Validation of SAGE III/ISS Solar Ozone Data with Correlative Satellite and Ground Based Measurements

H. J. Ray Wang[1], Robert Damadeo[2], David E. Flittner[2], Natalya A Kramarova[3], Ghassan Taha[4], Sean M. Davis[5], Anne M. Thompson[6], Susan E Strahan[4], Yuhang Wang[1], Lucien Froidevaux[7], Douglas A. Degenstein[8], Adam E. Edward Bourassa[8], Wolfgang Steinbrecht[9], Kaley A. Walker[10], Richard Querel[11], Thierry Leblanc[12], Sophie Godin-Beekmann[13], Dale F. Hurst[14], and Emrys Hall[15]

[1]Georgia Institute of Technology
[2]NASA Langley Research Center
[3]Science Systems and Applications, Inc.,
[4]Universities Space Research Association
[5]NOAA ESRL
[6]NASA-GODDARD
[7]JPL/California Institute of Technology, California, USA
[8]University of Saskatchewan
[9]Deutscher Wetterdienst
[10]University of Toronto
[11]National Institute of Water & Atmospheric Research (NIWA)
[12]California Institute of Technology
[13]LATMOS, CNRS, UVSQ, UPMC
[14]NOAA Earth System Research Laboratory
[15]NOAA Earth Systems Research Laboratory

November 22, 2022

Abstract

The Stratospheric Aerosol and Gas Experiment III on the International Space Station (SAGE III/ISS) was launched on February 19, 2017 and began routine operation in June 2017. The first two years of SAGE III/ISS (v5.1) solar ozone data were evaluated by using correlative satellite and ground-based measurements. Among the three (MES, AO3, and MLR) SAGE III/ISS solar ozone products, AO3 ozone shows the best accuracy and precision, with mean biases less than 5% for altitudes ~15–55 km in the mid-latitudes and ~20–55 km in the tropics. In the lower stratosphere and upper troposphere, AO3 ozone shows high biases that increase with decreasing altitudes and reach ~10% near the tropopause. Preliminary studies indicate that those high biases primarily result from the contributions of the oxygen dimer (O) not being appropriately removed within the ozone channel. The precision of AO3 ozone is estimated to be ~3% for altitudes between 20 and 40 km. It degrades to ~10–15% in the lower mesosphere (~55 km), and ~20–30% near the tropopause. There could be an altitude registration error of ~100 meter in the SAGE III/ISS auxiliary temperature and pressure profiles. This, however, does not affect retrieved ozone profiles in native number density on geometric altitude coordinates. In the upper stratosphere and lower mesosphere (~40–55 km) the SAGE III/ISS (and SAGE II) sunset ozone values are systematically higher than sunrise data by ~5–8% which are almost twice larger than what observed by other satellites or model predictions. This feature needs further study.

Validation of SAGE III/ISS Solar Ozone Data with Correlative Satellite and Ground Based Measurements

H.J. Ray Wang[1], Robert Damadeo[2], David Flittner[2], Natalya Kramarova[3], Ghassan Taha[4], Sean Davis[5], Anne M. Thompson[3], Susan Strahan[4], Yuhang Wang[1], Lucien Froidevaux[6], Doug Degenstein[7], Adam Bourassa[7], Wolfgang Steinbrecht[8], Kaley A. Walker[9], Richard Querel[10], Thierry Leblanc[11], Sophie Godin-Beekmann[12], Dale Hurst[13,14], Emrys Hall[13,14]

[1]School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA, USA

[2]NASA Langley Research Center, Hampton, VA, USA

[3]NASA Goddard Space Flight Center, Greenbelt, MD, USA

[4]Universities Space Research Association, NASA/Code 614, Greenbelt, MD, USA

[5]NOAA Earth System Research Laboratory Chemical Sciences Division, Boulder, CO, USA

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

[7]Institute of Space and Atmospheric Studies, University of Saskatchewan, Saskatoon, Saskatchewan, Canada

[8]Deutscher Wetterdienst, Hohenpeissenberg, Germany

[9]Department of Physics, University of Toronto, Toronto, Ontario, Canada

[10]National Institute of Water & Atmospheric Research (NIWA), Lauder, New Zealand

[11]Jet Propulsion Laboratory, California Institute of Technology, Wrightwood, CA 92397, USA

[12]LATMOS-ISPL, Université Paris 6 Pierre-et-Marie-Curie, Paris, France

[13]Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, USA

[14]NOAA Earth System Research Laboratory Global Monitoring Division, Boulder, CO, USA

Corresponding author: H.J. Ray Wang (raywang@eas.gatech.edu).

Key Points:

- Among the three SAGE III/ISS solar occultation retrievals, AO3 ozone product shows the best accuracy and precision.
- The mean biases of AO3 ozone are less than 5% for ~15–55 km in the mid-latitudes and ~20–55 km in the tropics. It increases to ~10% near the tropopause.
- The precision of AO3 ozone is ~3% for altitudes 20–40 km. It degrades to ~10–15% in the lower mesosphere (~55 km), and ~20–30% near the tropopause.

Abstract

The Stratospheric Aerosol and Gas Experiment III on the International Space Station (SAGE III/ISS) was launched on February 19, 2017 and began routine operation in June 2017. The first two years of SAGE III/ISS (v5.1) solar ozone data were evaluated by using correlative satellite and ground-based measurements. Among the three (MES, AO3, and MLR) SAGE III/ISS solar ozone products, AO3 ozone shows the best accuracy and precision, with mean biases less than 5% for altitudes ~15–55 km in the mid-latitudes and ~20–55 km in the tropics. In the lower stratosphere and upper troposphere, AO3 ozone shows high biases that increase with decreasing altitudes and reach ~10% near the tropopause. Preliminary studies indicate that those high biases primarily result from the contributions of the oxygen dimer (O_4) not being appropriately removed within the ozone channel. The precision of AO3 ozone is estimated to be ~3% for altitudes between 20 and 40 km. It degrades to ~10–15% in the lower mesosphere (~55 km), and ~20–30% near the tropopause. There could be an altitude registration error of ~100 meter in the SAGE III/ISS auxiliary temperature and pressure profiles. This, however, does not affect retrieved ozone profiles in native number density on geometric altitude coordinates. In the upper stratosphere and lower mesosphere (~40–55 km) the SAGE III/ISS (and SAGE II) sunset ozone values are systematically higher than sunrise data by ~5–8% which are almost twice larger than what observed by other satellites or model predictions. This feature needs further study.

1 Introduction

The Stratospheric Aerosol and Gas Experiment III on the International Space Station (SAGE III/ISS) is the second instrument from the SAGE III project. It was launched on a SpaceX Falcon 9/Dragon spacecraft on February 19, 2017 and began routine operation in June 2017. Similar to its predecessors, SAGE I (1979–1981), SAGE II (1984–2005), and SAGE III/M3M (2001–2006), SAGE III/ISS uses the solar occultation technique to retrieve vertical profiles of ozone (O_3), water vapor (H_2O), nitrogen dioxide (NO_2), and aerosol extinctions at multiple wavelengths (e.g., Mauldin et al., 1985; McCormick et al., 1993; Wang et al., 2006; Thomason et al., 2010). In addition, SAGE III can utilize the multi-spectral measurement of the oxygen A-band (758–771 nm) to derive vertical profiles of temperature and pressure (Pitts and Thomason, 2003). The SAGE series of observations has provided valuable data for understanding global ozone trends (SPARC/IO3C/GAW, 2019; WMO, 2018) and the impact of volcanoes and human activities on stratospheric aerosol (SPARC, 2006).

SAGE III/ISS can also observe the atmosphere at night by using the lunar occultation technique. Lunar occultation is achieved by rotating the solar attenuator out of the optical path and using a fully programmable Charged Couple Device (CCD) that enables selection of different spectral channels and integration times. The lunar observations can provide vertical profiles of ozone (O_3), nitrogen dioxide (NO_2), nitrogen trioxide (NO_3), and chlorine dioxide (OClO). A separate algorithm (e.g., Rault, 2005; Rault and Loughman, 2013) is being developed to retrieve trace gases from limb scattering measurements, which are still research products and not yet available to the public.

Unlike the first SAGE III instrument on the Meteor 3M spacecraft (SAGE III/M3M), which was in a sun synchronous orbit providing observations in the northern hemisphere at mid to high latitudes (~45°–80°N), and in the southern hemisphere at mid-latitudes (~35°–60°S), SAGE III/ISS is in a mid-inclination orbit (51.6°). The solar observations can provide near global (~70°S–70°N) measurements on a monthly basis with coverage similar to that of the

SAGE II measurements. There is, however, some loss of measurements due to the obscuration of the Sun by the ISS and limitations to operations due to spacecraft visits to ISS. The sampling coverage of SAGE III/ISS solar observations can be augmented by lunar measurements, which occur at locations and times not covered by solar observations.

In this paper, we evaluate the quality of SAGE III/ISS version 5.1 solar ozone data by comparisons with other independent measurements from satellites as well as ozonesondes and lidar. Section 2 describes the SAGE III retrieval algorithm, solar ozone products and some known anomalies in the current algorithm. The correlative satellite and ground-based datasets are described in section 3. Section 4 describes the coincident criteria and validation methodology. The comparison results are shown in section 5 and followed by the conclusions in section 6.

2 SAGE III/ISS solar ozone data

2.1 Instrument and retrieval overview

The SAGE III instruments makes solar occultation measurements by scanning a relatively small field-of-view (0.5 arcminutes in the vertical and 5.0 arcminutes in the horizontal) vertically across the face of the Sun and focusing the light into a simple grating spectrometer. The spectrometer uses a CCD array with 809 spectral columns with resolutions of ~1–2 nm that provide nearly continuous spectral coverage between ~280 and ~1035 nm as well as a single photodiode covering 1542 nm ± 15 nm. These 809 CCD pixels are then subsampled (i.e., read out individually or co-added or averaged with other pixels) into a number of “pixel groups” that change for different modes of operation. For solar occultation, there are 86 of these pixel groups (87 including the photodiode) that fall into 12 different channels illustrated in Fig. 1. For comparison, the central wavelength of the seven channels used by SAGE II are also shown in Figure 1.

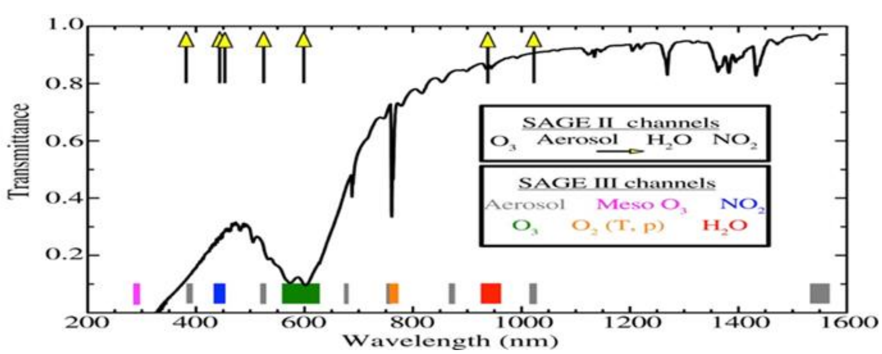


Figure 1: Sample wavelength dependence of atmospheric transmission in the lower stratosphere with locations of the different spectral channels used by the SAGE II (yellow arrows) and SAGE III/ISS instruments (colored boxes). The twelve SAGE III/ISS channels are color-coded by species of interest and are numbered from smallest to largest wavelength.

The current retrieval algorithm for SAGE III/ISS is version 5.1, which is essentially the same as that used for SAGE III/M3M. A complete description of the SAGE III retrieval algorithm is available in the SAGE III Algorithm Theoretical Basis Document: Solar and Lunar Algorithm (SAGE III ATBD, 2002). The algorithm consists of two main parts, the transmission algorithm and the species inversion algorithm. The transmission algorithm involves taking the raw uncalibrated radiance counts from the CCD (and photodiode) and converting them into line-of-sight (LOS) transmissions at each wavelength and tangent altitude. The species inversion algorithm uses these multi-wavelength LOS transmission profiles to derive vertical profiles of trace gas concentrations and aerosol extinction coefficients. This is done by first removing modeled contributions from Rayleigh scattering and O_4 absorption, then separating the remaining LOS transmission profiles into the contributions from each species of interest, and lastly inverting these LOS contributions into vertical profiles of concentration or extinction using a global fit inversion method (or a nonlinear Levenberg-Marquardt onion peeling method for water vapor or temperature/pressure retrievals).

The solar occultation retrieval for SAGE III actually produces three separate ozone products. The "MES" algorithm uses absorption features in the ultraviolet (<300 nm) to retrieve vertical profiles between ~45 and ~100 km. The other two use ozone absorption in the Chappuis band (near 600 nm) to retrieve vertical profiles from the surface or cloud top up to 70 km. Each uses the same pixel groups in the spectral channel surrounding 600 nm (Channel 5) but differ in how they treat aerosol and NO_2 within the retrieval. The "MLR" algorithm uses Channels 5 and 3 (~450 nm) to solve for both O_3 and NO_2 simultaneously while making an assumption about the spectral shape of aerosol extinction through each channel. The "AO3" algorithm removes the contributions from NO_2 that were solved in the MLR retrieval and then uses all of the data between Channels 4 and 11 (see Figure 1), excluding the O_2 A-band and the H_2O channels, to better constrain the influence of aerosol. The AO3 algorithm is similar to the retrieval used for the SAGE II instrument (e.g., Chu et al., 1989; Damadeo et al., 2013). It is worth noting that while the AO3 algorithm explicitly solves for aerosol extinction in each channel, this solution is not reported. Instead, the reported aerosol is computed as a residual while using the MLR solution for ozone and NO_2.

2.2 Known anomalies in version 5.1

The SAGE III/ISS instrument is by far the most/best characterized SAGE instrument. The detailed knowledge of the intricacies of the instrument's behavior and performance allow the SAGE III team to incorporate several new algorithms to improve the data quality. One such characterization is that of the spectral stray light within the spectrometer (reentrance spectra). While the instrument was still on the ground, a thorough characterization of the spectral stray light was performed on the instrument and one particular problem area was identified. A portion of the light incident on the UV range of the CCD actually comes from near the peak of Chappuis ozone absorption. This will have a negative impact on the mesospheric ozone retrieval and needs to be corrected. While a rudimentary correction is currently implemented, it stems from an ad-hoc correction derived for SAGE III/M3M data and does not use the most up-to-date information. As such, we do not recommend the MES ozone product for validation or research studies as it is still preliminary.

The SAGE III/ISS algorithm uses auxiliary temperature and pressure data from MERRA-2 (Modern-Era Retrospective analysis for Research and Applications, version 2) (GMAO, 2015, Gelaro et al., 2017), which is necessary for modeling refraction and molecular (Rayleigh) scattering. These data are provided with geopotential heights, which the SAGE algorithm converts to geometric altitudes at the location of the measurement. It has been discovered that this conversion between geopotential height and geometric altitude, which was actually copied from the SAGE II algorithm, was never thoroughly vetted and is more of an approximation (i.e., it assumes that the surface gravity is not latitude-dependent). As such, the current altitude registration of the meteorological products that pass through the algorithm, but not the retrieved profiles of species concentrations or aerosol extinctions, are biased on the order of 100 or so meters (altitude and latitude dependent) as shown in Figure 2 (see the Appendix for a recommended correction). The impact of this mis-registration would be most noticeable when converting SAGE III retrieved ozone from native number density on geometric altitude to mixing ratio on pressure (VMR/P) coordinates when using the reported temperatures and pressures in the SAGE data files, especially at higher altitudes (see Appendix). It is, of course, also noteworthy to point out that, since the code was present in the SAGE II v7.0 algorithm, that data product has a similar bias when making the same conversion to VMR/P coordinates.

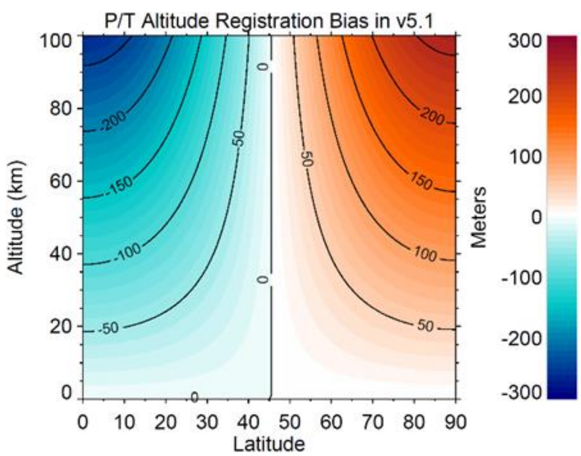


Figure 2: Estimated altitude registration errors in the reported SAGE III/ISS (v5.1) auxiliary temperature and pressure data.

Since aerosol measurements are intertwined with ozone measurements (i.e., through partitioning of the slant-path transmissions into the contributions from ozone, aerosol, and other interfering gases), assessing the quality of the aerosol product can also yield information about the quality of the ozone product. While aerosol extinctions at different wavelengths will vary with atmospheric conditions (e.g., total amount and type of aerosol from volcanoes and/or fires), it is expected that the "aerosol spectrum" (i.e., extinction as a function of wavelength) should be slowly varying and monotonic in almost all stratospheric conditions (Thomason et al., 2010). Instead, the aerosol spectrum derived from SAGE III/ISS measurements exhibits a "dip" near 600 nm that has different characteristics in different altitude regimes (latitude-dependent) as shown in Figure 3. At altitudes in the troposphere and lowermost stratosphere (below ~20 km in the tropics), this dip follows the shape of the ozone cross-sections and is systematically larger at lower altitudes. The primary contribution appears to be an error in the creation of the spectroscopic database for O_4 used by the retrieval algorithm (i.e., a preprocessing error, not an error in the source cross-sections themselves). This yields the incorrect spectroscopic shape of O_4, which aliases into the retrieval and results in a solution for ozone that should be too large (discussed later) when the contribution to extinction from O_4 is significant (i.e., scales with density squared). Since aerosol is solved as a residual using MLR ozone, any systematically large ozone would cause systematically small aerosol showing a wavelength-dependence that scales with the ozone cross-sections. At altitudes above the lowermost stratosphere (≳20 km in the tropics), this dip still follows the shape of the ozone cross-sections but scales with the ozone mixing ratio. A possible explanation for this is that the overall magnitude of the source ozone cross-section database is too large by 1–2% percent in the Chappuis relative to the other channels, but this requires further study. It is noteworthy that the magnitude of the dips is smaller in the aerosol data produced by the AO3 algorithm (not shown). This suggests that the use of additional aerosol channels in the retrieval better constrains the allowable shape of the aerosol spectrum, resulting in a potentially more robust aerosol data product. The SAGE team is investigating if the aerosol solution from the AO3 algorithm should be the released data product in future versions.

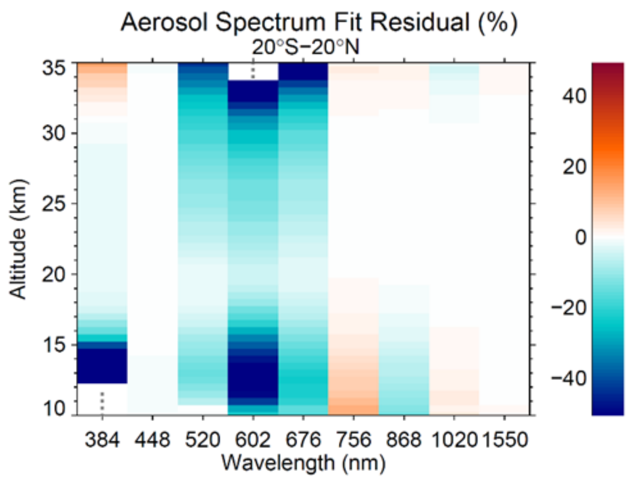


Figure 3: Residuals of a quadratic fit to the aerosol spectrum in log-log space using the aerosol extinctions at 448, 756, 868, 1020, and 1550 nm. The residuals are the median relative residuals of all SAGE III/ISS data from June 2017 to May 2019 between 20°S and 20°N. Results at mid-latitudes are similar, simply shifted down in altitude. Grey stippling denotes areas where aerosol extinction data does not exist. The median residuals in the channels used for the quadratic fit are <1% between ~20–30 km.

3 Correlative satellite and ground based ozone datasets

3.1 Aura MLS

The Earth Observing System (EOS) Microwave Limb Sounder (MLS) aboard the Aura satellite has provided daily global measurements of ozone (O_3) profiles and other trace gases from the upper troposphere to the upper mesosphere from August 2004 to present. Aura MLS measures thermal radiance emissions in 5 broad regions between 118 GHz and 2.5 THz by scanning the Earth's atmospheric limb vertically from the ground to ~90 km (Waters et al., 2006). Aura is in a sun-synchronous near-polar orbit with ascending equatorial crossing time of ~13:45 LT. Unlike the UARS MLS instrument, which observed limb emission in a direction perpendicular to the spacecraft flight direction, Aura MLS observes emission from the atmosphere directly ahead of the satellite. This results in near global-coverage from both daytime and nighttime measurements with ~3500 profiles each day.

Aura MLS ozone retrieved from the 240 GHz spectral region by using an optimal estimation approach (Rodgers, 2000; Livesey et al., 2006) is the standard reported ozone product. It has a vertical resolution of 2.5–3 km from the upper troposphere to the lower mesosphere, and ~5 km in the upper mesosphere. As indicated by comparisons with correlative measurements, the estimated accuracy of MLS v2.2 ozone is within about 5% for much of the stratosphere. The biases increase with decreasing altitudes, with some systematic positive biases of 10–20% in the lowest portion of the stratosphere (Froidevaux et al., 2008; Livesey et al., 2008) and ~20–30% in the upper troposphere (Jiang et al., 2007).

The latest Aura MLS v4.23 ozone data were used in this study. MLS v4.2x ozone profiles are very similar to v2.2 in the stratosphere and above, so the validation results for v2.2 product generally hold for the v4.2x product (Livesey et al., 2018). MLS v4.2x ozone profiles are retrieved on 12 surfaces per decade between 316 hPa and 1 hPa, twice as fine a resolution as that used in v2.2. There are several improvements in MLS v4.2x ozone retrievals. The high bias of MLS v2.2 ozone at 215 hPa is reduced in v4.2x. Compared to v3.3 ozone, v4.2x reduces the vertical oscillation behavior in the tropical upper troposphere and lower stratosphere (UT/LS) regions (although some oscillations still exist). The sensitivity of retrieved ozone to thick clouds is also improved in the v4.2x product. In this study, MLS v4.2x ozone data were screened based on the recommendations of Livesey et al. (2018).

3.2 OSIRIS

The Optical Spectrograph and InfraRed Imaging System (OSIRIS) on board the Odin satellite has been taking limb scattered measurements of the atmosphere from November 2001 to present. It operates at wavelengths of 280–810 nm, with a spectral resolution of ~1 nm (Llewellyn et al., 2004; McLinden et al., 2012). The Odin satellite has a polar orbit with equatorial crossing local times at ~6:00 PM (ascending node), and at 6:00 AM (descending node). OSIRIS can provide near global coverages (up to 82°) near the equinoxes, sunlit summer hemisphere and no coverage of mid to high latitude winter hemisphere.

The OSIRIS SaskMART v5.0x ozone data are retrieved using the multiplicative algebraic reconstruction technique (MART) (Degenstein et al., 2009; Roth et al., 2007), and the SASKTRAN spherical radiative transfer model (Bourassa et al., 2008, Zawada et al., 2015). The retrieval algorithm simultaneously uses and merges information from UV and VIS radiances. Ozone number density, NO_2, aerosol extinctions and albedo are retrieved from 60 km down to cloud tops (or 10 km during absence of clouds) with a vertical resolution of ~2 km at low altitudes. The resolution decreases toward higher altitudes and reaches ~3 km at 50 km.

Through inter-comparisons with other satellite and in-situ measurements, the OSIRIS ozone data show good agreement (within 5%) with correlative measurements for altitudes above 20 km. Between 20 km and the tropopause OSIRIS shows negative biases of ~5–20% for latitudes between 40°S and 40°N (Adams et al., 2014). It was also found that OSIRIS ozone biases depend on the OSIRIS optics temperature, retrieved aerosols, and albedo. The latest OSIRIS v5.10 ozone data, with a drift correction of sensor pointing bias, are used in this study. The drift in previous OSIRIS v5.07 ozone data (Hubert et al., 2016) is attributed to a changing bias in the procedure to determine the tangent altitudes of limb radiance profiles (Bourassa et al., 2018). There is no further filtering applied to OSIRIS data in this study since the OSIRIS v5.10 ozone profiles have been screened for outliers, based on the techniques described by Adams et al. (2013), prior to its distribution to the public.

3.3 ACE-FTS

The Atmospheric Chemistry Experiment – Fourier Transform Spectrometer (ACE-FTS) is a solar occultation instrument that records spectra between 2.2 and 13.3 μm (750–4400 cm^{-1}) at a high spectral resolution of 0.02 cm^{-1} (Bernath et al., 2005, 2017). ACE-FTS was launched on the SCISAT satellite in August 2003. Measurements are taken during each sunrise and sunset per orbit. ACE-FTS measurements are taken up to 30 times per day at sunrise and sunset. The volume mixing ratios of ozone and other trace gases as well as temperature and pressure are retrieved from cloud tops to ~100 km by a modified global fit approach based on the Levenberg-Marquardt nonlinear least-squares method (Boone et al., 2005). The final results are provided on the measurement (tangent height) grid, with vertical resolution of 3–4 km, and interpolated to a 1 km interval using a piecewise quadratic method.

When compared with Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) and Aura MLS, the ACE-FTS v3.5 ozone generally agree within 5% in the middle stratosphere (~20–45 km), and exhibit a positive bias of ~10–20% in the upper stratosphere and lower mesosphere (Sheese et al., 2017). ACE-FTS also tends to show negative bias with respect to MIPAS and MLS below 20 km. The negative bias increases with decreasing altitudes, and reaches ~20–30% near 10 km.

The ACE-FTS version 3.5 data extend from Feb. 2004 to March 2013. A new version number (version 3.6) is used for data onward when the version 3.5 processor was ported from a Unix to Linux based system. Although the ACE-FTS team just released version 4.0 data, we used version 3.6 data because version 3.5/3.6 data are still the recommended data set for scientific and validation studies at the time of writing. Data quality flags based on Sheese et al. (2015) are provided in version 3.5/3.6 netCDF files. All ACE-FTS data with a non-zero flag value were excluded from this study (ACE-FTS data usage guide and file description, 2017).

3.4 OMPS LP

The Ozone Mapping and Profiler Suite (OMPS) was launched in October 2011 on board the Suomi National Polar-orbiting Partnership (NPP) satellite. OMPS consists of three ozone-acquiring sensors (Flynn et al., 2006) designed to provide profile and total ozone measurements. All three sensors measure scattered solar radiances in overlapping spectral ranges and scan the same air masses within 10 min (Kramarova et al., 2014). The nadir module combines two sensors, the Total Column Nadir Mapper (TC-NM) for measuring total column ozone and the Nadir Profiler (NP) for ozone vertical profiles. The Limb Profiler (LP) module is designed to measure vertical profiles of ozone with higher vertical resolution (~2–3 km) from the upper troposphere to the mesosphere. In this study, we will use OMPS ozone profile products from the Limb Profiler (OMPS LP).

The OMPS LP sensor is based on principals tested in the 1990s by flying the Shuttle Ozone Limb Sounding Experiment on two space shuttle missions, STS-87 and STS-107 (Flittner et al., 2000; McPeters et al., 2000). OMPS LP measures solar radiances scattered from the atmospheric limb in UV and VIS spectral ranges to retrieve ozone profiles with a high vertical resolution. The OMPS LP algorithm retrieves ozone profiles independently from UV and VIS measurements using wavelengths pairs in the UV range and triplets in the VIS range (Rault and Loughman, 2013). Measured radiances are first normalized with radiances measured at 55.5 km and 40.5 km for UV and VIS retrievals respectively. In this study we use the most recent version

2.5 that was described and validated in Kramarova et al. (2018). Comparisons of ozone profiles derived from OMPS LP with MLS, OSIRIS and ACE-FTS demonstrated that between 18 and 42 km the mean biases are within ±10%, with the exception of the northern high latitudes where larger negative biases are observed between 20 and 32 km due to a thermal sensitivity issue (Kramarova et al., 2018). In the upper stratosphere and lower mesosphere (> 43 km) OMPS LP tends to have a negative bias against Aura MLS, ACE-FTS and OSIRIS instruments. In the UTLS below 15–18 km, especially in the tropics, negative biases increase up to ~30%. A positive drift of 0.5% yr-1 against MLS and OSIRIS was found that was more pronounced at altitudes above 35 km. Such a pattern is consistent with a possible 100 m drift in the LP sensor pointing detected in the analysis of LP radiances (Kramarova et al., 2018).

3.5 Ozonesondes

Ozonesondes are balloon-borne in situ instruments that can provide ozone profiles from the surface to the middle atmosphere (~30–35 km) with a high vertical resolution (~100–150 m). When standard operating procedures are followed, the three most commonly used sonde types produce consistent results. For altitudes between the tropopause and ~28 km, the systematic biases are less than 5% with precision better than 3% (Smit and ASOPOS panel, 2014). At higher and lower altitudes, the ozonesonde data quality degrades and the differences between different sonde types become larger. In the troposphere, the ECC type sondes have the best quality with estimated accuracy of 5–7 % and a precision of 3–5 % (Smit and ASOPOS panel, 2014). Ozonesonde data from the Southern Hemisphere Additional Ozonesondes (SHADOZ) network (Witte et al., 2017; Thompson et al., 2017), World Ozone and Ultraviolet Radiation Data Center (WOUDC, https://woudc.org), and National Oceanic & Atmospheric Administration (NOAA) (https://www.esrl.noaa.gov/gmd/ozwv/ozsondes/) are used to evaluate the SAGE III/ISS data. Ozonesonde stations used in this study can be seen in Table 1 in section 5.3.

3.6 Stratospheric ozone lidar

The Differential Absorption Lidar (DIAL) is a powerful technique to measure the vertical distribution of ozone in the stratosphere and troposphere with a vertical resolution of several hundred meters near tropopause to 3–5 km in the upper stratosphere (Godin et al., 1999). This technique uses two (or more) laser wavelengths which are chosen such that one has strong ozone absorption and the other has much lower absorption. The concentration of ozone is retrieved by measuring the different absorptions of the backscatter data at two wavelengths. The choice of selected laser wavelengths depends on whether the measurement is intended for the troposphere or stratosphere (Megie et al., 1985).

We used stratospheric ozone lidars in the Network for the Detection of Atmospheric Composition Change (NDACC, http://www.ndacc.org), which provide ozone number density vs geometric altitude profiles between the tropopause and 45–50 km. The precision of NDACC ozone lidar is ~1% up to 30 km, 2–5% at 40 km and 5–25% at 50 km (Keckhut et al., 2004). Intercomparisons of different processing algorithms within the NDACC network indicate that the biases in retrieved ozone are ~2% for altitudes between 20 and 35 km, and increase to ~5–10% at other altitudes (Keckhut et al., 2004). Those larger biases are due to lower signal to noise ratio or saturation of the detectors. By comparing lidars with ozonesondes and satellites, Nair et al. (2012) also showed biases less than ±5% in the lidar for altitudes between 20 and 40 km. We used data from five stratospheric ozone lidars in the NDACC networks (Table 2) that provide overlapping data with SAGE III/ISS in this study.

4 Methodology

To evaluate the quality of SAGE III/ISS ozone data with correlative measurements, we need to consider uncertainties from (1) spatial/temporal differences (mismatch), (2) different horizontal and vertical resolutions (smoothing), and (3) converting ozone profiles to different coordinates (auxiliary) (von Clarmann, 2006; Hubert et al., 2016). Common coincidence criteria are used to minimize the effect of spatial and temporal differences (i.e., mismatch error) between SAGE III/ISS and correlative measurements. For satellite comparisons, coincident profiles need to be on the same date with latitude difference less than ±2° and distance between them less than 1000 km. When there is more than one correlative ozone profile with a SAGE III/ISS ozone profile, the closest one in space is used. For comparisons with ground-based measurements, larger coincidence criteria are used, with temporal differences of ±24 hours, and spatial differences of ±5° in latitude and distance less than 1000 km. The larger coincidence criteria for ground-based measurements is to ensure there are enough correlative data to characterize the bias and precision of SAGE III ozone while minimizing the effects due to temporal and spatial variabilities

There is no good way to minimize the effect of different horizontal resolutions between instruments (e.g., satellite measurement vs ozonesondes); the ozone profiles from instruments with finer vertical resolution, however, can be smoothed before comparison to minimize the biases due to different vertical resolutions. For comparisons between SAGE III and MLS, the SAGE III ozone profiles were interpolated to MLS levels by using a least squares linear fit method recommended by the MLS science team (Livesey et al., 2018). The MLS averaging kernels and a priori profiles were not applied to interpolated SAGE III ozone profiles (e.g., Rodgers and Connor, 2003), because the effect of further smoothing by applying MLS averaging kernels has been shown to be very small (e.g., Adams et al., 2014). This is because the MLS averaging kernels are close to delta functions (sharply peaked and with vertical resolution comparable to the MLS retrieved profile level spacing). Finally, the MLS and SAGE III ozone number density profiles at varying geometric altitudes were linearly interpolated to every 1 km interval.

ACE-FTS ozone has a vertical resolution of ~3–4 km. Ozone data are retrieved at tangent altitudes, with vertical spacing of ~1.5 km at lower altitudes increasing to ~6 km in the mesosphere. Retrieved ozone profiles are then interpolated to a 1 km interval by using a piecewise quadratic method. To minimize the effect of different vertical resolutions, the SAGE III/ISS ozone profiles were first smoothed at ACE-FTS retrieved tangent altitudes by using a weighted Gaussian distribution function with a full width half maximum (FWHM) that approximates the vertical resolution of ACE-FTS (Kar et al., 2007; Sheese et al., 2017). The smoothed SAGE III ozone profiles were subsequently interpolated to a 1 km grid before comparing with ACE-FTS data. Alternatively, the SAGE III ozone profiles can be smoothed by a triangular function with full width at the bases equal to the vertical resolution of ACE-FTS (Dupuy et al., 2009). It has been found that the choice of smoothing function (e.g., triangular or Gaussian function) does not introduce systematic bias when comparing ozone profiles with different vertical resolutions although it may introduce a slight difference in random errors (Hubert et al., 2016). The OSIRIS and OMPS LP have similar vertical resolutions of ~2 km in most of stratosphere and ~3 km in the upper stratosphere and lower mesosphere. Similarly, the SAGE III ozone profiles were smoothed by the Gaussian distribution with FWHM corresponding to the vertical resolution of OSIRIS and OMPS LP. The ground-based ozonesondes and lidar (in the UT/LS regions) have better vertical resolution than SAGE III. Correlative ozone profiles from ozonesondes and lidar, therefore, were smoothed according to the SAGE III resolution (~1 km) before further inter-comparisons.

In order to compare collocated ozone profiles between SAGE III/ISS and correlative measurements, those profiles need to be on the same coordinate. Due to an altitude registration error in current SAGE III/ISS v5.1 temperature and pressure data (see discussion in section 2), we used ozone in the SAGE III native retrieval coordinate, number density on geometric altitude. Ozone profiles in different coordinates (e.g., mixing ratio on pressure or mixing ratio on geometric altitude) from Aura MLS, ACE-FTS and ozonesondes were converted to SAGE III native coordinates by using their own observed temperature data, except for Aura MLS. Although Aura MLS also measures temperatures and retrieves geopotential heights (GPH) along with each ozone profile, there are seasonally and latitudinally-repeating systematic errors in GPH (Livesey et al., 2018). The assimilated meteorology fields from the second Modern-Era Retrospective analysis for Research and Applications (MERRA-2) (GMAO, 2015), therefore, were used. The MERRA-2 temperatures (with resolution of 0.625° in longitude, 0.5° in latitude, 72 model layers from surface to 0.01 hPa, and every 3 hours), were first interpolated to MLS locations and pressure levels. The geopotential heights (GPH) at MLS pressure levels were then derived by using the hypsometric equation and reference altitude from MERRA-2. With interpolated MERRA-2 temperatures and geopotential heights corresponding to the MLS grid, the original MLS ozone profiles can be converted to number densities on geometric altitudes.

To assess the overall quality of SAGE III/ISS ozone data with correlative measurements, we use the following two metrics: the mean relative differences and the standard deviations of relative differences. The mean bias (relative difference), $\overline{D(z)}$, in percentage is defined as

$$\overline{D(z)} = 100\ X\ \frac{1}{n(z)} \sum_{i=1}^{n(z)} \frac{x_i^s(z) - x_i^c(z)}{x_i^c\ (z)}$$

where n(z) is the number of coincident profiles, $x^s(z)$ and $x^c(z)$ are ozone number density at a particular altitude (z) from SAGE III and correlative measurement, respectively. The SAGE III reported uncertainty along with retrieved ozone contains random errors from three primary sources: (1) line-of-sight optical depth measurement error, (2) estimated Rayleigh scattering, and (3) uncertainty associated with removal of contributions from interfering gases and aerosol (SAGE III ATBD, 2002). In order to verify SAGE III reported random errors and provide additional information regarding the significance of the bias and the upper limit of the precision of SAGE III/ISS ozone data, we calculate the standard deviation of bias-corrected differences. The de-biased standard deviation is a measure of the combined precision of instruments that are being compared (von Clarmann, 2006), and is represented as

$$\sigma(z) = \sqrt{\frac{1}{(n(z)-1)} \sum_{i=1}^{n(z)} \left(D_i(z) - \overline{D(z)}\right)^2}$$

where n(z) is the number of coincidences, $D_i(z)$ is the relative difference for the *i*th coincident pair, and $\overline{D(z)}$ is the mean relative difference at a particular altitude (z).

5 Results

5.1 Comparisons of the SAGE III/ISS solar ozone between AO3 and MLR algorithms and between sunrise and sunset measurements

As mentioned earlier in section 2, SAGE III/ISS produces two solar ozone products based on the ozone absorption in the Chappuis band by two different retrieval algorithms. The mean differences and reported uncertainties from these two ozone products are shown in Figure 4. The mean differences between AO3 and MLR ozone are negligible between 20 and 50 km, but become larger toward higher or lower altitudes. For altitudes above 50 km, the MLR ozone shows increasing high biases, reaching ~20–30% at 60 km. In the lower stratosphere below 20 km the MLR ozone also shows increasing high biases (with decreasing altitudes), as large as ~20% at 10 km. As expected both MLR and AO3 ozone show the smallest uncertainties around the ozone peak area. The uncertainties become larger toward higher and lower altitudes where there is less ozone or larger contributions from other interfering trace gases and aerosol in the retrieval algorithms. The reported uncertainties in MLR ozone are a few percent between 20 and 30 km. They become larger than 100% for altitudes above ~55 km and below 10 km. The mean uncertainties in AO3 ozone are approximately 2–3 times smaller than those of MLR ozone.

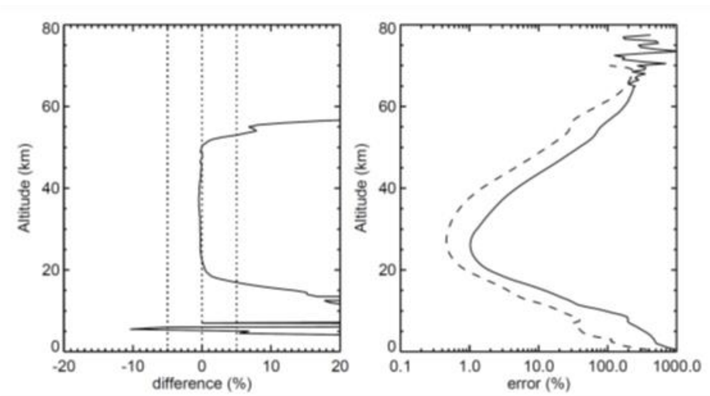


Figure 4: Mean relative differences between SAGE III/ISS MLR and AO3 solar ozone data (left). Percentage differences are represented as (MLR-AO3)/AO3*100%. Mean reported uncertainties in MLR (solid) and AO3 (dashed line) ozone profiles (right). Mean differences and uncertainties are based on all retrieved ozone profiles between June 2017 and May 2019.

By using the residual analysis detailed in Damadeo et al. (2014), we can get an assessment of random errors in AO3 and MLR ozone. The time series of observed ozone (averaged within a specific temporal/spatial window) contains information about the natural variability and instrument uncertainties. The natural variability of ozone can be approximated by a regression model with predictors for seasonal cycle, long term trend, quasi-biennial oscillation (QBO), solar cycle, etc. The spread of the residuals from the regression of observed ozone data can be used to ascertain the quality of the regression model and observed data itself. The total residuals consist of the correlated and uncorrelated residuals. The correlated residuals come from autocorrelation within the data and typically represent the natural variability that is not well represented by the regression model. Uncorrelated residuals represent a combination of measurement uncertainty and geophysical variability that is not well-sampled (e.g., zonal variability within the daily zonal means used for this analysis). For the purpose of this validation study, we only care to look at the uncorrelated residuals as an indication of data quality or precision. Since the choice of regression model has little bearing on the uncorrelated residuals, a rather simplistic model consisting only of a seasonal cycle was used for this analysis, applied to all SAGE III/ISS data between June 2017 and May 2019.

The spreads of the uncorrelated residuals from the regression of AO3 and MLR ozone are shown in Figure 5, which can provide an estimate of the upper limit of uncertainties in both datasets. This is an upper limit because zonal variability within each daily zonal mean used for this analysis will also increase the uncorrelated residuals. However, since the sampling is identical between the two data products, a direct comparison of the uncorrelated residuals yields information about the intrinsic data quality of each data product independent of any correlative source instrument. We can see that the uncorrelated residuals are similar throughout most of the stratosphere between the two products (~1–3%). The MLR ozone, however, is significantly noisier than the AO3 product both in the upper-most stratosphere and mesosphere as well as in the lowermost stratosphere and troposphere. These results are similar to those from a study (Wang et al., 2006) of SAGE III/M3M data using comparisons with other correlative data sets. While useful as an independent comparison of the relative data quality of the two data products, evaluating the statistics of the uncertainties (or precisions) for individual profiles via comparisons of correlative measurements can help mitigate the impact of the dynamical variability in the regression sample size (i.e., a daily zonal mean) and will be evaluated in later sections.

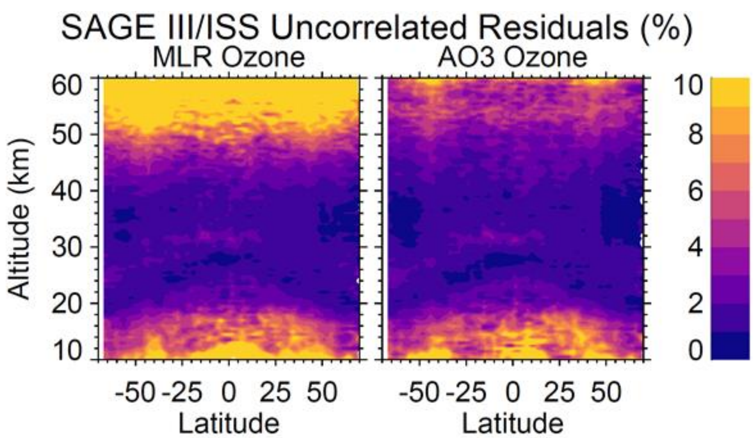


Figure 5: Standard deviations of the uncorrelated residuals in percentage as a function of latitude and altitude from the regression of SAGE III/ISS MLR (left) and AO3 (right) ozone data.

It has been reported that there is a difference in observed ozone values between sunrise and sunset from solar occultation instruments (Wang et al., 1996; Brühl et al., 1996; Kyrölä et al., 2013; Sakazaki et al., 2015). Measurements from the Halogen Occultation Experiment (HALOE), ACE–FTS, and Superconducting Submillimeter-Wave Limb-Emission Sounder (SMILES) show that the sunset values are higher than sunrise by 3–5% between 40 and 50 km (Sakazaki et al., 2015). SAGE II shows similar features as HALOE, ACE-FTS and SMILES, but the magnitude of sunrise/sunset differences is approximately twice as large as those from other satellites, especially in the tropics during January (Wang et al., 1996). Based on observations from SMILES and the Specified Dynamic version of the Whole Atmosphere Community Climate Model (SD-WACCM), Sakazaki et al. (2013, 2015) attributes the observed sunrise/sunset differences in the upper stratosphere to the vertical transport of atmospheric tidal winds, which reach a maximum in the tropics and during the winter season (Dec. to Feb.). The reason for the larger sunrise/sunset differences in SAGE II is not clear, but it is worth investigating whether a similar situation occurs in the SAGE III/ISS ozone data.

To investigate the sunrise/sunset differences in SAGE III/ISS retrieved ozone, we used two different methods. The first one is to apply the regression model described in Damadeo et al. (2018) to both SAGE II and SAGE III/ISS data simultaneously to derive the mean difference between sunrise and sunset data. There is currently insufficient sampling orthogonality within the SAGE III/ISS data set to differentiate seasonal variability from diurnal variability, so including SAGE II data (given its own diurnal cycle) helps constrain this. The lack of overlap between the two data sets is accounted for by considering SAGE III/ISS as an extension of the SAGE II product, which is acceptable since we are not interested in trend results in this work. The results are shown in Figure 6. Both AO3 and MLR ozone show similar results, with sunset values higher than sunrise by ~5–10% in the upper stratosphere, though the pattern of differences is more coherent for the AO3 product than the MLR product. The sunrise values, however, become slightly larger than sunset in the lower stratosphere below 25 km. The sunrise/sunset differences are also larger in the tropics than mid-latitudes. The vertical and latitudinal distributions of sunrise/sunset differences are consistent with the dynamical variations from atmospheric tidal winds (Sakazaki et al., 2013, 2015).

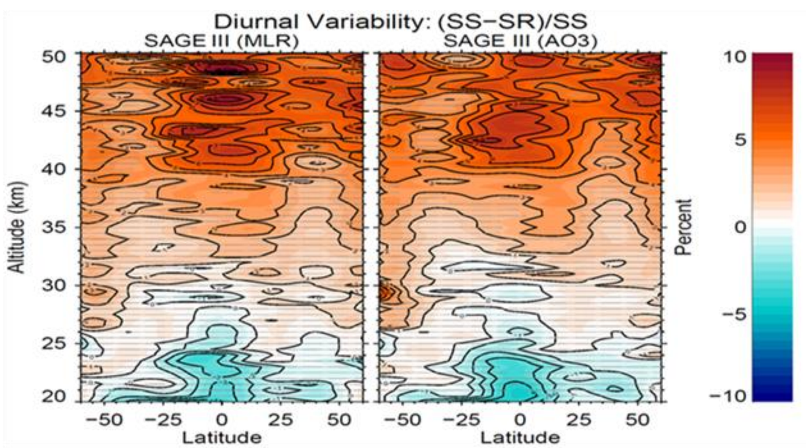


Figure 6: Mean differences between SAGE III/ISS sunrise (SR) and sunset (SS) ozone values from regression model analysis. Results from both MLR (left) and AO3 (right) algorithms are shown. The percentage difference is expressed as (SS-SR)/SS*100%. The stippling denotes regions that are not statistically significant at the 2-sigma level.

We also used Aura MLS as transfer standard to evaluate the differences between SAGE III/ISS sunrise and sunset measurements. Figure 7 shows comparison results between SAGE III/ISS AO3 ozone, separated by sunrise or sunset, and coincident Aura MLS nighttime measurements. As shown in Figure 7, SAGE III/ISS sunset values are systematically higher than sunrise values by ~5–8% for altitudes between 40 and 55 km. In the lower stratosphere between the tropopause and ~25 km, the sunrise values become slightly larger (less than 5%) than sunset values. Similar results were also found by using MLR ozone compared against collocated Aura MLS data, or comparing sunrise and sunset measurements directly (e.g., Wang et al., 1996) when they were observed on the same dates and approximately at the same locations (e.g., ±1° latitude, ±5° longitude, figures not shown). The reason for the large sunrise/sunset difference in SAGE retrieved ozone in the upper stratosphere is not clear, but since it occurs in both SAGE II and SAGE III/ISS, it could relate to the retrieval algorithm and needs further investigation.

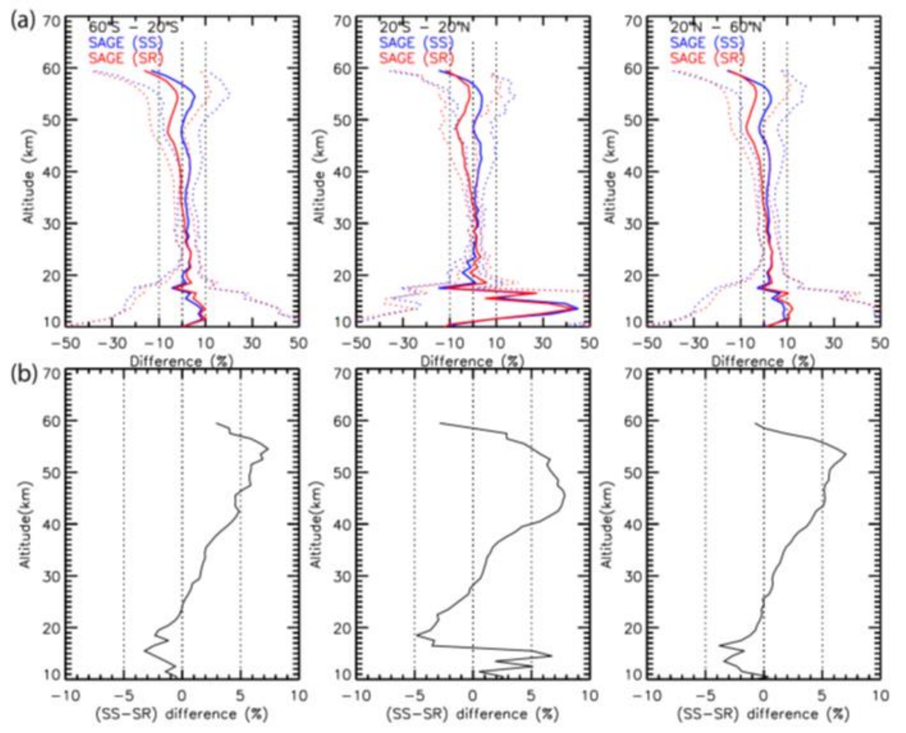


Figure 7: (a) The mean percentage differences (solid line) and standard deviations (dotted line) between SAGE III/ISS AO3 ozone and coincident Aura MLS nighttime measurements between June 2017 and May 2019 in three latitude bands, 60°–20°S, 20°S–20°N, and 20°–60°N. The means and standard deviations of relative differences are separated into SAGE sunrise (red) and sunset (blue) data. (b) The relative differences between SAGE III/ISS sunrise and sunset measurements by using coincident Aura MLS as a transfer standard. The percentage difference is represented as (SS-SR)/MLS*100.

5.2 Comparisons between SAGE III/ISS and other satellites

Among the correlative satellite instruments, the Aura MLS provides the most comprehensive global coverages (from 82°S–82°N) each day with the equatorial crossing time at ~1:45 am and 1:45 pm. The comparisons between SAGE III/ISS retrieved stratospheric ozone products and Aura MLS nighttime measurements are shown in Figure 8. We used MLS nighttime measurements to minimize the effect of ozone diurnal cycle on the differences between SAGE III and MLS, since the SAGE III measurements occur during sunrise and sunset which in general yield ozone values that are closer to nighttime than daytime ozone (Takatoshi et al., 2013; Parrish et al., 2014).

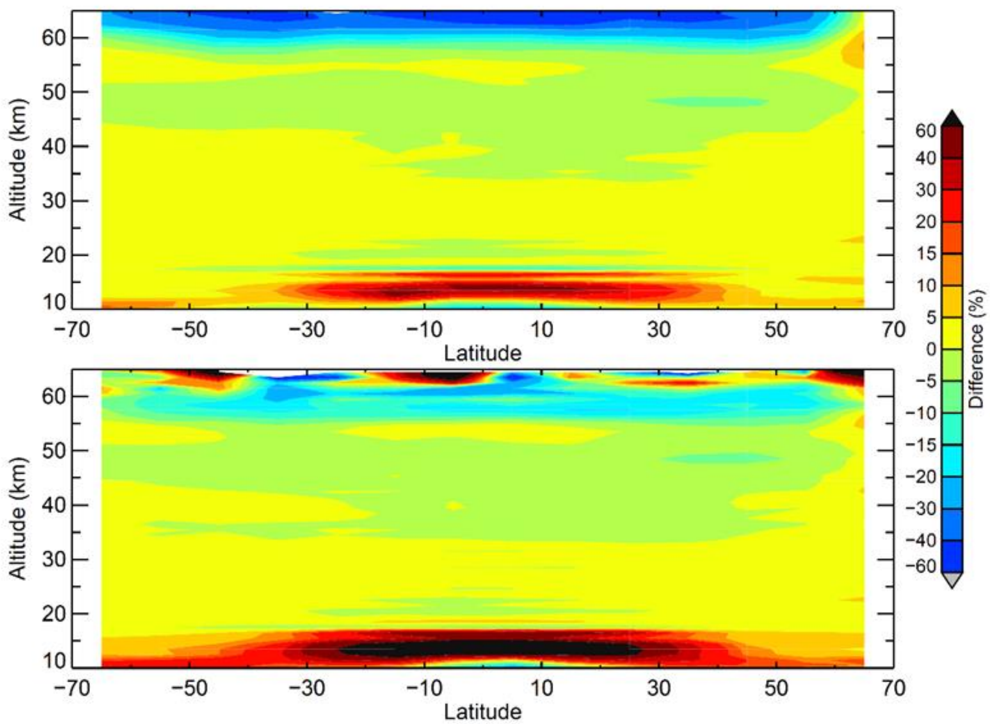


Figure 8: Mean differences between Aura MLS night time measurements and ozone retrieved from SAGE III/ISS AO3 (top) and MLR (bottom) algorithm as a function of latitude and altitude. Positive differences (in percentage) indicate the SAGE III/ISS ozone values are higher than Aura MLS.

SAGE III/ISS AO3 ozone shows very good agreement with Aura MLS for altitudes between ~20 and 55 km, with differences less than 5%. The differences become larger toward the lower stratosphere and upper troposphere and reach ~10% near the tropopause, with SAGE III ozone values higher than MLS. Above 55 km the SAGE III ozone values are systematically lower than those from Aura MLS with negative biases of ~10% at 60 km and 40–60% at 65 km. The larger biases (e.g., >40%) between SAGE III and Aura MLS in the mesosphere cannot be completely explained by the ozone diurnal cycle (e.g., sunrise/sunset vs nighttime) (Parrish et al., 2014). These biases could result from errors in the MERRA-2 temperature data in the mesosphere and/or deficiencies in SAGE III AO3 retrieval algorithm. We used MERRA-2 data to convert MLS ozone from mixing ratio and pressure coordinates to SAGE's native number density and geometric altitude coordinates. Any systematic error in auxiliary temperature and pressure data can lead to errors in converted MLS ozone profiles, but the evaluation of MERRA-2 temperature data in the mesosphere is outside the scope of this paper. Since the SAGE III AO3 ozone product is retrieved using the Chappuis band, the weakly attenuated signals in the mesosphere could yield degraded results in that region. Instead, the SAGE III/ISS MES algorithm may provide more information for mesospheric ozone after correcting for the stray light problem.

The SAGE III/ISS MLR ozone shows similar features as AO3 when compared against Aura MLS. The relative differences with MLS are less than 5% between 20 and 55 km for all latitudes. The differences, however, become larger at higher and lower altitudes. In the lower mesosphere above 60 km, SAGE III MLR ozone shows positive biases of 20% or more for some latitudes. This is contrary to what is expected from the ozone diurnal cycle. SAGE III MLR ozone also shows positive biases in the lower stratosphere, with mean differences of approximately 10–30% in the middle to high latitudes and greater than 60% near the tropical tropopause.

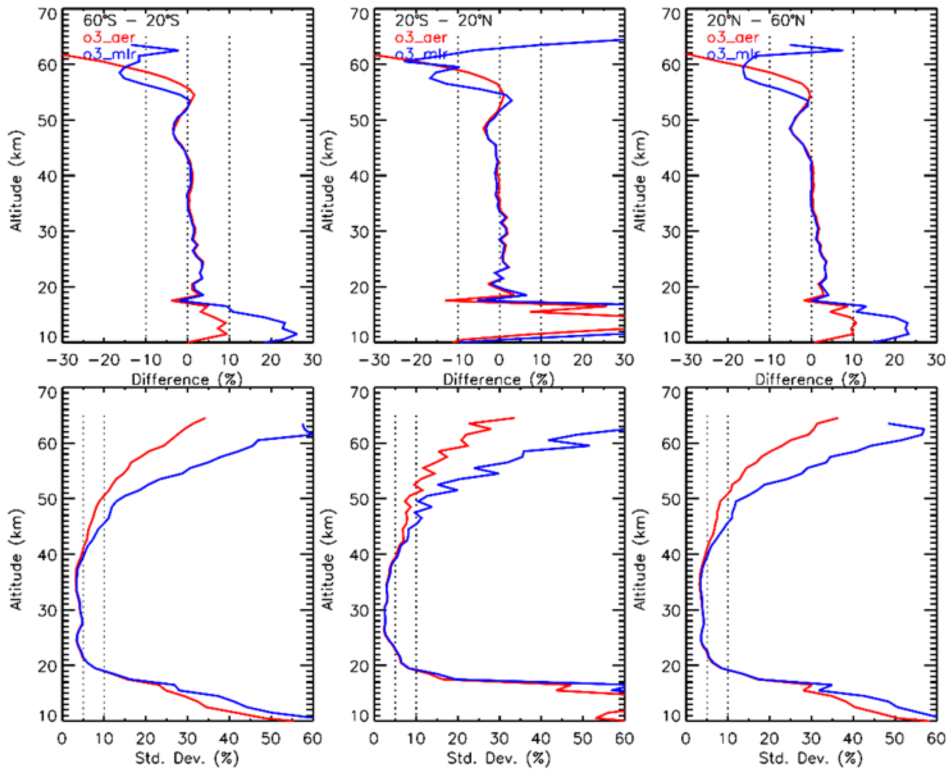


Figure 9: Mean differences (top panel) and standard deviations (bottom panel) between collocated Aura MLS and SAGE III/ISS ozone from AO3 (red) and MLR (blue) retrieval algorithm. Differences and standard deviations are derived in three broad latitude bands, 20°–60°S, 20°S–20°N, 20°–60°N, and represented as a percentage.

The mean relative differences and standard deviations between SAGE III/ISS AO3 and MLR ozone against Aura MLS are summarized in Figure 9. Between the two SAGE III retrieved solar ozone products, the AO3 shows overall better accuracy and precision than MLR ozone. The systematic biases in AO3 ozone are less than 3% from ~15 km to 55 km in the mid-latitudes and ~20 km to 55 km in the tropics. The biases increase with decreasing altitudes and reach ~10% near the tropopause. The differences between SAGE III AO3 and MLS also become larger for altitudes above 55 km due to an increase of the ozone diurnal cycle. The SAGE III/MLS differences oscillate with altitude in the lower stratosphere and upper troposphere (UT/LS) especially in the tropics. This mainly results from Aura MLS which reports ozone on a slightly finer vertical grid than its actual vertical resolution in that region (Livesey et al., 2018). SAGE III MLR ozone shows similar biases as AO3 for altitudes between 20 and 50 km, but the biases become larger outside those altitudes. This is consistent with the earlier results of direct comparisons between SAGE III AO3 and MLR ozone data (Figure 4). The MLR retrieved ozone also shows larger uncertainties than AO3 in the upper stratosphere and lower mesosphere (above

40 km) and in the UT/LS regions (below 20 km), as indicated by the larger standard deviations in Figure 9, which is consistent with results from the independent regression analysis shown in Figure 5. Similar features are also found in comparisons between SAGE III MLR ozone and other satellites (figures not shown). Because of the larger uncertainties and biases in MLR ozone for altitudes above 50 and below 20 km, we recommend using SAGE III AO3 ozone for scientific studies. In the following sections, we will just focus on validation results for SAGE III AO3 ozone.

The comparisons between SAGE III/ISS AO3 ozone and ACE-FTS, OSIRIS, OMPS LP are shown in Figure 10. Both SAGE III and ACE-FTS use solar occultation techniques to measure ozone. Due to limitation of the orbit geometry, there are no collocated SAGE III/ACE-FTS ozone profiles in the regions between equator and 20^0S, and poleward of 60^0S. The differences between SAGE III and ACE-FTS are in general within 5% between 15 and 45 km. Above 45 km SAGE III shows a negative bias of ~10%. Below 15–20 km, SAGE III values become larger than ACE-FTS by 10–20% in mid-latitudes (Figure 10a). This is consistent with an earlier study, which shows ACE-FTS v3.5 ozone has a positive bias of ~10–20% in the upper stratosphere and mesosphere (>45 km), and negative bias of 20–30% in the UT/LS (Sheese et al., 2017). SAGE III and OSIRIS show the best agreement between 20 and 50 km. The differences are generally within 5%, except in the northern hemisphere around 30 km, where the differences are slightly larger than 5% (Figure 10b). The reason for this hemispheric difference is not known but it doesn't occur in the comparisons between SAGE III against Aura MLS and ACE-FTS.

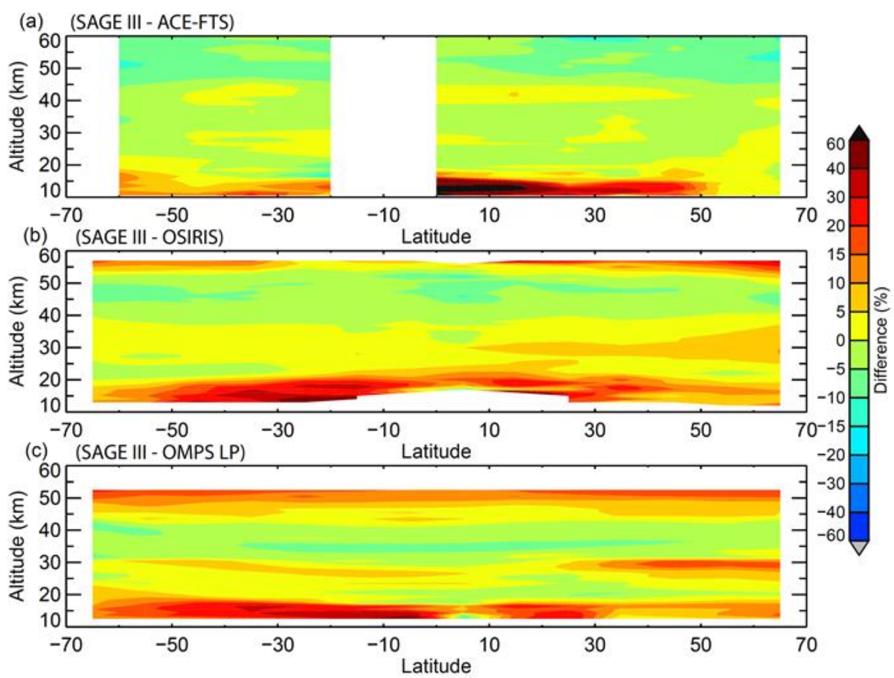


Figure 10: Mean differences between SAGE III/ISS AO3 ozone and correlative (a) ACE-FTS (b) OSIRIS and (c) OMPS LP measurements. Differences are represented as (SAGE-other)/other* 100%.

We can see that all satellite measurements show very good agreement, with differences less than 5%, in the middle stratosphere, except for OMPS LP in the northern mid-latitudes near

28–31 km (Figures 10c, 11). This is due to the thermal sensitivity problem in the OMPS LP instrument, which causes negative biases of 10–15% in retrieved ozone from the visible spectral ranges (Kramarova et al., 2018). In the upper stratosphere and lower mesosphere (e.g., above ~45 km) the differences between SAGE III and other correlative measurements become larger. This is due to the ozone diurnal cycle and/or known biases in those datasets. For example, SAGE III shows negative biases of 5–10% relative to ACE-FTS in the upper stratosphere and lower mesosphere. This is due to known positive biases in ACE-FTS ozone in those regions (Sheese et al., 2017). SAGE III also shows altitude dependent high biases versus OMPS LP, with mean differences of ~5% at 45 km and ~15–20% at 52 km (Figure 11). This is an artifact resulting from the known low biases (~10%) in OMPS LP ozone in the upper stratosphere and lower mesosphere (Kramarova et al., 2018) and the ozone diurnal cycle. In the upper stratosphere and mesosphere, the ozone levels show a strong depletion during the daytime and recover at night. The OMPS LP measurements mainly occurs during daytime (e.g., at local solar time ~1:30 PM), while SAGE III takes measurements during sunrise and sunset when ozone values are closer to nighttime measurements. The day-night ozone differences are ~10% at 50 km and increase to ~60% at 65 km (Parish et al., 2014). The low biases in OMPS LP ozone for altitudes above 45 km, therefore, would be further enhanced by the ozone diurnal cycle when compared with SAGE III, and result in altitude-dependent structure as shown in Figure 11.

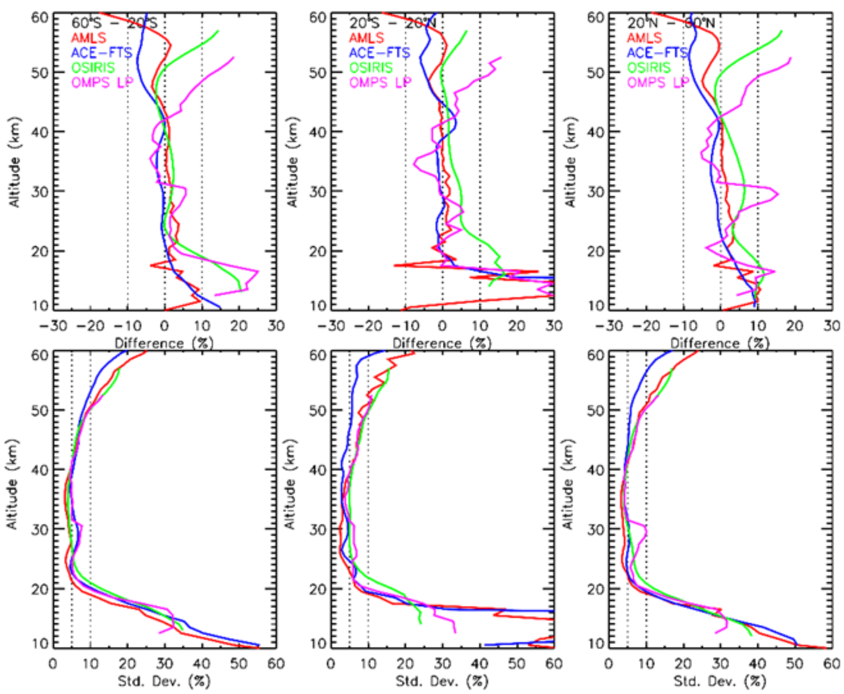


Figure 11: Mean differences (top) and standard deviations (bottom) between SAGE III/ISS AO3 against Aura MLS (red), ACE-FTS (blue), OSIRIS (green), and OMPS LP (pink) in three wide latitude bands.

The comparisons between SAGE III and OSIRIS ozone for altitudes above ~50 km show similar features (e.g., altitude-dependent biases) as those in SAGE III/OMPS LP comparisons. OSIRIS is on a sun synchronous satellite, which observes ozone mainly at local solar time between 6:30 and 7:30 am (closer to daytime ozone values). The observed differences between

SAGE III and OSIRIS for altitudes above 50 km are consistent with what we expect from day-night ozone differences. The effects of the ozone diurnal cycle on the comparisons between SAGE III and Aura MLS or ACE-FTS in the upper stratosphere and lower mesosphere are smaller. This is because MLS nighttime measurements (~1:45 am) were used in this study, and the ACE-FTS also makes measurements during local sunrise or sunset.

In the lower stratosphere and upper troposphere SAGE III ozone in general shows high biases against other correlative satellite measurements, with mean relative differences of ~5–10% against Aura MLS and ACE-FTS from 20 km down to the tropopause. Most, if not all, of this bias is likely the result of the O_4 spectroscopy problem discussed in section 2.2. The differences between SAGE III and OSIRIS and OMPS LP are larger (~10–20%) in the southern hemisphere mid-latitudes and in the tropics. This is most likely related to low biases in OSIRIS and OMPS LP ozone measurements in the UT/LS regions (Kramarova et al., 2018; Adams et al., 2014).

The standard deviations of relative differences between SAGE III and other satellite measurements, except ACE-FTS, show similar magnitudes and vertical structures. The smallest standard deviations of ~5% are found in the middle stratosphere (e.g., between 20 and 40 km). The standard deviations increase to ~10% at 50 km and ~20% at 60 km. The smaller standard deviations between SAGE III and ACE-FTS differences in the upper stratosphere and lower mesosphere are due to both instruments making observations during sunrise and sunset with smaller noise. Below 20 km the standard deviations also become larger. These increases result from both measurement uncertainties and mismatch (inexact coincidence) between SAGE III and other satellites. The lower stratosphere and upper troposphere is a challenging area for satellite ozone observations. SAGE III ozone in the UT/LS will be further evaluated by ground-based measurements in the following section.

5.3 Comparisons between SAGE III/ISS and ground-based measurements

The ozonesondes and stratospheric ozone lidars were used to further evaluate the SAGE III/ISS ozone in the UT/LS region. The geolocations and data sources of ozonesondes and lidar and number of coincident profiles found for each with SAGE III are listed in Table 1 and Table 2, respectively. For ozonesondes the tropical stations are mainly from the Southern Hemisphere ADditional OZonesondes (SHADOZ) network (Thompson et al., 2017; Witte et al., 2017). Although there are few coincident profiles (e.g., from 1 to 8) between SAGE III and individual ozonesonde stations in SHADOZ, the ozonesondes data have been processed with the same processing technique to minimize the inhomogeneities in ozonesonde data records. This enables us to group SHADOZ data in the tropics to provide better statistics for estimating SAGE III ozone biases in that region. Outside the tropical latitudes, ozonesondes from the WOUDC and NOAA Earth System Research Laboratory (ESRL) (Johnson et al., 2018) were used. There are five NDACC stratospheric ozone lidar stations that provide correlative measurements during the first two years of SAGE III operation (e.g., June 2017 to May 2019). Those stations are listed in Table 2.

Table 1 ozonesonde stations used in this study

Station	Latitude	Longitude	Data source	Coincident profiles
Hohenpeissenberg	47.80	11.00	WOUDC	53
Payerne	46.49	6.57	WOUDC	56
Trinidad Head	41.06	-124.15	NOAA	13
Boulder	39.95	-105.20	NOAA	21
Tsukuba	36.06	140.13	WOUDC	17
Huntsville	34.73	-86.85	NOAA	8
Hilo	19.40	-155.40	SHADOZ/NOAA	8
Costa Rica	9.94	-84.04	SHADOZ	5
Paramaribo	5.80	-55.20	SHADOZ	1
Kuala Lumpur	2.73	101.70	SHADOZ	7
Nairobi	-1.30	36.80	SHADOZ	6
Natal	-5.40	-35.40	SHADOZ	3
Ascension Is.	-7.56	-14.22	SHADOZ	5
Am. Samoa	-14.20	-170.60	SHADOZ/NOAA	3
Fiji	-18.10	178.40	SHADOZ	5
La Reunion Is.	-21.10	55.50	SHADOZ	1
Irene	-25.90	28.20	SHADOZ	4
Broadmeadows	-37.69	144.95	WOUDC	13
Lauder	-45.04	169.68	WOUDC	29
Macquarie Is.	-54.50	158.94	WOUDC	13

Table 2 Lidar data used in this study

Station	Latitude	Longitude	Data source	Coincident profiles
Hohenpeissenberg	47.80	11.00	NDACC	38
OHP	43.92	5.71	NDACC	46
Table Mtn.	34.5	-117.7	NDACC	45
Mauna Loa	19.47	-155.60	NDACC	30
Lauder	-45.04	169.68	NDACC	13

Due to limited coincident profiles between SAGE III and ground-based measurements the medians and spreads (defined as one-half of the differences between the 84th and 16th percentiles) of relative differences are better diagnostics to represent the biases and random errors in SAGE III retrieved ozone. The median and spread are the same as the mean and standard deviation when the statistical sample has a Gaussian distribution (e.g., Wang et al., 2002). The occurrence of outliers in the distribution, however, can lead to larger standard deviations and introduce a discrepancy between the mean and median for a non-Gaussian (asymmetric) distribution. For comparisons between SAGE III (or other satellites) and ground-based measurements, there could be outliers in the statistical sample due to anomalous data not being filtered out and/or large dynamic variability in the UT/LS (i.e., mismatch between SAGE III and ground-based measurements). The median and spread are more robust statistics to minimize the effect of outliers, especially for a distribution with small sample size.

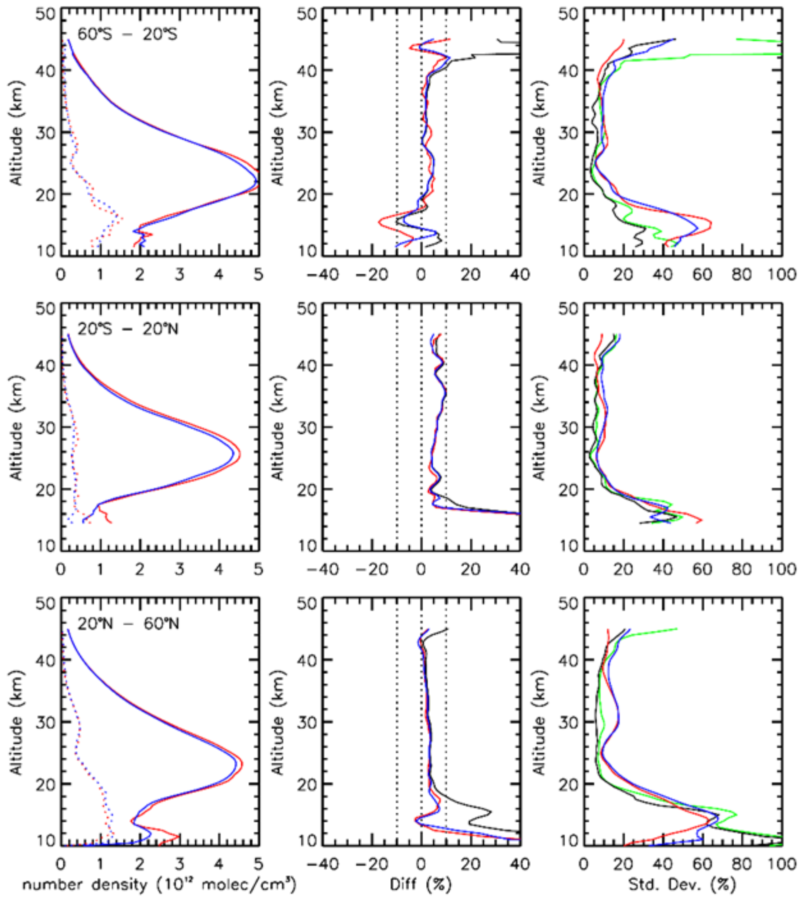


Figure 12: Comparisons between SAGE III/ISS and correlative lidar at three latitude bands, 60°S–20°S (top panel), 20°S–20°N (middle panel) and 20°N–60°N (bottom panel). The mean (solid lines) and standard deviation (dotted lines) of coincident SAGE III/ISS (red) and lidar (blue) ozone number density profiles are shown in the left panel. The relative percentage differences between SAGE III/ISS and lidar are shown in the middle panel. The mean and median of relative differences are indicated by the black and red colors, respectively. The blue lines indicates differences estimated from averaged ozone profiles (see text). In the right panel, the standard deviations of mean and 1-σ spreads of median differences are indicated by green and black lines, respectively. The standard deviations of coincident SAGE III/ISS (red) and lidar (blue) profiles are also shown.

The comparison results between SAGE III and lidar are shown in Figure 12. The analysis is performed by using all collocated profiles in three broad latitude bands, southern mid-latitudes (60°S–20°S), tropics (20°S–20°N), and northern mid-latitudes (20° N–60°N). There is only one lidar station, Lauder and Mauna Loa, located in the southern mid-latitude and tropics, respectively. For northern mid-latitudes, measurements from Hohenpeissenberg, Observatoire de Haute-Provence (OHP), and Table Mountain Facility are used. Both SAGE III and lidar show maximum ozone concentrations near 22–23 km in the mid-latitudes and 26–27 km in the tropics (Figure 12 left panel). The ozone variabilities indicated by the standard deviations generally increase from the upper stratosphere down to the lower stratosphere and upper troposphere. SAGE III and lidar observations show similar results with standard deviations between 10–20%

for altitudes between 20 and 40 km. The standard deviations increase to ~50–60 % in the UT/LS regions due to larger dynamic variability and smaller ozone amounts (Figure 12 right panel). The best agreements between SAGE III and lidar are found between 20 and 40 km. SAGE III shows a small positive bias of ~5% against all lidar observations except at Mauna Loa, where SAGE III ozone shows slightly larger high biases of ~5–10% between 30 and 40 km (Figure 12 middle panel). The reason for this is not clear, but SAGE III ozone is in good agreement (within 5%) with other satellites at the same altitude ranges in the tropics (Figure 11).

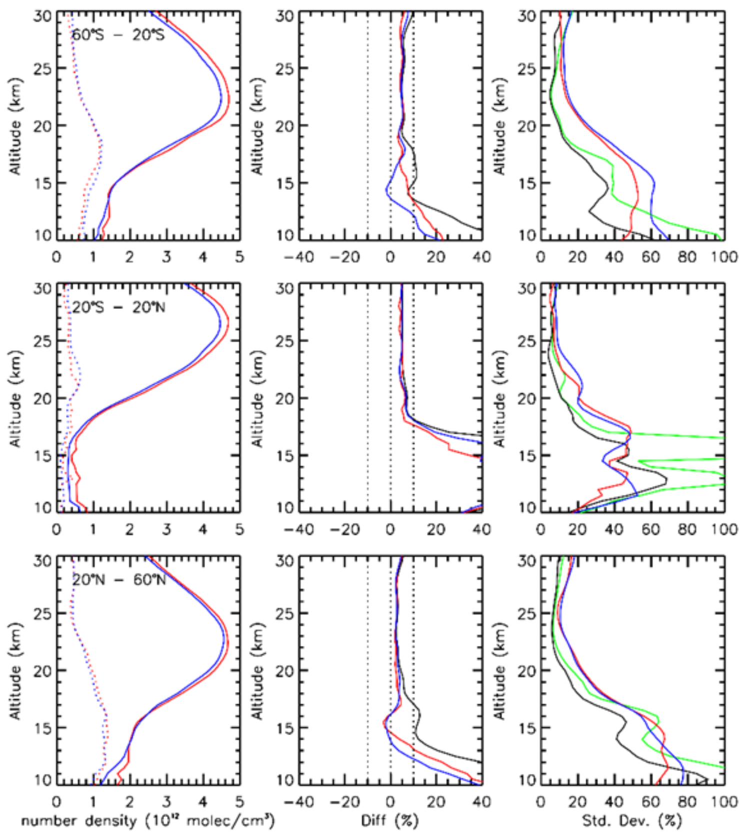


Figure 13: Similar to figure 12 but for comparisons between SAGE III/ISS and ozonesondes.

In the southern mid-latitudes above ~42 km, SAGE III and Lauder ozone lidar show mean differences of ~40% or larger and standard deviations greater than 60%. The median differences, however, are only ±10%. The larger mean differences and standard deviations, compared to medians and spreads, between SAGE III and Lauder in the upper stratosphere are due to outliers in the lidar measurements. Those outliers also contribute to larger standard deviations (by approximately a factor of 2 than SAGE III) in lidar observed ozone values (Figure 12 right panel).

In the lower stratosphere below 20 km, the systematic (median) differences between SAGE III and lidar measurements are within 10% except for Lauder. The systematic biases between SAGE III and lidar can be approximated (to first order) by the relative difference between averaged SAGE III and lidar ozone values (e.g., $(\overline{S} - \overline{L})/\overline{L}$, *where* $\overline{S}$ *and* $\overline{L}$ indicate averaged ozone values from all collocated SAGE III and lidar profiles, respectively). This method can also minimize the sensitivity of outliers. It yields similar results as those from the median of relative differences, except in the lower stratosphere at Lauder (Figure 12 middle panel). This is probably related to the fact that samples of coincident SAGE III and lidar ozone profiles at Lauder are too small (i.e., 13 profiles).

Similar analyses were performed between SAGE III and ozonesondes and the results are shown in Figure 13. In the mid-latitudes, SAGE III ozone values are generally biased high against ozonesondes with differences of ~5% for altitudes above 15 km. The biases increase toward the lower stratosphere and upper troposphere, and reach ~10% at 12–13km. The standard deviations (approximated by the spreads) of mean relative differences are ~5% near the ozone peak and become larger at higher and lower altitudes. The standard deviations increase to ~30–40% at 15 km and ~50% near the tropopause. The comparisons between SAGE III and ozonesondes in the tropics show similar vertical structure as those in the mid-latitudes. SAGE III ozone values are systematically higher than sonde ozone values by ~5% for altitudes above 20 km. The biases increase rapidly toward the UT/LS, and reaches ~10% at 17–18 km and ~40% (or higher) at 15–16 km. It should be noted that comparison results for altitudes below 17 km in the tropics are not robust because both the standard deviations and spreads of relative differences are larger than those of SAGE III and ozonesondes measurements and combined uncertainties (Figure 13). Similar situations also occurs for altitudes below 12 km in the mid-latitudes.

5.4 Estimated accuracies and precisions of SAGE III/ISS AO3 ozone

The comparisons between SAGE III/ISS solar ozone data and correlative satellite and ground-based measurements are summarized in Figure 14. Since there is known thermal sensitivity issue in the OMPS LP ozone data (Kramarova et al., 2018), the OMPS LP data between 28 and 32 km (e.g., Figure 11) were filtered before calculating the means and standard deviations of relative differences between SAGE III and other satellites. There is no additional filtering for Aura MLS, ACE-FTS, OSIRIS, lidar and ozonesonde data. The median and spread are used for comparisons between SAGE III and ground based measurements for reasons discussed earlier. Based on these correlative measurements, the accuracy of SAGE III/ISS AO3 ozone in the stratosphere is better than 5% for altitudes down to 15 km in the mid-latitudes and 20 km in the tropics. The accuracy degrades toward lower altitudes and reaches ~10% at the tropopause. In the southern hemisphere mid-latitudes the SAGE III/ISS ozone show lager than 10% positive bias near 15 km comparing to correlative satellite data (Figure 14). This is due to larger biases between SAGE III/ISS and OMPS LP and OSIRIS in that region (e.g., Figure 11). The SAGE III/ISS, however, shows much better agreement (<10%) with Aura MLS and ozonesondes in the same region. The larger biases (>5%) between SAGE III/ISS and other satellites for altitudes above ~50 km is due to the diurnal cycle effects not being removed from the comparisons which has been discussed earlier in section 5.2.

The standard deviation of relative differences between SAGE III/ISS and correlative measurements can be used as an approximation of measurement uncertainty in the SAGE III instrument. It, however, becomes invalid when the uncertainties (random error) of correlative measurements become larger and/or the uncertainties due to temporal/spatial differences are large. The variance of the differences between SAGE III and collocated measurements contains uncertainties from not only SAGE but also correlative measurements and from uncertainties associated with natural variability (e.g., Sofieva et al., 2014).

$$\sigma^2(x_s - x_c) = \sigma^2(x_s) + \sigma^2(x_c) + \sigma^2(nat)$$

where x_s and x_c are SAGE III and correlative measurements, respectively. The $\sigma^2(nat)$ is the variance contributed by the natural variability, which can be minimized by using coincident criteria. The uncertainties of satellite measurements generally become larger toward the UT/LS regions. This can be seen in Figure 14, where the standard deviations of relative differences between SAGE III and correlative satellite measurements increase from ~5% at 20 km to ~50–60% near 10 km. Although the ground-based measurements (e.g., ozonesondes) have better precisions in the UT/LS region, the mismatch errors between SAGE III and ground-based measurements are larger (e.g., due to larger coincident criteria). Furthermore, the satellite measurements cover a larger air mass while ground-based observations represent a much smaller area. The different horizontal resolution (e.g., smoothing error) could further enhance the mismatch error. Due to the above-mentioned reasons, the standard deviations between SAGE III and ground-based measurements are similar or even larger than those in SAGE III and satellite comparisons (Figure 14).

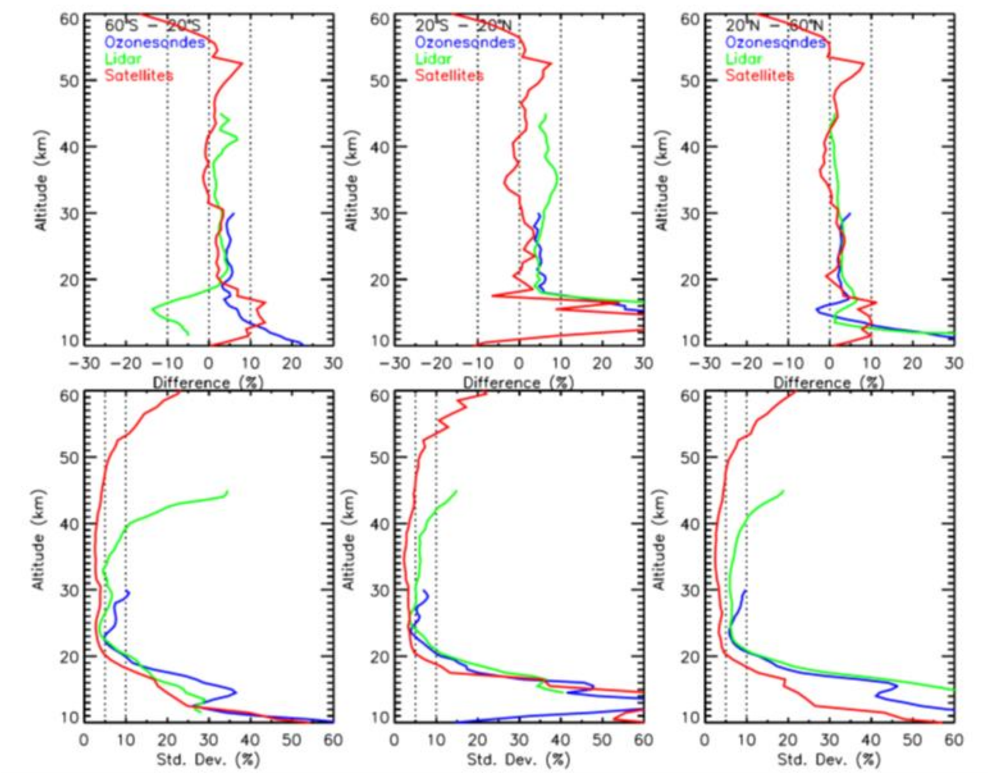


Figure 14: Mean (or median) differences (top panel), and the standard deviations (or spreads) of differences (bottom panel) between SAGE III-ISS ozone and coincident measurements from satellites (red line), lidar (green line) and ozonesondes (blue line) at three latitude bands, 60°S–20°S (left column), 20°S–20°N (middle column) and 20°N–60°N (right column). Results between SAGE III and ground based measurements (lidar and ozonesondes) are based on medians and spreads, while comparisons between SAGE III and other satellites are based on mean and standard deviation.

To better assess the precisions of SAGE III ozone measurements especially in the UT/LS, we used the method in Fioletov et al. (2006). In Fioletov et al. (2006), it is assumed that paired measurements are perfectly collocated (i.e., no mismatch error). In reality it is almost impossible to have SAGE III and correlative measurements at the same location and time. The effect of spatial and temporal differences, however, could be minimized by using tighter coincident criteria. We used smaller coincident criteria of latitude differences within ±1°, longitude differences within ±5°, and the closest in time within the same day for this purpose. The estimated precisions of SAGE III AO3 ozone based on comparisons with correlative Aura MLS and OMPS LP data are shown in Figure 15. We did not use other correlative satellite or ground-based measurements because there were fewer coincident profiles with SAGE III compared to those with Aura MLS and OMPS LP.

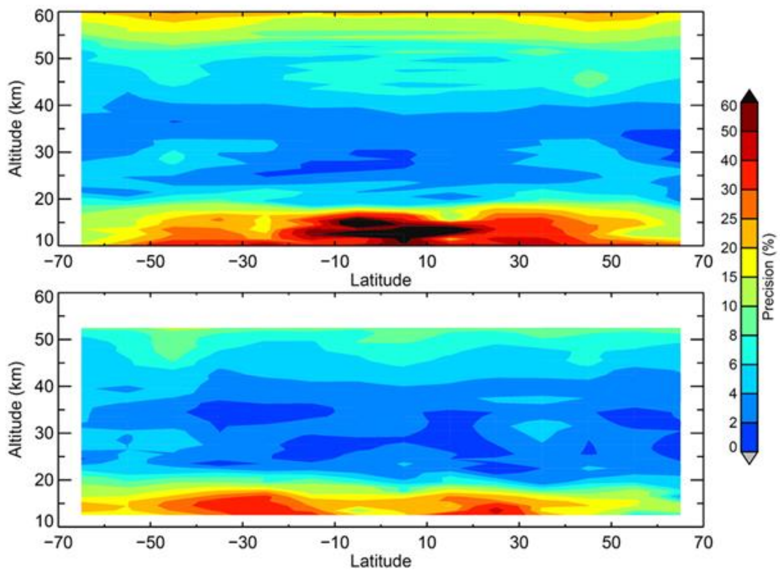


Figure 15: Estimated precisions in SAGE III/ISS AO3 ozone based on compassions with collocated Aura MLS (top) and OMPS LP (bottom) data between June 2017 and May 2019.

By comparing SAGE III/ISS against collocated Aura MLS measurements the estimated precision of SAGE III ozone is approximately 3% (e.g., 2–4%) between 20 and 40 km, and ~10–15% at 55 km (Figure 15). Below 20 km, the precisions of SAGE III ozone degrade toward lower altitudes and reach ~20–30% near the tropopause. Similar results can be seen in the comparisons between SAGE III and OMPS LP except in the tropical UT/LS region. Since both analyses, between SAGE III and Aura MLS and OMPS LP, show consistent results, this indicates that the derived precisions of SAGE III ozone data are robust. The estimated precisions of SAGE III ozone shown in Figure 15 are in general slightly larger than the random errors reported by the SAGE retrieval algorithm (e.g., Figure 4). This is probably due to the small residual effect of spatial and temporal differences between SAGE III and correlative measurements (mismatch error cannot be completely removed from the analyses by the coincident criteria).

6 Conclusions

The Stratospheric Aerosol and Gas Experiment III on the International Space Station (SAGE III/ISS) was launched in February 2017 and started routine operation in June 2017. It is the second SAGE III instrument but with better latitudinal coverage. Similar to SAGE II, it provides near global observations on a monthly basis. The first two years of SAGE III/ISS version 5.1 solar ozone data were evaluated by using correlative measurements from satellites (Aura MLS, ACE-FTS, OSIRIS, OMPS LP) and ground-based instruments (lidar and ozonesondes). There are three retrieved ozone products, denoted as AO3, MLR, and MES, from SAGE III solar occultation measurements. The first two (AO3 and MLR) algorithms both use ozone absorption in the Chappuis band but different methods to separate ozone and other interfering gases from the observed slant path radiances (SAGE III ATBD, 2002). The third algorithm (MES) uses ozone absorption in the ultraviolet band, which can provide better ozone signals at higher altitudes (e.g., above 45 km). The MES retrieval algorithm, however, is affected by a spectral stray light problem, which has not been properly corrected. The MES ozone product, therefore, is currently not recommended for scientific studies.

To evaluate the quality of SAGE III/ISS solar ozone data, appropriate procedures have been applied to SAGE III and correlative measurements to minimize the biases and uncertainties associated with mismatch (spatial/temporal differences) and different smoothing (e.g., resolutions) in respective observations. The coincident criteria are a trade-off between mismatch uncertainties and large sample size (number of coincident profiles), especially for comparisons between SAGE III and ground-based measurements. There is no good way to remove the horizontal component of smoothing differences, which, however, would be reflected as random errors in statistics with a sufficiently large sample size (e.g., Cortesi et al., 2007). The method recommended by the instrument science team or Gaussian kernel (e.g., Kar et al., 2007; Sheese et al., 2017) was applied to the profiles with finer vertical resolution to remove/minimize the vertical component of smoothing differences. Since there are altitude registration errors of approximately 100 m in the auxiliary temperature and pressure profiles in SAGE III/ISS version 5.1 data, we used ozone number density on geometric altitude as the common coordinate for comparisons. The altitude registration errors in SAGE III temperature and pressure profiles are due to a simplistic approximation in the geopotential height to geometric altitude conversion. It should be noted that this error would not affect SAGE III ozone on its native retrieved grids, number density and geometric altitude, unless the profiles are converted to mixing ratio on pressure coordinate by using the auxiliary temperature and pressure profile accompanying each ozone profile.

For ozone retrieved from the AO3 and MLR algorithm, it was found that MLR ozone has larger biases (e.g., by 10% or higher) and uncertainties (by a factor of 2 to 3) in the UT/LS and above the upper stratosphere by comparisons with correlative measurements or using residual analyses (Damadeo et al., 2014). These results are similar to a previous study (Wang et al., 2006) for the SAGE III/M3M instrument. SAGE III/ISS AO3 ozone show very good agreement with correlative measurements, with mean biases less than 5% for altitudes down to ~15 km in the mid-latitudes and ~20 km in the tropics. The differences become larger in the lower mesosphere (e.g., 10–15% near 60 km), which mainly results from the ozone diurnal cycle not being removed from the comparisons. In the lower stratosphere and upper troposphere, the SAGE III/ISS AO3 ozone show systematic high biases that increase with decreasing altitudes, and reach ~10% near the tropopause. The precision of SAGE III/ISS AO3 ozone is estimated to be ~3% between 20 and 40 km. The precisions degrades toward higher and lower altitudes due to smaller signal to noise ratio in Chappuis band and large natural variability in the UT/LS region. The estimated precision in AO3 ozone is ~10–15% in the lower mesosphere (55 km), and ~20–30% near the tropopause.

The sunrise/sunset differences in SAGE III/ISS retrieved ozone were examined by regression analyses and comparisons with correlative Aura MLS data. It was found that SAGE III sunset ozone values are systematically larger than sunrise values by ~5–8%, at 40–55 km with mean differences larger in the tropics than at mid-latitudes. In the lower stratosphere below ~25 km, the sunrise values become slightly larger than sunset values by a few percent. The vertical and latitudinal distribution of sunrise/sunset differences in observed ozone is consistent with the vertical transport of atmospheric tidal winds (Sakazaki et al., 2013). The magnitude of sunrise/sunset differences in SAGE III/ISS retrieved ozone in the upper stratosphere, however, are almost twice as large as those observed from other satellites and model prediction (Sakazaki et al., 2015). The reason for this is not clear and needs further investigation. The SAGE III retrieval algorithm team is investigating the high biases in retrieved ozone in the UT/LS region. Preliminary studies indicate that the oxygen dimer O_2-O_2 (or O_4) spectroscopy used in the current v5.1 retrieval algorithm could primarily contribute to the observed high biases in ozone. It was also found that an under estimation of aerosol contribution in the ozone absorption band could indicate a potentially small high bias in stratospheric ozone in both the AO3 and MLR algorithms. The effects are more pronounced in the MLR than the AO3 algorithm. This is consistent with our validation results, which show altitude-dependent high biases in both MLR and AO3 retrieved ozone for altitudes below 15–20 km. The biases in MLR ozone are also larger than those in AO3. Further analyses will be made in the future by applying updated O_4 spectroscopy and aerosol clearing procedures in the retrieval algorithm to quantify these effects on retrieved ozone in the upper troposphere and lower stratosphere.

Appendix

As a known anomaly in v5.1, Section 2.2 describes an altitude registration bias in the reported pressure and temperature profiles that are passed through the algorithm. This Appendix details a recommended conversion from which Figure 2 derives. The process involves three simple steps: 1) convert the geometric altitude array upon which the pressures and temperatures are reported (Z_{OLD}) back to the original geopotential heights (Z_{Φ}) using the approximation used in the v5.1 algorithm, 2) convert the geopotential heights to geometric altitude (Z_{NEW}) using a better model, and 3) remap the reported pressures and temperatures on the new geometric altitudes to the desired grid (such as the original grid) using your favorite interpolation scheme. Step 1 is very straightforward, and comes from the overly simplistic assumption that the surface gravity is the same everywhere and is equal to the mean surface gravity (g_0) defined as 9.80665 m/s^2):

$$Z_{\Phi} = \frac{Z_{OLD} * R_{EARTH}}{R_{EARTH} + Z_{OLD}}$$

Step 2 is also straightforward:

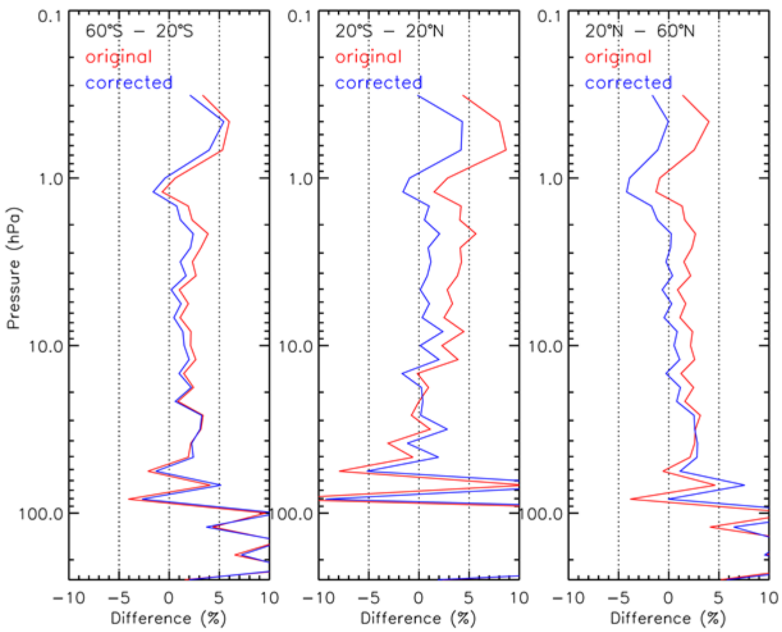
$$Z_{NEW} = \frac{Z_{\Phi} * R_{EARTH}}{\frac{g(\theta)}{g_0} * R_{EARTH} - Z_{\Phi}}$$

where g is the surface gravity at a particular geodetic latitude (θ, or "map" latitude). While the model of surface gravity is always being updated, the SAGE algorithm makes use of the World Geodetic System 1984 model (WGS84, updated in 2004) (NIMA Technical Report, 1997) and thus this provides the recommendation for g:

$$g(\theta) = 9.7803253359 \frac{1 + 0.00193185265241 * \mathrm{SIN}^2(\theta)}{\sqrt{1 - 0.00669437999013 * \mathrm{SIN}^2(\theta)}}$$

924 It is important to note that the latitude-dependence of R_{EARTH} should be taken into account for all
925 of these calculations.

926



927

928 **Figure S1:** Mean differences between SAGE III/ISS AO3 ozone and collocated Aura MLS data
929 at three latitude bands 60°S–20°S (left column), 20°S–20°N (middle column), and 20°N–60°N
930 (right column). SAGE ozone profiles are converted to MLS coordinates by using reported (red)
931 and bias corrected (blue) temperature and pressure profiles. The percentage difference is
932 calculated as (SAGE-MLS)/MLS*100%.

933 To evaluate the effect of altitude registration bias in the reported temperature and
934 pressure profiles on ozone, SAGE III/ISS AO3 ozone data were compared against collocated
935 Aura MLS nighttime measurements on volume mixing ratio and pressure coordinates (VMR/P).
936 The coincidence criteria are the same as those described in section 4. SAGE III/ISS AO3 ozone
937 profiles were converted to VMR/P by using accompanying temperature and pressure profiles.
938 The mean biases between SAGE and MLS are generally within 5% between ~83 and 0.3 hPa
939 except in the tropics, where larger biases (>5%) are found below ~46 and above 1 hPa (Figure
940 S1). It should be noted that the differences between SAGE and MLS in the tropics show an
941 altitude-dependent structure. SAGE ozone shows increasing positive biases for altitudes above
942 the ozone peak while increasing negative biases below the ozone peak. This is due to the altitude
943 registration errors in reported temperature and pressure profiles that are more pronounced in the
944 tropics than mid-latitudes (Figure 2). After correcting the altitude registration errors in the
945 reported temperature and pressure profiles the SAGE ozones show better agreement with MLS
946 data without the altitude-dependent feature. The mean differences in general are less than 3% for

altitudes between 1 and ~83 hPa in the mid-latitudes and between 1 and ~56 hPa in the tropics (Figure S1).

Acknowledgments

The authors would like to thank Bryan Johnson at Earth System Research Laboratory, Global Monitoring Division, NOAA for some ozonesonde data in NOAA's network. The lidar data used in this publication were obtained from the Network for the Detection of Atmospheric Composition Change (NDACC) and are publicly available (see http://www.ndacc.org). We also want to thank WMO/GAW Ozone Monitoring Community, World Meteorological Organization-Global Atmosphere Watch Program (WMO-GAW)/World Ozone and Ultraviolet Radiation Data Centre (WOUDC) for ozonesonde data. Retrieved May 2019 from https://woudc.org. A list of all contributors is available on the website. doi:10.14287/10000008. This work was funded by the National Aeronautics and Space Administration (NASA) on Grant Number 80NSSC18K0710. The Atmospheric Chemistry Experiment (ACE) is a Canadian-led mission mainly supported by the Canadian Space Agency.

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/amt6-1447-2013, 2013.

Adams, C., Bourassa, A. E., Sofieva, V., Froidevaux, L., McLinden, C. A., Hubert, D., Lambert, J.-C., Sioris, C. E., and Degenstein, D. A.: Assessment of Odin-OSIRIS ozone measurements from 2001 to the present using MLS, GOMOS, and ozonesondes, *Atmos. Meas. Tech.*, 7, 49-64, https://doi.org/10.5194/amt-7-49-2014, 2014.

Bernath, P. F., McElroy, C. T., Abrams, M. C., Boone, C. D., Butler, M., Camy-Peyret, C., Carleer, M., Clerbaux, C., Coheur, P.-F., Colin, R., DeCola, P., DeMazière, M., Drummond, J. R., Dufour, D., Evans, W. F. J., Fast, H., Fussen, D., Gilbert, K., Jennings, D. E., Llewellyn, E. J., Lowe, R. P., Mahieu, E., McConnell, J. C., McHugh, M., McLeod, S. D., Michaud, R., Midwinter, C., Nassar, R., Nichitiu, F., Nowlan, C., Rinsland, C. P., Rochon, Y. J., Rowlands, N., Semeniuk, K., Simon, P., Skelton, R., Sloan, J. J., Soucy, M.-A., Strong, K., Tremblay, P., Turnbull, D., Walker, K. A., Walkty, I., Wardle, D. A., Wehrle, V., Zander, R., and Zou, J.: Atmospheric Chemistry Experiment (ACE): Mission overview, *Geophys. Res. Lett.*, 32, L15S01, doi:10.1029/2005GL022386, 2005.

Boone, C. D., Nassar, R., Walker, K. A., Rochon, Y., McLeod, S. D., Rinsland, C. P., and Bernath, P. F.: Retrievals for the Atmospheric Chemistry Experiment Fourier-transform spectrometer, *Appl. Optics*, 44(33), 7218–7231, 2005.

Boone, C. D., Walker, K. A., and Bernath, P. F.: Version 3 Retrievals for the Atmospheric Chemistry Experiment Fourier Transform Spectrometer (ACE-FTS), The Atmospheric Chemistry Experiment ACE at 10: A Solar Occultation Anthology, A. Deepak Publishing, Hampton, Virginia, USA, 103–127, 2013.

Bourassa, A. E., Degenstein, D. A., and Llewellyn, E. J.: SASKTRAN: 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, 2008.

Bourassa, A. E., Roth, C. Z., Zawada, D. J., Rieger, L. A., McLinden, C. A., and Degenstein, D. A.: Drift-corrected Odin-OSIRIS ozone product: algorithm and updated stratospheric ozone trends, *Atmos. Meas. Tech.*, 11, 489-498, https://doi.org/10.5194/amt-11-489-2018, 2018.

Brühl, C., Drayson, S. R., Russell III, J. M., Crutzen, P. J., McInerney, J. M., Purcell, P. N., Claude, H., Gernandt, H., McGee, T. J., McDermid, I. S., and Gunson, M. R.: Halogen Occultation Experiment ozone channel validation, *J. Geophys. Res*., 101, 10,217–10,240, 1996.

Chu, W. P., McCormick, M. P., Lenoble, J., Broniez, C., and Pruvost, P.: SAGE II inversion algorithm, *J. Geophys. Res.*, 94, 8339–8351, doi:10.1029/JD094iD06p08447, 1989

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

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

Damadeo, R. P., Zawodny, J. M., and Thomason, L. W.: Reevaluation of stratospheric ozone trends from SAGE II data using a simultaneous temporal and spatial analysis, *Atmos. Chem. Phys.*, 14, 13455-13470, https://doi.org/10.5194/acp-14-13455-2014, 2014.

Damadeo, R, Zawodny, J. M., Remsberg, E. E., and Walker, K. A.: The impact of nonuniform sampling on stratospheric ozone trends derived from occultation instruments, *Atmos. Chem. Phys.*, 18, 535–554, https://doi.org/10.5194/acp-18-535-2018, 2018.

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.

Dupuy, E., Walker, K. A., Kar, J., Boone, C. D., McElroy, C. T., Bernath, P. F., Drummond, J. R., Skelton, R., McLeod, S. D., Hughes, R. C., Nowlan, C. R., Dufour, D. G., Zou, J., Nichitiu, F., Strong, K., Baron, P., Bevilacqua, R. M., Blumenstock, T., Bodeker, G. E., Borsdorff, T., Bourassa, A. E., Bovensmann, H., Boyd, I. S., Bracher, A., Brogniez, C., Burrows, J. P., Catoire, V., Ceccherini, S., Chabrillat, S., Christensen, T., Coffey, M. T., Cortesi, U., Davies, J., De Clercq, C., Degenstein, D. A., De Mazière, M., Demoulin, P., Dodion, J., Firanski, B., Fischer, H., Forbes, G., Froidevaux, L., Fussen, D., Gerard, P., GodinBeekmann, S., Goutail, F., Granville, J., Griffith, D., Haley, C. S., Hannigan, J. W., Höpfner, M., Jin, J. J., Jones, A., Jones, N. B., Jucks, K., Kagawa, A., Kasai, Y., Kerzenmacher, T. E., Kleinböhl, A., Klekociuk, A. R., Kramer, I., Küllmann, H., Kuttippurath, J., Kyrölä, E., Lambert, J.-C., Livesey, N. J., Llewellyn, E. J., Lloyd, N. D., Mahieu, E., Manney, G. L., Marshall, B. T., McConnell, J. C., McCormick, M. P., McDermid, I. S., McHugh, M., McLinden, C. A., Mellqvist, J., Mizutani, K., Murayama, Y., Murtagh, D. P., Oelhaf, H., Parrish, A., Petelina, S. V., Piccolo, C., Pommereau, J.-P., Randall, C. E., Robert, C., Roth, C., Schneider, M., Senten, C., Steck, T., Strandberg, A., Strawbridge, K. B., Sussmann, R., Swart, D. P. J., Tarasick, D. W., Taylor, J. R., Tétard, C., Thomason, L. W., Thompson, A. M., Tully, M. B., Urban, J., Vanhellemont, F., Vigouroux, C., von Clarmann, T., von der Gathen, P., von Savigny, C., Waters, J. W., Witte, J. C., Wolff, M., and Zawodny, J. M.: Validation of ozone measurements from the Atmospheric Chemistry Experiment (ACE), *Atmos. Chem. Phys*., 9, 287–343, doi:10.5194/acp-9-287-2009, 2009.

Fioletov, V. E., Tarasick, D. W., and Petropavlovskikh, I.: Estimating ozone variability and instrument uncertainties from SBUV(/2), ozonesonde, Umkehr, and SAGE II measurements: Short-term variations, *J. Geophys. Res.*, 111, D02305, doi:10.1029/2005JD006340, 2006.

Flittner, D., Bhartia, P. K., and Herman, B. M.: O3 profiles retrieved from limb scatter measurements: Theory, *Geophys. Res. Lett.*, 27, 2601–2604, 2000.

Flynn, L. E., Seftor, C. J., Larsen, J. C., and Xu, P.: The Ozone Mapping and Profiler Suite, in: Earth Science Satellite Remote Sensing, edited by: Qu, J. J., Gao, W., Kafatos, M., Murphy, R. E., and Salomonson, V. V., Springer, Berlin, 279–296, https://doi.org/10.1007/978-3-540-37293-6_15, 2006.

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.

Gelaro, R., and Coauthors, 2017: The Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2). *J. Climate*, 30, 5419–5454, https://doi.org/ 10.1175/JCLI-D-16-0758.1.

Global Modeling and Assimilation Office (GMAO) (2015), tavg3_3d_asm_Nv: MERRA-2 3D IAU State, Meteorology time-averaged 3-hourly (model-level, 0.625x0.5L72), version 5.12.4, Greenbelt, MD, USA: Goddard Space Flight Center Distributed Active Archive Center (GSFC DAAC), Accessed May 2019 at doi: 10.5067/SUOQESM06LPK.

Godin, S., Carswell, A. I., Donovan, D. P., Claude, H., Steinbrecht, W., McDermid, I. S., McGee, T. J., Gross, M. R., Nakane, H., Swart, D. P. J., Bergwerff, H. B., Uchino, O., von der Gathen, P., and Neuber, R.: Ozone differential absorption lidar algorithm intercomparison, *Appl. Opt.*,38, 6225–6236, doi:10.1364/AO.38. 006225, 1999.

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

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 Zablocki, G.: Validation of Aura Microwave Limb Sounder Ozone by ozonesonde and lidar measurements, *J. Geophys. Res.*, 112, D24S34, doi:10.1029/2007JD008776, 2007.

Johnson, Bryan J.; Cullis, Patrick D.; and NOAA ESRL (2018): Earth System Research Laboratory Ozone Water Vapor Group Ozonesonde Measurements, Version 1. NOAA National Centers for Environmental Information. doi:10.7289/V5CC0XZ1 [assess in May 2019]

Kar, J., McElroy, C. T., Drummond, J. R., Zou, J., Nichitiu, F., Walker, K. A., Randall, C. E., Nowlan, C. R., Dufour, D. G., Boone, C. D., Bernath, P. F., Trepte, C. R., Thomason, L. W., and McLinden, C.: Initial comparison of Ozone and NO2 profiles from ACE-MAESTRO with Balloon and Satellite Data, *J. Geophys. Res.*, 112, D16301, doi:10.1029/2006JD008242, 2007

Keckhut, P., McDermid, S., Swart, D., McGee, T., Godin Beekmann, S., Adriani, A., Barnes, J., Baray, J.-L., Bencherif, H., Claude, H., di Sarra, A., Fiocco, G., Hansen, G., Hauchecorne, A., Leblanc, T., Lee, C., Pal, S., Megie, G., Nakane, H., Neuber, R., Steinbrecht, W., and Thayer, J.: Review of ozone and temperature lidar validations performed within the framework of the Network for the Detection of Stratospheric Change, *J. Environ. Monit.*, 6, 721–733, doi:10.1039/b404256e, 2004.

Kramarova, N. A., Nash, E. R., Newman, P. A., Bhartia, P. K., McPeters, R. D., Rault, D. F., Seftor, C. J., Xu, P. Q., and Labow, G. J.: Measuring the Antarctic ozone hole with the new Ozone Mapping and Profiler Suite (OMPS), *Atmos. Chem. Phys.*, 14, 2353–2361, https://doi.org/10.5194/acp-14-2353-2014, 2014.

Kramarova, N. A., Bhartia, P. K., Jaross, G., Moy, L., Xu, P., Chen, Z., DeLand, M., Froidevaux, L., Livesey, N., Degenstein, D., Bourassa, A., Walker, K. A., and Sheese, P.: Validation of ozone profile retrievals derived from the OMPS LP version 2.5 algorithm against correlative satellite measurements, *Atmos. Meas. Tech.*, 11, 2837-2861, https://doi.org/10.5194/amt-11-2837-2018, 2018.

Kyrölä, E., Laine, M., Sofieva, V., Tamminen, J., Päivärinta, S.-M., Tukiainen, S., Zawodny, J., and 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.*, 13, 10645–10658, doi:10.5194/acp-13-10645-2013, 2013.

Livesey, N. J., W. V. Snyder, W. G. Read, and P. A. Wagner (2006), Retrieval algorithms for the EOS Microwave Limb Sounder (MLS), *IEEE Trans. Geosci. Remote Sens.*, 44(5), 1144–1155.

Livesey N. J., M. J. Filipiak, L. Froidevaux, W. G. Read, A. Lambert, M. L. Santee, J. H. Jiang, H. C. Pumphrey, J. W. Waters, R. E. Cofield, D. T. Cuddy, W. H. Daffer, B. J. Drouin, R. A. Fuller, R. F. Jarnot, Y. B. Jiang, B. W. Knosp, Q. B. Li, V. S. Perun, M. J. Schwartz, W. V. Snyder, P. C. Stek, R. P. Thurstans, P. A. Wagner, M. Avery, E. V. Browell, J.-P. Cammas, L. E. Christensen, G. S. Diskin, R.-S. Gao, H.-J. Jost, M. Loewenstein, J. D. Lopez, P. Nedelec, G. B. Osterman, G. W. Sachse, C. R. Webster, Validation of Aura Microwave Limb Sounder O3 and CO observations in the upper troposphere and lower stratosphere, *J. Geophys. Res.*, 113,D15S02,doi:10.1029/2007JD008805, 2008.

Livesey, Nathaniel J, William G. Read, Paul A. Wagner, Lucien Froidevaux, Alyn Lambert, Gloria L. Manney, Luis F. Millán Valle, Hugh C. Pumphrey, Michelle L. Santee, Michael J. Schwartz, Shuhui Wang, Ryan A. Fuller, Robert F. Jarnot, Brian W. Knosp, Elmain Martinez, Richard R. Lay, Earth Observing System (EOS) Aura Microwave Limb Sounder (MLS) Version 4.2x Level 2 data quality and description document, JPL D-33509 Rev.D 2018.

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., Sjoberg, 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.

Mauldin, L. E., Zaun, N. H., McCormick, M. P., Guy, J. H., and Vaughn, W. R.: Stratospheric Aerosol and Gas Experiment II Instrument: A Functional Description, *Opt. Eng.*, 24, 307–312, doi:10.1117/ 12.7973473, 1985.

Mégie, G., G. Ancellet, and J. Pelon, Lidar measurements of ozone vertical profiles, *Appl. Opt.*, 24, 3454-3453, 1985

McCormick, M. P., J. M. Zawodny, W. P. Chu, J. W. Baer, J. Guy, and A. Ray, Stratospheric Aerosol and Gas Experiment III (SAGE III), *SPIE International Symposium for Optical Engineering* - Orlando, FL, 1993

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.

McPeters R. D., Janz, S. J., Hilsenrath, E., and Brown, T. L.: The retrieval of O3 profiles from limb scatter measurements: Results from the Shuttle Ozone Limb Sounding Experiment, *Geoph. Res. Lett.*, 27, 2597–2600, 2000.

Nair, P. J., Godin-Beekmann, S., Froidevaux, L., Flynn, L. E., Zawodny, J. M., Russell III, J. M., Pazmiño, A., Ancellet, G., Steinbrecht, W., Claude, H., Leblanc, T., McDermid, S., van Gijsel, J. A. E., Johnson, B., Thomas, A., Hubert, D., Lambert, J.-C., Nakane, H., and Swart, D. P. J.: Relative drifts and stability of satellite and ground-based stratospheric ozone profiles at NDACC lidar stations, *Atmos. Meas. Tech*., 5, 1301–1318, doi:10.5194/amt-5-1301-2012, 2012.

NIMA Technical Report TR8350.2: Department of Defense World Geodetic System 1984, Its Definition and Relationships With Local Geodetic Systems, 3rd Edition, 4 July 1997.

Parrish,A.,Boyd,I.S.,Nedoluha,G.E.,Bhartia,P.K.,Frith,S.M., Kramarova, N. A., Connor, B. J., Bodeker, G. E., Froidevaux, L., Shiotani, M., and Sakazaki, T.: Diurnal variations of stratospheric ozone measured by ground-based microwave remote sensing at the Mauna Loa NDACC site: measurement validation and GEOSCCM model comparison, *Atmos. Chem. Phys.*, 14, 7255–7272, https://doi.org/10.5194/acp-14-7255-2014, 2014.

Pitts, M. C., and L. W. Thomason (2003), SAGE III temperature and pressure retrievals: Initial results, *Proc. SPIE Int. Soc. Opt. Eng*., 482, 62– 70.

Rault, D. F. (2005), Ozone profile retrieval from Stratospheric Aerosol and Gas Experiment (SAGE III) limb scatter measurements, *J. Geophys. Res*.,110, D09309, doi:10.1029/2004JD004970

Rault, D. F. and Loughman, R. P.: The OMPS Limb Profiler Environmental Data Record Algorithm Theoretical Basis Document and Expected Performance, *IEEE T. Geosci. Remote Sens.*, 51, 2505–2527, https://doi.org/10.1109/TGRS.2012.2213093, 2013.

Rodgers, C. D. (2000), Inverse Methods for Atmospheric Science, Theory and Practice, 238 pp., World Sci., Hackensack, N.J.

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.

SAGE III ATBD: SAGE III Algorithm Theoretical Basis Document: Solar and Lunar Algorithm, Earth Observing System Project Science Office web site online available at: https://eospso.gsfc.nasa.gov/sites/default/files/atbd/atbd-sage-solar-lunar.pdf, 2002

Sakazaki, T., Fujiwara, M., Mitsuda, C., Imai, K., Manago, N., Naito, Y., Nakamura, T., Akiyoshi, H., Kinnison, D., Sano, T., Suzuki, M., and Shiotani, M.: Diurnal ozone variations in the stratosphere revealed in observations from the Superconducting Submillimeter-Wave Limb-Emission Sounder (SMILES) on board the International Space Station (ISS), *J. Geophys. Res*. Atmos., 118, 2991–3006, doi:10.1002/jgrd.50220, 2013.

Sakazaki, T., Shiotani, M., Suzuki, M., Kinnison, D., Zawodny, J. M., McHugh, M., and Walker, K. A.: Sunset–sunrise difference in solar occultation ozone measurements (SAGE II, HALOE, and ACE–FTS) and its relationship to tidal vertical winds, *Atmos. Chem. Phys*., 15, 829-843, https://doi.org/10.5194/acp-15-829-2015, 2015

Sheese,P.E., Boone, C.D.,and Walker,K.A.: Detecting physically unrealistic outliers in ACE-FTS atmospheric measurements, *Atmos. Meas. Tech*., 8, 741–750, https://doi.org/10.5194/amt-8741-2015, 2015.

Smit, H. G. J. and the Panel for the Assessment of Standard Operating Procedures for Ozonesondes (ASOPOS): Quality Assurance and Quality Control for Ozonesonde Measurements in GAW, WMO Global Atmosphere Watch report No. 201, World Meteorological Organization, available at: https://library.wmo.int/pmb_ged/gaw_201_en.pdf (last access: 22 July 2019), 2014.

Sofieva, V. F., Tamminen, J., Kyrölä, E., Laeng, A., von Clarmann, T., Dalaudier, F., Hauchecorne, A., Bertaux, J.-L., Barrot, G., Blanot, L., Fussen, D., and Vanhellemont, F.: Validation of GOMOS ozone precision estimates in the stratosphere, *Atmos. Meas. Tech*., 7, 2147–2158, https://doi.org/10.5194/amt-7-2147-2014, 2014.

Sheese, P. E., Walker, K. A., Boone, C. D., Bernath, P. F., Froidevaux, L., Funke, B., and von Clarmann, T.: ACE-FTS ozone, water vapour, nitrous oxide, nitric acid, and carbon monoxide profile comparisons with MIPAS and MLS, *J. Quant. Spectrosc. Ra*., 186, 63–80, https://doi.org/10.1016/j.jqsrt.2016.06.026, 2017.

SPARC/IO3C/GAW, 2019: SPARC/IO3C/GAW Report on Long-term Ozone Trends and Uncertainties in the Stratosphere. I. Petropavlovskikh, S. Godin-Beekmann, D. Hubert, R. Damadeo, B. Hassler, V. Sofieva (Eds.), SPARC Report No. 9, GAW Report No. 241, WCRP-17/2018, doi: 10.17874/f899e57a20b

SPARC, 2006: SPARC Assessment of Stratospheric Aerosol Properties (ASAP). L. Thomason and Th. Peter (Eds.), SPARC Report No. 4, WCRP-124, WMO/TD – No. 1295

Thomason, L. W., Moore, J. R., Pitts, M. C., Zawodny, J. M., and Chiou, E. W.: An evaluation of the SAGE III version 4 aerosol extinction coefficient and water vapor data products, *Atmos. Chem. Phys.*, 10, 2159-2173, https://doi.org/10.5194/acp-10-2159-2010, 2010

Thompson, A. M., J. C. Witte, C., Sterling, A., Jordan, B. J., Johnson, S. J. Oltmans, ... Thiongo, K. (2017). First reprocessing of Southern Hemisphere Additional Ozonesondes (SHADOZ) ozone profiles (1998-2016): 2. Comparisons with satellites and ground-based instruments. *Journal of Geophysical Research: Atmospheres*, 122, 13,000-13,025. https://doi.org/10.1002/2017JD 027406.

von Clarmann, T.: Validation of remotely sensed profiles of atmospheric state variables: strategies and terminology, *Atmos. Chem. Phys*., 6, 4311-4320, https://doi.org/10.5194/acp-6-4311-2006, 2006.

Wang, H. J., D. M. Cunnold, and X. Bao, A critical analysis of SAGE ozone trends, J. *Geophys. Res*.,101, 12,495–12,514, 1996.

Wang, H. J., D. M. Cunnold, L. W. Thomason, J. M. Zawodny, and G. E. Bodeker, Assessment of SAGE version 6.1 ozone data quality, *J. Geophys. Res*.,107(D23), 4691, doi:10.1029/2002JD002418, 2002.

Wang, H.-J., Cunnold, D. M., Trepte, C., Thomason, L. W., Zawodny J. M.: SAGE III solar ozone measurements: Initial results, *Geophys. Res. Lett*., 33, L03805, doi:10.1029/2005GL025099, 2006.

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.

Witte, J.C., A. M. Thompson, H. G. J. Smit, M. Fujiwara, F. Posny, Gert J. R. Coetzee, E. T. Northam, B. J. Johnson, C, W, Sterling, M. Mohamad, S. Ogino, A. Jordan, and F. R. da Silva (2017), First reprocessing of Southern Hemisphere ADditional OZonesondes (SHADOZ) profile records (1998-2015): 1. Methodology and evaluation, *J. Geophys. Res. Atmos*., 122, 6611-6636. https://doi.org/10.1002/2016JD026403.

WMO (World Meteorological Organization), Scientific Assessment of Ozone Depletion: 2018, Global Ozone Research and Monitoring Project–Report No. 58, 588 pp., Geneva, Switzerland, 2018.

Zawada DJ, Dueck SR, Rieger LA, Bourassa AE, Lloyd ND, Degenstein DA, High-resolution and Monte Carlo additions to the SASKTRAN radiative transfer model, *Atmos. Meas. Tech*., 8, 2609-2623, 2015