. A., McGee, T. J., Twigg, L. W., Sumnicht, G. K., Jucks, K. W., Margitan, J. J., Sen, B., Stachnik, R. A., Toon, G. C., Bernath, P. F., Boone, C. D., Walker, K. A., Filipiak, M. J., Harwood, R. S., Fuller, R. A., Manney, G. L., Schwartz, M. J., Daffer, W. H., Drouin, B. J., Cofield, R. E., Cuddy, D. T., Jarnot, R. F., Knosp, B. W., Perun, V. S., Snyder, W. V., Stek, P. C., Thurstans, R. P., and Wagner, P. A.: Validation of Aura Microwave Limb Sounder stratospheric ozone measurements, *J. Geophys. Res.*, 113, D15S20, doi:10.1029/2007JD008771, 2008.
- Gaffen, D. J., Sargent, M. A., Habermann, R. E., and Lazante, J. R.: Sensitivity of Tropospheric and Stratospheric Temperature Trends to Radiosonde Data Quality, *J. Climate*, 13, 1776–1796, 2000.
- GOMOS: GOMOS Level 2 Product Quality Readme File, ENVI-GSOP-EOGD-QD-12-0117, Issue 1.0, 1–15, available at: http://earth.eo.esa.int/pcs/envisat/gomos/documentation/RMF_0117_GOM_NL__2P_Disclaimers.pdf (last access: 28 February 2012), 2012.
- Harvey, V. L., Hitchman, M. H., Pierce, R. B., and Fairlie, T. D.: Tropical aerosol in the Aleutian High, *J. Geophys. Res.*, 104, 6281–6290, 1999.
- Hassler, B., Bodeker, G. E., and Dameris, M.: Technical Note: A new global database of trace gases and aerosols from multiple sources of high vertical resolution measurements, *Atmos. Chem. Phys.*, 8, 5403–5421, doi:10.5194/acp-8-5403-2008, 2008.
- Holland, P. W. and Welsch, R. E.: Robust regression using iteratively reweighted least-squares, *Commun. Stat. A-Theor.*, 6, 813–827, doi:10.1080/03610927708827533, 1977.
- Hubert, D., et al.: Ground-based network assessment of fourteen occultation/limb ozone profilers over 1984–2012, *Atmos. Meas. Tech. Discuss.*, in preparation, 2014.
- Jiang, Y. B., Froidevaux, L., Lambert, A., Livesey, N. J., Read, W. G., Waters, J. W., Bojkov, B., Leblanc, T., McDermid, I. S., Godin-Beekmann, S., Filipiak, M. J., Harwood, R. S., Fuller, R. A., Daffer, W. H., Drouin, B. J., Cofield, R. E., Cuddy, D. T., Jarnot, R. F., Knosp, B. W., Perun, V. S., Schwartz, M. J., Snyder, W. V., Stek, P. C., Thurstans, R. P., Wagner, P. A., Allaart, M., Andersen, S. B., Bodeker, G., Calpini, B., Claude, H., Coetzee, G., Davies, J., De Backer, H., Dier, H., Fujiwara, M., Johnson, B., Kelder, H., Leme, N. P., König-Langlo, G., Kyro, E., Laneve, G., Fook, L. S., Merrill, J., Morris, G., Newchurch, M.,

- Oltmans, S., Parrondos, M. C., Posny, F., Schmidlin, F., Skrivankova, P., Stubi, R., Tarasick, D., Thompson, A., Thouret, V., Viatte, P., Vömel, H., von Der Gathen, P., Yela, M., and Zabolck, G.: Validation of Aura Microwave Limb Sounder Ozone by ozonesonde and lidar measurements, *J. Geophys. Res.*, 112, D24S34, doi:10.1029/2007JD008776, 2007.
- Keckhut, P., Hauchecorne, A., Blanot, L., Hocke, K., Godin-Beekmann, S., Bertaux, J.-L., Barrot, G., Kyrölä, E., van Gijssels, J. A. E., and Pazmino, A.: Mid-latitude ozone monitoring with the GOMOS-ENVISAT experiment version 5: the noise issue, *Atmos. Chem. Phys.*, 10, 11839–11849, doi:10.5194/acp-10-11839-2010, 2010.
- Kyrölä, E., Tamminen, J., Leppelmeier, G. W., Sofieva, V., Hasinen, S., Bertaux, J. L., Hauchecorne, A., Dalaudier, F., Cot, C., Korabely, O., Fanton d'Andon, O., Barrot, G., Mangin, a., Théodore, B., Guirlet, M., Etanchaud, F., Snoeij, P., Koopman, R., Saavedra, L., Fraisse, R., Fussen, D. and Vanhellemont, F.: GOMOS on Envisat: an overview, *Adv. Space Res.*, 33, 1020–1028, doi:10.1016/S0273-1177(03)00590-8, 2004.
- Kyrölä, E., Tamminen, J., Sofieva, V., Bertaux, J. L., Hauchecorne, A., Dalaudier, F., Fussen, D., Vanhellemont, F., Fanton d'Andon, O., Barrot, G., Guirlet, M., Mangin, A., Blanot, L., Fehr, T., Saavedra de Miguel, L., and Fraisse, R.: Retrieval of atmospheric parameters from GOMOS data, *Atmos. Chem. Phys.*, 10, 11881–11903, doi:10.5194/acp-10-11881-2010, 2010.
- Livesey, N. J., Read, W. G., Froidevaux, L., Lambert, A., Manney, G. L., Pumphrey, H. C., Santee, M. L., Schwartz, M. J., Wang, S., Cofield, R. E., Cuddy, D. T., Fuller, R. A., Jarnot, R. F., Jiang, J. H., Knosp, B. W., Stek, P. C., Wagner, P. A., and Wu, D. L.: EOS MLS Version 3.3 Level 2 data quality and description document, available at: http://mls.jpl.nasa.gov/data/v3-3_data_quality_document.pdf (last access: 2 January 2014), 2011.
- Llewellyn, E. J., Lloyd, N. D., Degenstein, D. A., Gattinger, R. L., Petelina, S. V., Bourassa, A. E., Wiensz, J. T., Ivanov, E. V., McDade, I. C., Solheim, B. H., McConnell, J. C., Haley, C. S., Von Savigny, C., Sioris, C. E., McLinden, C. A., Griffioen, E., Kaminski, J., Evans, W. F. J., Puckrin, E., Strong, K., Wehrle, V., Hum, R. H., Kendall, D. J. W., Matsushita, J., Murtagh, D. P., Brohede, S., Stegman, J., Witt, G., Barnes, G., Payne, W. F., Piché, L., Smith, K., Warshaw, G., Deslauniers, D., Marchand, P., Richardson, E. H., King, R. A., Wevers, I., McCreath, W., Kyrölä, E., Oikarinen, L., Leppelmeier, G. W., Auvinen, H., Mégie, G., Hauchecorne, A., Lefèvre, F., De La Nöe, J., Ricaud, P., Frisk, U., Sjöberg, F., Von Schéele, F., and Nordh, L.: The OSIRIS instrument on the Odin spacecraft, *Can. J. Phys.*, 82, 411–422, doi:10.1139/P04-005, 2004.
- McLinden, C. A. and Fioletov, V.: Quantifying stratospheric ozone trends: Complications due to stratospheric cooling, *J. Geophys. Res.*, 38, L03808, doi:10.1029/2010GL046012, 2011.
- McLinden, C. A., Haley, C. S., and Sioris, C. E.: Diurnal effects in limb scatter observations, *J. Geophys. Res.*, 111, D14302, doi:10.1029/2005JD006628, 2006.
- McLinden, C. A., Fioletov, V. E., Haley, C. S., Lloyd, N., Roth, C., Degenstein, D., Bourassa, A., McElroy, C. T., and Llewellyn, E. J.: An evaluation of Odin/OSIRIS limb pointing and stratospheric ozone through comparisons with ozonesondes, *Can. J. Phys.*, 85, 1125–1141, doi:10.1139/P07-112, 2007.
- McLinden, C. A., Bourassa, A. E., Brohede, S., Cooper, M., Degenstein, D. A., Evans, W. J. F., Gattinger, R. L., Haley, C. S., Llewellyn, E. J., Lloyd, N. D., Loewen, P., Martin, R. V., McConnell, J. C., McDade, I. C., Murtagh, D., Rieger, L., von Savigny, C., Sheese, P. E., Sioris, C. E., Solheim, B. and Strong, K.: OSIRIS: A decade of scattered light, *B. Am. Meteorol. Soc.*, 93, 1845–1863, doi:10.1175/BAMS-D-11-00135.1, 2012.
- Murtagh, D., Frisk, U., Merino, F., Ridal, M., Jonsson, A., Stegman, J., Witt, G., Jiménez, C., Megie, G., Noë, J., De Ricaud, P., Baron, P., Pardo, J. R., Llewellyn, E. J., Degenstein, D. A., Gattinger, R. L., Lloyd, N. D., Evans, W. F. J., McDade, I. C., Haley, C. S., Sioris, C., Savigny, V., Solheim, B. H., McConnell, J. C., Richardson, E. H., Leppelmeier, G. W., Auvinen, H., and Oikarinen, L.: Review: An overview of the Odin atmospheric mission, *Can. J. Phys.*, 80, 309–319, doi:10.1139/P01-157, 2002.
- Nagatani, R. M. and Rosenfield, J. E.: Temperature, net heating and circulation, in *The Atmospheric Effects of Stratospheric Aircraft: Report of the 1992 Models and Measurements Workshop*, NASA Ref. Publ. 1291, edited by: Remsberg, E. E. and Prather, M. J., A1–A47, 1993.
- Natarajan, M., Deaver, L. E., Thompson, E., and Magill, B.: Impact of twilight gradients on the retrieval of mesospheric ozone from HALOE, *J. Geophys. Res.*, 110, D13305, doi:10.1029/2004JD005719, 2005.
- Nair, P. J., Godin-Beekmann, S., Kuttippurath, J., Ancellet, G., Goutail, F., Pazmiño, A., Froidevaux, L., Zawodny, J. M., Evans, R. D., Wang, H. J., Anderson, J., and Pastel, M.: Ozone trends derived from the total column and vertical profiles at a northern mid-latitude station, *Atmos. Chem. Phys.*, 13, 10373–10384, doi:10.5194/acp-13-10373-2013, 2013.
- Ozone_cci: Ozone_cci User Requirement Document, 2.1, 1–43, available at: http://www.esa-ozone-cci.org/?q=webfm_send/37 (last access: 5 December 2012), 2011.
- Randel, W. J., Shine, K. P., Austin, J., Barnett, J., Claud, C., Gillett, N. P., Keckhut, P., Langematz, U., Lin, R., Long, C., Mears, C., Miller, A., Nash, J., Seidel, D. J., Thompson, D. W. J., Wu, F., and Yoden, S.: An update of observed stratospheric temperature trends, *J. Geophys. Res.*, 114, D02107, doi:10.1029/2008JD010421, 2009.
- Rodgers, C. D. and Connor, B. J.: Intercomparison of remote sounding instruments, *J. Geophys. Res.*, 108, 4116, doi:10.1029/2002JD002299, 2003.
- Roth, C. Z., Degenstein, D. A., Bourassa, A. E., and Llewellyn, E. J.: The retrieval of vertical profiles of the ozone number density using Chappuis band absorption information and a multiplicative algebraic reconstruction technique, *Can. J. Phys.*, 85, 1225–1243, doi:10.1139/P07-130, 2007.
- Schwartz, M. J., Lambert, A., Manney, G. L., Read, W. G., Livesey, N. J., Froidevaux, L., Ao, C. O., Bernath, P. F., Boone, C. D., Cofield, R. E., Daffer, W. H., Drouin, B. J., Fetzer, E. J., Fuller, R. A., Jarnot, R. F., Jiang, J. H., Jiang, Y. B., Knosp, B. W., Krüger, K., Li, J.-L. F., Mlynarczyk, M. G., Pawson, S., Russell III, J. M., Santee, M. L., Snyder, W. V., Stek, P. C., Thurstans, R. P., Tompkins, A. M., Wagner, P. A., Walker, K. A., Waters, J. W., and Wu, D. L.: Validation of the Aura Microwave Limb Sounder temperature and geopotential height measurements, *J. Geophys. Res.*, 113, D15S11, doi:10.1029/2007JD008783, 2008.
- SI²N: SI²N Initiative, available at: http://igaco-o3.fmi.fi/VDO/working_groups.html (last access: 22 May 2012), 2012.
- Smit, H. G. J., Straeter, W., Johnson, B. J., Oltmans, S. J., Davies, J., Tarasick, D. W., Hoegger, B., Stubi, R., Schmidlin, F. J.,

- Northam, T., Thompson, A. M., Witte, J. C., Boyd, I., and Posny, F.: Assessment of the performance of ECC-ozone sondes under quasi-flight conditions in the environmental simulation chamber: Insights from the Juelich Ozone Sonde Intercomparison Experiment (JOSIE), *J. Geophys. Res.*, 112, D19306, doi:10.1029/2006JD007308, 2007.
- Sofieva, V. F., Vira, J., Kyrölä, E., Tamminen, J., Kan, V., Dalaudier, F., Hauchecorne, A., Bertaux, J.-L., Fussen, D., Vanhellefont, F., Barrot, G., and Fanton d'Andon, O.: Retrievals from GOMOS stellar occultation measurements using characterization of modeling errors, *Atmos. Meas. Tech.*, 3, 1019–1027, doi:10.5194/amt-3-1019-2010, 2010.
- SPARC/IOC/GAW: Assessment of Trends in the Vertical Distribution of Ozone, SPARC Report No.1, WMO Ozone Research and Monitoring Project Report No. 43, edited by: Harris, N., Hudson, R., and Phillips, C., May 1998, available at: http://www.sparc-climate.org/fileadmin/customer/6_Publications/SPARC_reports_PDF/1_Ozone_SPARCreportNo1_May1998_redFile.pdf (last access: 2 January 2014), 1998.
- Tamminen, J., Kyrölä, E., Sofieva, V. F., Laine, M., Bertaux, J.-L., Hauchecorne, A., Dalaudier, F., Fussen, D., Vanhellefont, F., Fanton-d'Andon, O., Barrot, G., Mangin, A., Guirlet, M., Blanot, L., Fehr, T., Saavedra de Miguel, L., and Fraisse, R.: GOMOS data characterisation and error estimation, *Atmos. Chem. Phys.*, 10, 9505–9519, doi:10.5194/acp-10-9505-2010, 2010.
- van Gijssel, J. A. E., Swart, D. P. J., Baray, J.-L., Bencherif, H., Claude, H., Fehr, T., Godin-Beekmann, S., Hansen, G. H., Keckhut, P., Leblanc, T., McDermid, I. S., Meijer, Y. J., Nakane, H., Quel, E. J., Stebel, K., Steinbrecht, W., Strawbridge, K. B., Tatarov, B. I., and Wolfram, E. A.: GOMOS ozone profile validation using ground-based and balloon sonde measurements, *Atmos. Chem. Phys.*, 10, 10473–10488, doi:10.5194/acp-10-10473-2010, 2010.
- Waters, J. W., Froidevaux, L., Harwood, R. S., Jarnot, R. F., Pickett, H. M., Read, W. G., Siegel, P. H., Cofield, R. E., Filipiak, M. J., Flower, D. A., Holden, J. R., Lau, G. K., Livesey, N. J., Manney, G. L., Pumphrey, H. C., Santee, M. L., Wu, D. L., Cuddy, D. T., Lay, R. R., Loo, M. S., Perun, V. S., Schwartz, M. J., Stek, P. C., Thurstans, R. P., Boyles, M. A., Chandra, K. M., Chavez, M. C., Chen, G., Chudasama, B. V., Dodge, R., Fuller, R. A., Girard, M. A., Jiang, J. H., Jiang, Y., Knosp, B. W., Labelle, R. C., Lam, J. C., Lee, K. A., Miller, D., Oswald, J. E., Patel, N. C., Pukala, D. M., Quintero, O., Scaff, D. M., Snyder, W. Van, Tope, M. C., Wagner, P. A., and Walch, M. J.: The Earth Observing System Microwave Limb Sounder (EOS MLS) on the Aura Satellite, *IEEE T. Geosci. Remote*, 44, 1075–1092, 2006.
- Weatherhead, E. C., Reinsel, G. C., Tiao, G. C., Meng, X., Choi, D., Cheang, W.-K., Keller, T., DeLuisi, J., Wuebbles, D. J., Kerr, J. B., Miller, A. J., Oltmans, S. J., and Frederick, J. E.: Factors affecting the detection of trends?: Statistical considerations and applications to environmental data range, *J. Geophys. Res.*, 103, 17149–17161, 1998.
- WMO: Scientific assessment of ozone depletion 2010, Rep. 52, Global ozone research monitoring project, Geneva, 2010.
- WOUDC: World Ozone and Ultraviolet Radiation Data Centre, available at: <http://www.woudc.org/> (last access: 1 July 2013), 2012.