Atmos. Meas. Tech. Discuss., 8, 6661–6757, 2015
www.atmos-meas-tech-discuss.net/8/6661/2015/
doi:10.5194/amtd-8-6661-2015

Atmospheric Measurement Techniques Discussions

Ground-based assessment of the bias and long-term stability of fourteen limb and occultation ozone profile data records

D. Hubert[1], J.-C. Lambert[1], T. Verhoelst[1], J. Granville[1], A. Keppens[1], J.-L. Baray[2], U. Cortesi[3], D. A. Degenstein[4], L. Froidevaux[5], S. Godin-Beekmann[6], K. W. Hoppel[7], E. Kyrölä[8], T. Leblanc[9], G. Lichtenberg[10], C. T. McElroy[11], D. Murtagh[12], H. Nakane[13,14], J. M. Russell III[15], J. Salvador[16], H. G. J. Smit[17], K. Stebel[18], W. Steinbrecht[19], K. B. Strawbridge[20], R. Stübi[21], D. P. J. Swart[22], G. Taha[23,24], A. M. Thompson[24], J. Urban[12,†], J. A. E. van Gijsel[25], P. von der Gathen[26], K. A. Walker[27,28], E. Wolfram[16], and J. M. Zawodny[29]

[1]Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium
[2]Laboratoire de Météorologie Physique, Observatoire de Physique du Globe de Clermont-Ferrand, Clermont-Ferrand, France
[3]Istituto di Fisica Applicata "Nello Carrara" del Consiglio Nazionale delle Ricerche, Sesto Fiorentino, Italy
[4]Institute of Space and Atmospheric Studies, University of Saskatchewan, Saskatoon, SK, Canada
[5]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA

[6]Laboratoire Atmosphère Milieux Observations Spatiales, Université de Versailles Saint-Quentin en Yvelines, Centre National de la Recherche Scientifique, Paris, France
[7]Naval Research Lab, Washington, DC, USA
[8]Finnish Meteorological Institute, Helsinki, Finland
[9]Jet Propulsion Laboratory, California Institute of Technology, Wrightwood, CA, USA
[10]German Aerospace Center (DLR), Remote Sensing Technology Institute, Oberpfaffenhofen, Germany
[11]York University, Toronto, ON, Canada
[12]Department of Earth and Space Sciences, Chalmers University of Technology, Göteborg, Sweden
[13]Kochi University of Technology, Kochi, Japan
[14]National Institute for Environmental Studies, Tsukuba, Ibaraki, Japan
[15]Department of Atmospheric and Planetary Science, Hampton University, VA, USA
[16]CEILAP-UNIDEF (MINDEF-CONICET), UMI-IFAECI-CNRS-3351, Villa Martelli, Argentina
[17]Research Centre Jülich, Institute for Energy and Climate Research: Troposphere (IEK-8), Jülich, Germany
[18]Norwegian Air Research Institute (NILU), Kjeller, Norway
[19]Meteorologisches Observatorium, Deutscher Wetterdienst, Hohenpeissenberg, Germany
[20]Air Quality Processes Research Section, Environment Canada, Toronto, ON, Canada
[21]Payerne Aerological Station, MeteoSwiss, Payerne, Switzerland
[22]National Institute for Public Health and the Environment (RIVM), Bilthoven, the Netherlands
[23]Universities Space Research Association, Greenbelt, MD, USA
[24]NASA Goddard Space Flight Center, Greenbelt, MD, USA
[25]Royal Netherlands Meteorological Institute (KNMI), De Bilt, the Netherlands
[26]Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Potsdam, Germany
[27]Department of Physics, University of Toronto, Toronto, ON, Canada
[28]Department of Chemistry, University of Waterloo, Waterloo, ON, Canada
[29]NASA Langley Research Center, Hampton, VA, USA
[†]deceased

AMTD
8, 6661–6757, 2015

Received: 19 March 2015 – Accepted: 24 May 2015 – Published: 02 July 2015

Correspondence to: D. Hubert (daan.hubert@aeronomie.be)

Published by Copernicus Publications on behalf of the European Geosciences Union.

AMTD
8, 6661–6757, 2015

Abstract

The ozone profile records of a large number of limb and occultation satellite instruments are widely used to address several key questions in ozone research. Further progress in some domains depends on a more detailed understanding of these data sets, especially of their long-term stability and their mutual consistency. To this end, we make a systematic assessment of fourteen limb and occultation sounders that, together, provide more than three decades of global ozone profile measurements. In particular, we consider the latest operational Level-2 records by SAGE II, SAGE III, HALOE, UARS MLS, Aura MLS, POAM II, POAM III, OSIRIS, SMR, GOMOS, MIPAS, SCIAMACHY, ACE-FTS and MAESTRO. Central to our work is a harmonized and robust analysis of the comparisons against the ground-based ozonesonde and stratospheric ozone lidar networks. It allows us to investigate, from the ground up to the stratopause, the following main aspects of data quality: long-term stability, overall bias, and short-term variability, together with their dependence on geophysical parameters and profile representation. In addition, it permits us to quantify the overall consistency between the ozone profilers. Generally, we find that between 20–40 km, the satellite ozone measurement biases are smaller than ±5 %, the short-term variabilities are better than 5–12 % and the drifts are at most ±5 % decade^{-1} (and ±3 % decade^{-1} for a few records). The agreement with ground-based data degrades somewhat towards the stratopause and especially towards the tropopause, where natural variability and low ozone abundancies impede a more precise analysis. A few records deviate from the preceding general remarks, in part of the stratosphere; we identify biases of 10 % and more (POAM II and SCIAMACHY), markedly higher single-profile variability (SMR and SCIAMACHY), and significant long-term drifts (SCIAMACHY, OSIRIS, HALOE, and possibly GOMOS and SMR as well). Furthermore, we reflect on the repercussions of our findings for the construction, analysis and interpretation of merged data records. Most notably, the discrepancies between several recent ozone profile trend assess-

AMTD
8, 6661–6757, 2015

ments can be mostly explained by instrumental drift. This clearly demonstrates the need for systematic comprehensive multi-instrument comparison analyses.

1 Introduction

Long-term global observations of the distribution and evolution of ozone are vital to improve our current understanding of atmospheric processes, and thereby to allow more robust projections of the recovery of the ozone layer and climate change. Measurements of the vertical profile of ozone have been carried out over the last few decades by a large number of instruments, operating in-situ or from remote vantage points, on the ground and in space (for an overview, see Hassler et al., 2014). These indisputibly show globally declining ozone levels during the 1980s and a large part of the 1990s in the lower and upper stratosphere ($\sim$ 5–7 % decade^{-1}), and to a lesser extent in the middle stratosphere as well (1–2 % decade^{-1}) (WMO, 2014; Harris et al., 2015). Furthermore, the observed loss rates are in excellent agreement with expectations for the chemical destruction of ozone by man-made halocarbons (WMO, 2014). The abundancies of these substances have decreased significantly over the past 15–20 years (WMO, 2011), as a result of the Montreal Protocol and its subsequent adjustments and amendments. It is therefore generally expected that the recovery of the ozone layer is currently ongoing, albeit in an atmosphere with concomitant increases in greenhouse gas concentrations and changes in residual circulation (Waugh et al., 2009; Oman et al., 2010). While observations provide substantial evidence for the levelling off of the downward trend around 1997 at most latitudes and altitudes, i.e. the first phase of recovery, it is less clear whether they support an upward trend in recent years (Harris et al., 2015). Whether the onset of the second stage has been detected (or not) is one of the key questions in current ozone research, a debate that is hampered by two factors. First, the small magnitude of the increases in ozone (a few percent) when compared to its natural variability. This can only be remedied by longer time series. And second, the

AMTD
8, 6661–6757, 2015

lack of appropriate knowledge of the uncertainties in the observational records. Shedding more light on the latter issue is the main objective of this paper.

Limb and occultation sounders are of prime interest for e.g. ozone profile trend assessments, as they provide near-global coverage at reasonably high vertical resolution. However, satellite instruments are rarely operational for much more than a decade, so their records are generally combined for long-term studies. The uncertainties (overall bias, short-term variability and long-term stability) in the resulting combined data set are an intricate combination of the uncertainties inherited from the contributing data sets and those introduced by the merging algorithm. Tummon et al. (2015) recently noted that the former source of error tends to dominate over the latter, thereby demonstrating the need for a detailed characterization of each individual record and especially of their mutual consistency.

Numerous validation studies have been published in recent years (for an overview, see Lambert et al., 2015), but some important gaps remain. First of all, there are no comprehensive multi-instrument assessments of most limb/occultation sounders using ground-based data as a reference. Also satellite intercomparison studies rarely cover more than a handful of records (see e.g. Dupuy et al., 2009; Jones et al., 2009; Laeng et al., 2014). Tegtmeier et al. (2013) conducted perhaps the most complete assessment so far, of the ozone climatologies from 18 sounders. Like most works, it was dedicated to the quantification of bias patterns and shorter-term variability, but not to a detailed assessment of the stability on decadal time scales. However, precise estimates of instrumental drift are crucial for a sound determination of the significance of trend results. Just a few (in some cases indirect) drift estimates are available from ground-based comparisons (e.g. Terao and Logan, 2007; Nair et al., 2012), and from satellite intercomparisons (e.g. Jones et al., 2009; Mieruch et al., 2012; Adams et al., 2014; Eckert et al., 2014). Moreover, no works comprise all the records considered in the recent trend assessments, by e.g. the World Meteorological Organisation (WMO, 2014) or within the SPARC/IO3C/IGACO-O3/NDACC (SI2N) initiative (for an overview, see Harris et al., 2015). Finally, the quality of auxiliary pressure and temperature pro-

AMTD
8, 6661–6757, 2015

files plays a role too, as it unavoidably affects the quality of ozone data when used for the conversion to another profile representation, a typical step in the merging process. At the moment, very little information on this latter aspect of data quality is available.

Our objective is to shed more light on these three missing pieces of information. We therefore perform an exhaustive assessment, from the ground up to the stratopause, of the latest operational Level-2 ozone profile data sets collected by 14 limb/occultation instruments over the period 1984–2013: SAGE II (v7), SAGE III (v4), HALOE (v19), UARS MLS (v5), Aura MLS (v3.3), POAM II (v6), POAM III (v4), OSIRIS (v5.07), SMR (v2.1), GOMOS (IPF 6), MIPAS (ML2PP 6), SCIAMACHY (SGP 5), ACE-FTS (v3) and MAESTRO (v1.2). Each satellite data set is compared to the observations by the ground-based ozonesonde and stratospheric ozone lidar networks, thereby acting as a pseudo-global, independent and well-characterized transfer standard. The robust analysis of co-located satellite-ground profile pairs allows us to quantify overall bias, short-term variability and long-term stability of the satellite records, and their dependence on altitude, latitude and season. Methodology and results for the native profile representation of each record are described in Sects. 3–5. In Sect. 6 we investigate whether the accompanying ancillary meteorological data impact ozone data quality in non-native representations.

The adoption of a harmonized analysis framework permits us to bring all single-instrument results together, and examine the mutual consistency between instruments of each quality indicator (Sect. 7). We report the general tendencies and several peculiarities, most notably a few instruments that drift significantly at some altitudes. Finally, we frame our findings within the broader context (Sect. 8), by commenting on current challenges related to verifying user requirements, and by highlighting the implications of our results for the design of merging schemes. Perhaps the most tangible outcome of our study is the successful interpretation of discrepancies in recent trend studies in terms of instrumental drift. It demonstrates that our work can contribute to a better exploitation of the limb and occultation ozone profile data sets. This should, in the end,

AMTD
8, 6661–6757, 2015

be beneficial not only for trend assessments and the related merging activities, but also for other applications, such as trend attribution studies or model evaluations.

2 Ozone profile records

What follows is a description of the ozone profile data sets used in this work, together with a compilation of the results of previous validation and intercomparison studies. We start by introducing the central data sets of this analysis, those collected by ozonesonde and stratospheric lidar instruments. Tables 1 and 2 list the ground stations considered for this work. Then, we describe the latest operational records of fourteen limb and occultation sounders, together spanning nearly three decades (October 1984 to May 2013). Table 3 summarises technical details of the satellite records. An exhaustive review of both satellite and ground-based instruments can be found in Hassler et al. (2014).

2.1 Ground-based network observations

2.1.1 Ozonesondes

In-situ measurements of ozone are carried out regularly by ozonesondes onboard small meteorological balloons launched at numerous sites around the world (Smit and ASOPOS-panel, 2014). The ozonesondes measure the vertical profile of ozone partial pressure with 100–150 m vertical resolution from the ground to the burst point of the balloon, typically between 30–35 km. An interfaced radiosonde provides the pressure (P), temperature (T) and GPS data necessary to geolocate each measurement or to convert ozone partial pressure to other units.

Different types of ozonesondes were developed over the years. Those still in use today are based on the electrochemical reaction of ozone with a potassium iodide sensing solution. Nowadays more than 80 % of the stations launch an electrochemical concentration cell (ECC) sonde (Komhyr, 1969). The Brewer-Mast (BM) sonde (Brewer

AMTD
8, 6661–6757, 2015

and Milford, 1960) has mostly been used by the early sounding stations with long data records, while the Japanese stations fly a carbon iodine cell (KC) sonde (Kobayashi and Toyama, 1966).

Laboratory tests and field campaigns indicated that between the tropopause and ~28 km these three sonde types produce consistent results when standard operating procedures are followed (Smit and ASOPOS-panel, 2014). The bias is smaller than ±5 % and the precision remains within ~3 % (Smit et al., 2007; Thompson et al., 2007; Deshler et al., 2008). Above 28 km, the bias increases for all types but the ECC instruments seem most reliable. At these altitudes the ECC sondes exhibit a small positive bias, whereas the BM and KC sondes tend to have a large negative and positive bias, respectively, and should therefore be used with care. Below the tropopause, because of lower ozone concentrations, the precision worsens slightly to 3–5 % for ECC and more for the other types. The tropospheric bias also becomes larger, ranging between ±5–7 %. Other factors besides ozonesonde type influence the data quality by up to a few percent. For a detailed overview we refer to Smit and ASOPOS-panel (2014).

We use the ozonesonde data acquired by the Network for the Detection of Atmospheric Composition Change (NDACC, http://www.ndacc.org), WMO's Global Atmospheric Watch network (GAW, http://www.woudc.org) and the Southern Hemisphere Additional Ozonesondes network (SHADOZ, http://croc.gsfc.nasa.gov/shadoz, Thompson et al. (2012)). Together these networks cover 82.5° N to 90.0° S, with the earliest ozonesonde flights dating back to the 1960s. Many participating stations provide soundings at least once a week. The sonde stations considered in this work are listed in Table 1, together with the total number of profiles over the analysis period.

2.1.2 Stratospheric ozone lidars

A differential absorption lidar (DIAL) operates mostly during clear-sky nights, emitting quasi-simultaneously two pulsed laser beams at wavelengths with a different ozone absorption cross-section. The backscattered signal is integrated over a few hours to retrieve the vertical distribution of ozone (Mégie et al., 1977). A stratospheric ozone

AMTD
8, 6661–6757, 2015

lidar system operates at 308 nm and 353 or 355 nm, which makes it sensitive from the tropopause up to about 45–50 km with a vertical resolution that worsens with altitude from 0.3 to 3–5 km, depending on the system. The profiles are reported as ozone number densities versus geometric altitude. We convert the data to other profile representations using *P*/*T* fields from the ERA-Interim reanalysis (Dee et al., 2011).

The DIAL technique is in principle self-calibrating since the ozone profile is retrieved directly from the returned signals without introducing instrumental constants, and therefore attractive for long-term monitoring. However, some artefacts can degrade the quality of the measurements, e.g. in the lower stratosphere interference by volcanic aerosols and saturation of the data acquisition system, and in the high stratosphere presence of signal-induced noise (Godin-Beekmann et al., 2003). Unreliable measurements can be discarded based on the reported precisions, which have been shown to be realistic (Godin et al., 1999). The bias and precision are about ±2 % between 20–35 km, increasing to ±5–10 % outside this altitude range where the signal-to-noise ratio deteriorates (Keckhut et al., 2004). The NDACC lidar retrieval algorithms were extensively intercompared, and the profile measurements validated against the mobile reference lidar (McGee et al., 1991), ozonesondes, microwave radiometers and various satellite instruments (Keckhut et al., 2004; Nair et al., 2012). The network appears homogeneous, with biases between stations within about 2 %. The agreement with space-based observations over the range 20–40 km is within ±5 %, with a decadal stability better than ±5 % decade^{-1} (Nair et al., 2012).

We consider data from thirteen stratospheric ozone lidars that have been or still are operational in the NDACC network, since the beginning of the 1990s. The lidar network covers 80.0° N to 66.7° S, but most sites are located in the Northern Hemisphere. The lidar stations considered in this work are listed in Table 2, together with the total number of profiles over the analysis period.

AMTD
8, 6661–6757, 2015

2.2 Satellite observations

2.2.1 SAGE II v7.0

The second Stratospheric Aerosols and Gas Experiment (SAGE II) (Mauldin et al., 1985) was deployed by the National Aeronautics and Space Administration (NASA) on its Earth Radiation Budget Satellite (ERBS) in a 57° inclined orbit. The seven-channel photometer recorded absorption spectra during solar occultations from which vertical profiles of ozone are retrieved using the onion-peeling technique, primarily using data from channel 3 (centered on 600 nm). Observations were made from October 1984 to August 2005 and cover 80° N to 80° S.

In this paper we use the v7.0 data set, produced with an algorithm primarily derived from the SAGE III v4.0 algorithm (Damadeo et al., 2013). Ozone number densities are reported on a fixed altitude grid with step 0.5 km. The retrieved profiles have 1 km vertical resolution and are recommended for scientific use from the cloud top up to 60 km. The data set includes pressure and temperature profiles from the Modern-Era Retrospective Reanalysis for Research and Applications (MERRA), extended to higher altitudes using the shape of profiles from the Global Reference Atmosphere Model (GRAM-95).

Early studies indicate that the v7.0 algorithm produces on average 1–2 % lower ozone values than its predecessor, v6.2, primarily as a result of changed spectroscopic data. In addition, it provides improved consistency in the vertical dependence of the comparisons to SAGE III v4.0, and it is expected to be more robust for trend analysis (Damadeo et al., 2013). Sunrise profiles tend to be biased low by 8–10 % relative to sunset data above 35 km (Kyrölä et al., 2013; Damadeo et al., 2014; Sakazaki et al., 2015). The interfering effect of aerosols and clouds on the v7.0 data quality can be reduced by applying a slight modification of the Wang et al. (2002) data screening (https://eosweb.larc.nasa.gov/project/sage2/sage2_release_v7_notes).

AMTD
8, 6661–6757, 2015

2.2.2 SAGE III v4.0

The third Stratospheric Aerosols and Gas Experiment (SAGE III) (Mauldin et al., 1998) was deployed by NASA on the Russian METEOR-3M platform in a polar sun-synchronous orbit, crossing the equator northward around 09:15 local time (LT). This SAGE II successor has a significantly improved spectral resolution and also exploits the solar occultation technique to retrieve vertical profiles of ozone, from measurements in the Chappuis band. Observations were made from March 2002 to November 2005, covering 50–80° N for sunset and 30–50° S for sunrise events.

In this paper we use the v4.0 data set (Thomason et al., 2010). In particular, we consider the product retrieved with the multiple linear regression technique. It should perform better in the lower stratosphere and troposphere than another, SAGE II v6.2 like retrieval technique (Wang et al., 2006). Ozone number densities are reported on a fixed altitude grid with step 0.5 km. The profiles have about 1 km vertical resolution and are suitable for scientific use between 6–85 km. The data set includes P/T profiles from reanalyses by the National Centers for Environmental Prediction (NCEP), extended at high altitudes using the shape of GRAM-95 profiles.

Comparison of the previous data set, v3.0, to ground- and space-based observations showed agreement within 5 % down to 17 km and a high bias up to 10 % at 13 km (Wang et al., 2006). The v3.0 record furthermore appears stable relative to that of SAGE II v6.2. Unreliable measurements were removed by the SAGE III team, so no additional screening is required (NASA LaRC, 2004).

2.2.3 HALOE v19

The Halogen Occultation Experiment (HALOE) (Russell III et al., 1993) was launched by NASA on the Upper Atmosphere Research Satellite (UARS) in a 57° inclined orbit. The solar occultation technique is exploited to retrieve vertical profiles of ozone from the broadband radiometer measurements in the 9.6 μm band. Observations were made from October 1991 to November 2005 and cover 80° N to 80° S.

AMTD
8, 6661–6757, 2015

We consider the latest data release, version 19. Ozone volume mixing ratio (VMR) is reported on a fixed pressure grid with 30 levels per decade. The retrieved profiles have about 2–3 km vertical resolution (Nazaryan et al., 2005) and are recommended for scientific use between 15 and 60 km. The data set includes P/T profiles, composed of retrieved and NCEP reanalysis information.

Validation studies show a general agreement with other satellite and ground-based instruments within 5–10 % above 20 km (Morris et al., 2002; Nazaryan et al., 2005). In the upper stratosphere a bias of ∼ 5 % exists between sunset and sunrise occultations (Sakazaki et al., 2015). Furthermore, the drift of HALOE relative to SAGE II is in general insignificant and remains below 10 % decade^{-1} (Morris et al., 2002; Nazaryan et al., 2005). Cirrus clouds were found to have an interfering effect on the data quality, and this effect was reduced by applying the screening suggested by Hervig and McHugh (1999).

2.2.4 UARS MLS v5

The UARS platform also carried the Microwave Limb Sounder (UARS MLS) (Barath et al., 1993), whose three radiometers recorded the microwave emission spectrum at the limb. The data from the 205 GHz radiometer are used to extract vertical profiles of ozone with an optimal estimation method. Observations over a latitude range of 34–80° were carried out from September 1991 to July 1999, but with a much reduced duty cycle after 1994.

In this paper we use the data set produced with the version 5 algorithm (Livesey et al., 2003). Ozone VMR is reported on a fixed pressure grid with 6 levels per decade. The retrieved profiles have about 3–5 km vertical resolution and are suitable for scientific use between 100–0.22 hPa. The data set includes retrieved profiles of temperature and geopotential height (GPH).

Comparisons to correlative measurements indicate a positive ozone bias of 2–4 % for pressures less than 50 hPa, increasing to 10–15 % at lower altitudes (Livesey et al., 2003). The absolute pointing of UARS MLS drifted upwards by 600 m from 1991 to

AMTD
8, 6661–6757, 2015

1997 (Livesey et al., 2003, Fig. 1). This may impact ozone if the GPH profiles are used for conversions between vertical coordinates. Prescriptions to screen ozone data of lower quality are detailed in Livesey et al. (2003). In particular, data after the 15 June 1997 switch-off of the 63 GHz radiometer should be discarded for trend analyses, due to possible biases between both operational modes.

2.2.5 Aura MLS v3.3

The second Microwave Limb Sounder (Aura MLS) (Waters et al., 2006) is deployed by NASA on its EOS-Aura platform in a polar sun-synchronous orbit, crossing the equator northward around 13:45 LT. Its five radiometers record microwave emission at the limb from which vertical profiles of ozone are retrieved, mainly using the 240 GHz radiometer data. Observations are ongoing since August 2004 until today and cover 82° N to 82° S.

In this paper we investigate the v3.3 data set (Livesey et al., 2013b). Ozone VMR is reported on a fixed pressure grid with 12 levels per decade up to 1 hPa and coarser at higher altitudes. The retrieved profiles have about 2.5–4 km vertical resolution and are suitable for scientific use from 262–0.02 hPa. The data set includes GPH and T profiles retrieved by the same processor.

The previous data release, v2.2, was extensively compared to satellite and ground-based instruments, showing agreement within 5–10 % in most of the stratosphere, and a positive bias up to 20–30 % at 216 hPa (Jiang et al., 2007; Froidevaux et al., 2008). The v3.3 ozone data are within 1–2 % of v2.2, except in the tropical upper troposphere lower stratosphere (UTLS) (Livesey et al., 2013a). Although the high bias at 216 hPa is now much reduced, a persistent vertical oscillation pattern appears in the overlying levels, in particular at low latitudes. Comparisons to MIPAS and OSIRIS data indicated relative drifts in ozone less than ±3 % decade^{-1} (Eckert et al., 2014; Adams et al., 2014). However, the absolute pointing of Aura MLS v3.3 drifts downwards, but not at a constant rate. The 100 hPa level decreased by 100 m in GPH between 2005 and 2009, and stabilized after that (Livesey et al., 2015), corroborating an earlier report for v2.2 data of a large drop in GPH at the start of the mission (Schwartz et al., 2008). This

AMTD
8, 6661–6757, 2015

may impact ozone if the GPH profiles are used for conversions between vertical coordinates. Recommendations for data screening are described in Livesey et al. (2013b), and implemented for this work.

2.2.6 POAM II v6

The second Polar Ozone and Aerosol Measurement (POAM II) instrument (Glaccum et al., 1996) was deployed by the Naval Research Laboratory (NRL) on the third French Satellite Pour l'Observation de la Terre (SPOT-3) platform in a polar sun-synchronous orbit, crossing the equator southward at 10:30 LT. The nine-channel photometer recorded absorption spectra during solar occultations from which vertical profiles of ozone are retrieved using measurements from channel 4 (centered on 601.4 nm). Observations were made from September 1993 to November 1996, at mid to high latitudes.

In this paper we use the latest data release, version 6 (Lumpe et al., 1997). Ozone number densities are reported on a fixed altitude grid with step 1 km. The retrieved profiles have about 1 km vertical resolution and are recommended for scientific use between 15–50 km. The data set includes P and T profiles from UKMO.

Comparisons to satellite and ground-based observations indicate a negative bias of about 5 % above 20 km, and possibly larger at lower altitudes (Rusch et al., 1997; Deniel et al., 1997; Danilin et al., 2002). Large polar stratospheric clouds (PSC) interfere with the retrieval, but have been screened by removing outliers that correlate with large 1 μm aerosol extinction measurements.

2.2.7 POAM III v4

The third Polar Ozone and Aerosol Measurement (POAM III) instrument (Lucke et al., 1999) was deployed by NRL on the fourth French Satellite Pour l'Observation de la Terre (SPOT-4) platform in a polar sun-synchronous orbit, crossing the equator southward at 10:30 LT. The instrument design and solar occultation measurement principle

AMTD
8, 6661–6757, 2015

are similar to that of POAM II. Vertical profiles of ozone are retrieved using measurements from channel 4 (centered on 603 nm). Observations were made from March 1998 to December 2005, at mid to high latitudes.

In this paper we use the latest data release, version 4 (Lumpe et al., 2002). Ozone number densities are reported on a fixed altitude grid with step 1 km. The retrieved profiles have about 1–2 km vertical resolution and are recommended for scientific use between 10–60 km. The data set includes P and T profiles from UKMO.

Version 4 offers minor improvements relative to the previous release v3. The latter was extensively validated using observations from aircraft, balloons and satellite instruments (Randall et al., 2003), showing an agreement of ±5 % from 13 to 60 km and a 15–20 % high bias below 12 km. However, there are indications that the Northern Hemisphere (orbital sunrise) data are biased low by up to 5 % relative to the Southern Hemisphere (orbital sunset) data between 25–50 km. No signs of long-term drifts relative to other instruments were noticed so far. A description of the screening procedure for data contaminated by sunspot and aerosol feedback errors can be found in Naval Research Lab (2005).

2.2.8 OSIRIS v5.07

The Optical Spectrograph and InfraRed Imaging System (OSIRIS) (McLinden et al., 2012) is deployed by the Canadian Space Agency (CSA) on the Swedish Odin platform in a polar sun-synchronous orbit, crossing the equator southward around 06:30 LT. The optical spectrograph records limb-scattered sunlight from which vertical profiles of ozone are retrieved using a multiplicative algebraic reconstruction technique that combines measurements from the Chappuis and Hartley-Huggins bands. Observations have been ongoing since October 2001 until today, and cover mainly the sunlit hemisphere up to 82° latitude.

In this paper we use the v5.07 data set (Degenstein et al., 2009). Ozone number densities are reported on a fixed altitude grid with step 1 km. The retrieved profiles have about 1–2 km vertical resolution and are suitable for scientific use from the cloud

AMTD
8, 6661–6757, 2015

top up to 60 km. The data set includes P and T profiles from operational ECMWF analyses.

Version 5.07 ozone was compared to several satellite and ground-based data records (Adams et al., 2013, 2014), with a general agreement better than 5 % above 21 km. At lower altitudes a positive bias of up to 20 % appears. The positive bias of 2–4 % around 22 km at high latitudes is possibly a result of the interference of aerosols. In addition, the data set seems to be stable relative to other data records, mostly remaining within 3 % per decade (Adams et al., 2014). Any data affected by clouds or cosmic rays were removed by the OSIRIS team from the released data set, following the procedure in Adams et al. (2013).

2.2.9 SMR v2.1

The Sub-Millimetre Radiometer (SMR) (Frisk et al., 2003) is the second instrument on the Swedish Odin platform. Its four receivers record microwave emission at the limb from which vertical profiles of ozone are retrieved with an optical estimation method. Observations cover 82° N to 82° S and are ongoing since June 2001 until today. Before April 2007 half of the observations were dedicated to astronomical purposes, with less atmospheric data as a consequence.

In this paper we use the Chalmers v2.1 ozone profiles retrieved from the 501.8 GHz receiver data (Urban et al., 2005). Ozone VMR is reported on a variable altitude grid with a step of ∼ 1.5 km in the stratosphere. The retrieved profiles have about 3 km vertical resolution and are suitable for scientific use from ∼ 18 up to 60 km. The data set includes P and T profiles from operational ECMWF analyses.

Comparisons to ground-based, balloon and satellite data show a typical agreement within 10 % in the stratosphere, with a small negative bias (Urban et al., 2005; Jones et al., 2007; Jégou et al., 2008). Due to the weak ozone line the signal-to-noise ratio is quite low and the resulting single-scan precision considerable (about 0.25–1.5 ppmv or 20–25 %). When noise is an issue, Urban et al. (2005) suggest to average profiles

AMTD
8, 6661–6757, 2015

in the logarithmic VMR domain. Lower quality measurements were screened with the prescriptions found in Jones et al. (2009).

2.2.10 GOMOS IPF 6.01

The instrument Global Ozone Monitoring by Occultation of Stars (GOMOS) (Bertaux et al., 2010) was deployed by the European Space Agency (ESA) on its Environmental satellite (Envisat) in a polar sun-synchronous orbit, crossing the equator southward around 10:00 LT. Its medium resolution spectrometer recorded absorption spectra between 250–680 nm during stellar occultations from which vertical profiles of ozone are retrieved using the onion-peeling technique. Observations covered 90° N to 90° S between July 2002 and April 2012. An instrument problem in January 2005 decreased the measurement frequency until the end of the mission.

In this paper we use the latest data set by ESA's operational IPF 6.01 processor (Kyrölä et al., 2010). Ozone number densities are reported on an altitude grid with variable spacing of 0.2–1.6 km, depending on the obliquity of the occultation. The retrieved profiles have about 2–3 km vertical resolution and are recommended for scientific use between 15–100 km. The data set includes P and T profiles from operational ECMWF analyses (below 1 hPa) and from the MSIS90 model (above 1 hPa).

Validation studies of the IPF 6.01 data set are under way, indicating slightly lower ozone values and fewer outliers relative to IPF 5.00. Studies of the previous data releases showed that the data quality depends on the illumination condition. Although the star characteristics are expected to be important as well, this could so far not be confirmed by comparisons to ground-based data (van Gijsel et al., 2010). The highest quality profiles, obtained from occultations of hot and bright stars on the dark limb, agree with correlative data within 5–10 % from 20 km up to 40 km. Below 20 km, a large positive bias appears due to the interference of ozone and aerosol retrieval with aerosol models (Tamminen et al., 2010). Nair et al. (2011) reported negative drifts up to 18 % decade^{-1} near 20 km relative to the lidar at Observatoire de Haute-Provence (43.9° N, 5.7° E), while van Gijsel et al. (2010) found no obvious drift rela-

AMTD
8, 6661–6757, 2015

tive to ozonesonde, lidar and microwave network data in the yearly bias results. The screening procedure recommended by the GOMOS team is implemented here (ESA, 2012a).

2.2.11 MIPAS ML2PP 6.0

The Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) (Fischer et al., 2008) was another ESA instrument aboard the Envisat platform. Its Fourier transform spectrometer recorded thermal IR emission at the limb from which vertical profiles of ozone are retrieved by means of a non-linear least-squares fit modified by the Levenberg Marquardt method for minimization of the chi-square function and using a global fit strategy. Observations covered 90° N to 90° S between July 2002 and April 2012. In March 2004, the instrument experienced a major anomaly, but operations resumed in January 2005 with reduced spectral resolution and increased vertical sampling. The second part of the mission is typically referred to as the Optimized Resolution[1] (OR)phase, whereas the first part is called the Full Resolution (FR) phase.

Several MIPAS ozone profile data sets exist (Laeng et al., 2015), here we focus on the latest version of ESA's operational ML2PP 6.0 processor (Raspollini et al., 2013). We furthermore consider only the nominal measurement mode profiles, which constitute about 70 % of the total data set. Ozone VMR is reported on a variable pressure grid (corresponding to retrieved pressure values at the measurement tangent altitudes) with spacing between 2–4 km in the stratosphere. The retrieved profiles have about 2–4 km vertical resolution, depending on whether in OR or FR phase, and are suitable for scientific use between 6–68 km. The data set includes P and T profiles retrieved by the same processor. In addition, vertical averaging kernels are provided for each profile.

Extensive validation studies of ML2PP 6.0 data pointed to a positive bias of roughly 5 % in most of the stratosphere, which is also present for three other scientific MIPAS processors (Laeng et al., 2015). Ceccherini et al. (2013) anticipate a bias between

[1] Some authors prefer the term Reduced Resolution (RR).

AMTD
8, 6661–6757, 2015

the observations in the OR and FR phase, as the reduction in spectral resolution necessitated the selection of different retrieval microwindows. Such a possible jump in the timeseries needs to be considered in trend studies that use the complete data set (Raspollini et al., 2013). Eckert et al. (2014) reported a small negative drift of up to -3 ppmv decade^{-1} ($\sim$ 2–4 % decade^{-1}) between scientific MIPAS data from Karlsruhe Institute of Technology and other satellite instruments such as Aura MLS and OSIRIS. Since the bias between FR/OR profiles is included in their regression model, the drift is believed to be related to detector aging effects. We implemented the screening recommendations of ESA (2012b).

2.2.12 SCIAMACHY SGP 5.02

The SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY (SCIAMACHY) (Bovensmann et al., 1999) was the third atmospheric chemistry instrument onboard the Envisat platform. Vertical profiles of ozone are retrieved from measurements of limb-scattered UV-visible sunlight. Observations were made from August 2002 to April 2012 and covered 82° N to 82° S.

Several SCIAMACHY limb ozone profile data sets exist (Rozanov et al., 2007). Here we focus on the latest version of the operational SGP 5.02 processor (Lichtenberg, 2011). Since it does not exploit the UV measurements the ozone retrieval is limited to 40 km. Ozone number density is reported on a fixed altitude grid with 1.75–3.5 km spacing in the stratosphere. The retrieved profiles have about 3 km vertical resolution and are suitable for scientific use between 15–40 km. The data set includes P and T profiles from the climatology by McLinden et al. (2002). In addition, vertical averaging kernels are provided for each profile. A detailed bottom-up analysis of the error budget of the scientific ozone data set by IUP Bremen indicated a random component of 10–15 % (Rahpoe et al., 2013).

Early validation studies on a limited data set indicated that the SGP 5 profiles agree within 5–10 % with ground-based measurements, with no signs of very large significant drifts (ESA, 2013). In the Tropics the bias is larger, peaking up to +20 % around 18 km.

AMTD
8, 6661–6757, 2015

2.2.13 ACE-FTS v3.0

The Atmospheric Chemistry Experiment Fourier Transform Spectrometer (ACE-FTS) instrument (Bernath et al., 2005) is deployed by CSA on its SCISAT platform, in a 74° inclined orbit. The Fourier transform spectrometer records high resolution absorption spectra in the mid infrared during solar occultations, from which vertical profiles of ozone are retrieved using measurements around 10 μm. Operational observations are ongoing since February 2004 until today and cover 85° N to 85° S, but mostly high latitudes.

In this paper we study the latest v3.0 data set (Boone et al., 2013) reporting ozone VMR on a fixed altitude grid with step 1 km. The data on a coarser, profile-dependent measurement grid are not considered here. The retrieved profiles have about 3–4 km vertical resolution and have an altitude range from the cloud top up to 95 km. The data set includes P and T profiles retrieved by the same processor.

Initial ACE-FTS intercomparison studies show that v3.0 data between 15–35 km is similar to that of the previous data release, v2.2 update. Outside this altitude range the new ozone data is reduced on average by 5–10 % (Waymark et al., 2013). It is anticipated that the lower ozone above 35 km will decrease the positive bias identified in extensive comparisons of the v2.2 update release to satellite and ground-based measurements (Dupuy et al., 2009). These indicated a typical positive bias of about 5 % between 15–45 km, increasing to 20 % in the upper stratosphere. Sakazaki et al. (2015) recently reported sunrise-sunset biases in the upper stratosphere of about 5 %. Due to problems with auxiliary inputs the v2.2 and v3.0 data products from after September 2010 should be discarded (Boone et al., 2013). New versions (v2.5/3.5) have been produced that correct this issue and continue these data sets to the present. Comparisons to MIPAS data do not indicate a large relative drift, although the uncertainty is quite large due to a short overlap period (Eckert et al., 2014).

AMTD
8, 6661–6757, 2015

2.2.14 MAESTRO v1.2

The Measurements of Aerosol Extinction in the Stratosphere and Troposphere Retrieved by Occultation (MAESTRO) (McElroy et al., 2007) is the second instrument onboard SCISAT, and optically nearly co-aligned with ACE-FTS. Its dual spectrophotometer records absorption spectra in the visible and ultraviolet during solar occultations from which vertical profiles of ozone are retrieved. Operational observations are ongoing since February 2004 and cover 85° N to 85° S, but most data are taken at high latitudes.

In this paper we study the v1.2 data set from the visible (VIS) spectrometer, which is in agreement with the UV ozone profile data below 30 km but provides better data at higher altitudes (McElroy et al., 2007). Ozone VMR is reported on a variable altitude grid with step between 0.3–2 km in the stratosphere, depending on the obliquity of the occultation. The retrieved profiles have about 1.5 km vertical resolution and cover from the cloud top up to 50 km. The data set does not include P and T profiles, these can be obtained from coincident, co-aligned ACE-FTS measurements.

Comparisons of the VIS ozone profiles to ozonesonde and satellite measurements indicated a mean agreement of about 5–10 % over 20–40 km (Kar et al., 2007; Dupuy et al., 2009). In the UTLS MAESTRO underestimates ozone by up to 50 % with respect to all correlative instruments. Above 35 km, sunrise ozone measurements are up to 20 % lower than sunset observations. This bias could be caused by an error in the altitude registration due to timing errors between MAESTRO and ACE-FTS. All data after September 2010 should be discarded because the ACE-FTS P and T data, used for the tangent pointing of MAESTRO, are affected by problems with auxiliary inputs. The procedure used by Kar et al. (2007) can be used for data screening.

AMTD
8, 6661–6757, 2015

3 Analysis approach and data preprocessing

A careful design of the analysis allows us not only to obtain robust estimates of the data quality of the individual satellite records but also, and this is one of our primary objectives, to assess their mutual consistency. Prerequisite to achieving these goals is a good understanding of the metrological aspects. Our analysis approach is therefore based on three principles that reduce confounding methodological biases. First of all, we use a single analysis and software framework. Second, all satellite records are compared to the same reference data, from ground-based observations. And finally, the manipulation of satellite data is kept to a strict minimum. In this section we describe the general aspects of the analysis. A detailed account of how decadal stability, bias and short-term variability are estimated follows in Sects. 4 and 5.

The ozonesonde and lidar networks provide vertical ozone profiles of well-documented quality and serve as suitable transfer standards on a pseudo-global scale and from the ground up to the stratopause. We compare the satellite profiles to co-located ground-based measurements in relative units

$$\Delta x_{ij}(l) = 100 \times \frac{x_{ij,\mathrm{sat}}(l) - x'_{ij,\mathrm{gnd}}(l)}{x'_{ij,\mathrm{gnd}}(l)}. \quad (1)$$

Here, $x_{ij,\mathrm{sat}}(l)$ and $x'_{ij,\mathrm{gnd}}(l)$ represent respectively satellite and (vertically smoothed and representation-transformed) ground-based ozone at grid level l of co-location pair i for correlative instrument j. The advantage of relative over absolute differences is that atmospheric cycles are mostly divided out and that it allows a direct comparison between the results in different ozone coordinates. A disadvantage, however, is that relative differences are sensitive to low ozone values, often leading to pronounced values in and below the UTLS and in the upper stratosphere.

The uncertainty on Δx is determined by several factors besides pure measurement and retrieval uncertainties ($\mathbf{S}_{x_{\mathrm{sat}}}$, $\mathbf{S}_{x_{\mathrm{gnd}}}$) because satellite and ground-based instruments perceive the variable atmosphere differently. Vertical and horizontal resolutions differ

AMTD
8, 6661–6757, 2015

and the probed air masses rarely coincide perfectly in space and time. In addition, the comparison can only be done when both profiles are expressed in the same representation. As a result, the total comparison error budget contains terms related to the differences in smoothing, the spatio-temporal mismatch of the co-locations and the auxiliary data used to transform between profile representations. When correlations between the terms are disregarded, the total error covariance matrix (including systematic and random components) becomes $\mathbf{S}_{\Delta x} = \mathbf{S}_{x_{\text{sat}}} + \mathbf{S}_{x_{\text{gnd}}} + \mathbf{S}_{\text{smoothing}} + \mathbf{S}_{\text{mismatch}} + \mathbf{S}_{\text{auxiliary}}$ (von Clarmann, 2006). When Δx data are averaged or regressed the co-located profile sample may not be sufficiently representative of the actual state of the studied parameter (ozone differences). Toohey et al. (2013) recently showed the importance of $\mathbf{S}_{\text{sampling}}$ for trace gas climatologies. Estimating the sampling uncertainty for our purposes remains outside the scope of this work however.

The next paragraphs describe the data preprocessing scheme in which the reduction of smoothing, mismatch and auxiliary uncertainties play a central role. Preprocessing starts off by removing the unreliable measurements following the guidelines of the data providers. Table 3 lists the recommended screening procedure references for the satellite records. Ground-based data are filtered using general criteria, removing measurements with larger uncertainties: below the 5 hPa level ($\sim$ 33 km) for ozonesondes and outside the 15–47 km range for lidars. In addition, we reject measurement levels with clearly unphysical readings ($O_3 < 0$, $p < 0$ hPa, $T < 0$ K or $T > 400$ K) or during unrealistic jumps in pressure ($\mathrm{d}p/\mathrm{d}t > 0$ and $\mathrm{d}z > 0.1$ km). Entire profiles are discarded from further analysis when (a) more than half of the levels are tagged bad, or (b) less than 30 levels are tagged good.

The definition of the co-location window is a trade-off between mismatch and sampling uncertainties. We found that a maximum horizontal distance Δr of 500 km between the profiles is optimal, given the typical horizontal resolution of order a few hundred km of the satellite and ground-based measurements. The maximal temporal separation Δt is 6 h for UARS MLS, Aura MLS, MIPAS and SMR, and 12 h for the other instruments. When multiple satellite profiles are present in the co-location win-

AMTD
8, 6661–6757, 2015

dow around a ground-based profile only the pair closest in space and time is retained, defined by $\sqrt{\Delta r^2 + v_{\text{wind}}^2 \Delta t^2}$ with $v_{\text{wind}} = 100\,\text{km}\,\text{h}^{-1}$. Multiple co-locations occur mostly between polar orbiting instruments and high latitude stations. Mismatch uncertainties $\mathbf{S}_{\text{mismatch}}$ increase when and where atmospheric inhomogeneities are larger. The systematic component is generally less than a few percent, while the random component is typically 5 % but can reach 20 % during Antarctic ozone hole conditions (Cortesi et al., 2007; De Clercq, 2009).

There is no well-established method in the community to remove the horizontal component of the smoothing error. Instead we refer to the model-based estimates for the specific case of MIPAS comparisons (Cortesi et al., 2007; De Clercq, 2009), which showed that the horizontal smoothing uncertainty mainly has a random nature and is of similar magnitude as the mismatch uncertainty. The vertical component on the other hand can be mostly removed by smoothing the ground-based profiles. We use a triangular response function with base width that follows the altitude-dependent satellite resolution. The exception is the MIPAS analysis, for which we smoothed with the vertical averaging kernel (AK) and a-priori of the co-located MIPAS profile. Such an AK smoothing was initially also tried for the SCIAMACHY analysis. Unfortunately it introduced peculiar and unexpected vertical oscillations in the comparisons, so we resorted to the triangular method for SCIAMACHY. The comparison results are somewhat sensitive to the shape of the smoothing function (we also tried rectangular and Gaussian windows), but most of the vertical smoothing error is removed. We estimate that the residual vertical smoothing uncertainty is less than a few percent.

In a final preprocessing step the data are transformed to the same profile representation, defined by the ozone coordinate (number density or VMR), the vertical coordinate (altitude or pressure) and the levels of the vertical grid. Here, we focus on the satellite instrument's native representation, see Fig. 1. Mainly because it is closest to the retrieved information, but also because users will use it as a starting point to convert to another representation if their application requires that. They can use the auxiliary pressure and/or temperature profiles provided along with the ozone profiles in many

AMTD
8, 6661–6757, 2015

satellite records for this purpose. As we have seen in Sect. 2.2, these auxiliary data originate from different sources which may lead to a representation-dependence of the mutual consistency of the satellite data quality. This is discussed further in Sect. 6. Until then, only the correlative data are converted when needed. Ozonesonde data are transformed with the help of P and T measurements from the attached radiosonde, and lidar data using ERA-Interim fields. The quality of these ancillary data has been investigated by various authors (e.g. Sun et al. (2013); Stauffer et al. (2014); Simmons et al. (2014); Inai et al. (2015)). The regridding to the satellite's vertical grid is based on a pseudo-inverse interpolation method (Calisesi et al., 2005). Since the ground-based grid is more finely resolved than the satellite grid, the associated regridding uncertainties are generally negligible. We note that the SMR, GOMOS, MIPAS and MAESTRO profiles are inevitably regridded as well, since the grid is variable. In these cases the levels of the comparison grid are selected to reflect the average spacing between two lines of sight.

To conclude this section we repeat the importance of using a single analysis and code framework. Apart from some unavoidable preprocessing steps, the data and analysis flow is identical for all fourteen satellite comparison studies. In this, way the methodological biases are mostly identical and, hence, unlikely responsible for eventually observed differences between the satellite records. This approach will be exploited in Sect. 7. The next two sections present a detailed assessment of the bias, the short-term variability and the decadal stability of each satellite record individually.

4 Decadal stability

We estimate the decadal stability of satellite data through a robust analysis of the satellite-ground comparison time series. This is a two-step process, in which the drift is first estimated at each ground station and subsequently averaged over the ozonesonde and lidar networks. The focus of this section is on the decadal stability of the individual satellite records, in their native profile representation. Later on we expand the discus-

AMTD
8, 6661–6757, 2015

sion to the drift consistency between profile representations (Sect. 6) and between satellite records (Sect. 7).

4.1 Methodology

4.1.1 Time series analysis at individual stations

We first estimate the drift of the satellite data at each ground station. The comparison time series can contain large gaps and/or outliers, see e.g. the GOMOS comparisons in Fig. 2 (top panel). Hence robust techniques are needed to estimate not only the drift but also its uncertainty (Muhlbauer et al., 2009; Croux et al., 2004). To this end we use an iterative Tukey-bisquare reweighted least-squares procedure to fit the daily averaged relative difference time series to a linear regression model

$$\Delta x_{ij}(l) = \alpha_j(l)\,(t_i - t_0) + \beta_j(l) + e_{ij}(l). \quad (2)$$

With $\Delta x_{ij}(l)$ as in Eq. (1) at time t_i and grid level l, and the fit residual $e_{ij}(l)$. In this model, the fit parameter $\alpha_j(l)$ represents the linear drift of the satellite data relative to the ground-based record j, whereas $\beta_j(l)$ is the bias between both records at reference time t_0. Time series with less than ten data points are not regressed. The significance of the estimated $\hat{\alpha}(l)$ is tested using a robust estimate of its standard deviation $\hat{\sigma}_\alpha(l)$ proposed by Street et al. (1988), a slightly modified version of the ordinary least-squares expression. Figure 2 illustrates three time series with superimposed regression results (left panels, blue line) and the corresponding 95 % confidence level intervals for $\hat{\alpha}(l)$ (right panels, vertical dashed blue lines).

4.1.2 Aggregation into ground network average

In a second step, the drift estimates $\hat{\alpha}_j(l)$ are averaged over various ground stations $j = \{1, \ldots, N\}$. None of the satellite records exhibit a clear latitudinal structure of drift (see for instance HALOE and Aura MLS at 25 km in Fig. 3). Therefore, we average

AMTD
8, 6661–6757, 2015

the results over the entire sonde and over the entire lidar network. The figure also shows a clear variability in the regression uncertainty across the network, so each station estimate is weighted by the inverse of its variance $w_j(l) = \hat{\sigma}_{\alpha,j}^{-2}(l)$. The network-averaged drift

$$\bar{\alpha}(l) = \frac{\sum_j w_j(l)\,\hat{\alpha}_j(l)}{\sum_j w_j(l)}, \tag{3}$$

has a standard deviation $\sigma_{\bar{\alpha}}(l) = 1/\sqrt{\sum_j w_j(l)}$.

The single-site drift uncertainties alone do not always explain the observed variability of the drift estimates over the network. When the number of stations is large enough ($N \gtrsim 20$) the distribution of $v_j(l) = (\hat{\alpha}_j - \bar{\alpha})/\hat{\sigma}_{\alpha,j}$ should have unit variance for realistic estimates $\hat{\sigma}_\alpha$ of the variance of $\hat{\alpha}$. That is typically not the case for the dense samplers, that tend to have larger variance as illustrated for instance for Aura MLS in Fig. 3 (right). This suggests an unaccounted for source of uncertainty, likely related to inhomogeneities or differences in sampling across the ground-based network. We follow an ad-hoc approach to incorporate this unknown component, by scaling the uncertainty up

$$\sigma^*_{\bar{\alpha}}(l) = \kappa(l) \times \sigma_{\bar{\alpha}}(l), \tag{4}$$

so that the reduced $\chi^2(l) = \sqrt{\frac{1}{N-1}\sum_j v_j(l)^2}$ becomes unity. We also assume, conservatively, that the original regression uncertainty does not overestimate the true uncertainty, hence $\kappa(l) = \max\left\{\chi^2(l), 1\right\}$. In the following, this adjusted standard deviation $\sigma^*_{\bar{\alpha}}(l)$ is used to test the significance of the drift averages at the 5 % level.

4.1.3 Sensitivity to analysis parameters

The importance of correct single station uncertainties $\hat{\sigma}_{\alpha,j}(l)$ is evident for the calculation of both the weighted mean and its uncertainty. The analytic expression of Street

AMTD
8, 6661–6757, 2015

et al. (1988) was therefore cross-checked with a bootstrapping technique (Efron and Tibshirani, 1986). For every time series, we regressed 500 bootstrapped samples to reconstruct the distribution of $\hat{\alpha}_j(l)$ (Fig. 2, right). The 2.5 % and 97.5 % quantiles define the 95 % confidence interval (light red area) that is compared to the analytic expression (vertical dashed blue lines). The results confirmed the validity of the analytic uncertainties. Replacing the latter by the bootstrap-derived uncertainties in Eq. (3) typically changes $\bar{\alpha}$ by less than 0.3 % decade^{-1}. Figure 2 (left) also illustrates the outcome of other sensitivity checks, such as changing the temporal resolution of the time series prior to regression (from daily to monthly, green curve) or adding a 1-year harmonic component to the regression model (orange curve). The results were very consistent as well, which shows the robustness of the baseline method.

4.2 Selection of ground sites

Several lidar stations were discarded from the stability analysis, either because of poorer sampling or because of less well understood features in the time series. All high latitude sites (no data during polar summer) and several mid-latitude records fall in the first category. Two stations were assigned to the second category, based on the vertical drift structure of various satellite instruments (Fig. 4). The six ozone lidar records in this figure were also studied by Nair et al. (2012). At Hohenpeißenberg (47.8° N, 11.0° E) and Tsukuba (36.0° N, 140.1° E) the earlier SAGE II data exhibit a near-zero drift (grey line), whereas the more recent satellite records all drift in a similar way relative to the lidar data above 25 km (coloured lines). This suggests a systematic change in both lidar records in recent years. Inspection of the time series indeed showed that the Hohenpeißenberg lidar reported more ozone for a few years after 2007. Similarly, the Table Mountain lidar (34.4° N, 117.7° W) (McDermid et al., 1990) measured higher ozone relative to satellite instruments during 2007–2008 (Nair et al., 2012). This bias disappeared again in later years to leave the satellite drift estimates nearly unchanged. One exception is Aura MLS since the temporary lidar bias occurred close to the start of the mission. We nevertheless keep this lidar station for the drift analysis, together with

AMTD
8, 6661–6757, 2015

the data from Observatoire de Haute-Provence (43.9° N, 5.7° E), Mauna Loa (19.5° N, 155.6° W) (McDermid et al., 1995) and Lauder (45.0° N, 169.7° E). A similar procedure was followed to discard about 20 ozonesonde records from the drift analysis. Since there are many sonde stations the network-average should be quite representative of the actual satellite drift. Later on, we describe the remarkable agreement between the ozonesonde and lidar-derived drift results, strengthening the confidence in the stability of these ground networks (Fig. 5). Tables 1 and 2 list the stations used for the drift analysis (last column).

4.3 Results

Below we report on the vertical structure of the network-averaged drift (and its significance) of each satellite record, summarised in Fig. 5 and Table 4.

4.3.1 SAGE II

The very long record of SAGE II, spanning 21 years, allows a detailed analysis of its stability. The most precise estimates have 1σ uncertainty of 0.8 % decade^{-1} and 1.6 % decade^{-1} at sonde and lidar stations respectively. On average the single station drift uncertainty is $\sim$ 4 % decade^{-1}. The drift results are furthermore very consistent from one station to another, with a spread of 2–3 % decade^{-1} at 25 km (Fig. 4, grey line). The sonde and lidar derived estimates are very consistent as well. When aggregated over the entire ground network a significant SAGE II drift should be detectable at the 1–2 % decade^{-1} level, depending on altitude.

In the middle and upper stratosphere, between 20–40 km, the average drift is generally slightly negative (Fig. 5). The negative drift remains smaller than 1–2 % decade^{-1} and is therefore not significant. At lower altitudes the drift becomes gradually more pronounced, but is never significant either as a result of the increased atmospheric variability or noise in the SAGE II record. We therefore conclude that the SAGE II record is stable relative to the ground measurements, at least within 2 % decade^{-1}.

AMTD
8, 6661–6757, 2015

4.3.2 SAGE III

SAGE III collected data for only 3.5 years, which excludes the upper stratosphere from our study as no lidar sites provide sufficient statistics. Between 20–30 km the minimal drift uncertainty is 6 % decade^{-1}, while that of most stations is easily twice as high.

SAGE III ozone decreases relative to ground measurements, by −(2–10) % decade^{-1} in the middle stratosphere and somewhat more at lower altitudes (Fig. 5). The significance is by far insufficient however for a 2σ detection. The detection limit for the network-averaged drift is at best 6 % decade^{-1} between 20–30 km. In the lower stratosphere the threshold rapidly worsens to 10–20 % decade^{-1} due to the increased comparison noise from natural variability and instrumental noise. We therefore conclude that SAGE III is stable within about ±10 % decade^{-1}.

4.3.3 HALOE

The 14-year HALOE record allows a quite detailed study of the stability as well. The typical uncertainty at single stations is 5 % decade^{-1}, which is comparable to the variability of the spread between stations (Fig. 3, top left, light grey band). The 2σ detection threshold for the network average is 2–3 % decade^{-1} or more.

Above the 100 hPa level we observe a negative drift of about −(2–7) % decade^{-1} (Fig. 5). The result is significant between 10–50 hPa for both the ozonesonde and the lidar comparisons. Figure 3 demonstrates that negative drifts are found across the entire ground network (left panel), all centered around the network-averaged value (right). At lower and higher altitudes the drift is less than ±5 % decade^{-1} with an uncertainty of 2–7 % decade^{-1} and hence not significant. No coordinate dependence was found for the HALOE drift results (Fig. 9), so it can not be explained by drifting auxiliary data of the correlative records (Sect. 6).

Two earlier studies concluded that HALOE does not drift significantly relative to SAGE II, at least not more than ±10–15 % decade^{-1} (Morris et al., 2002; Nazaryan et al., 2005). Due to the longer data record considered here, the more frequent sam-

AMTD
8, 6661–6757, 2015

pling and the stability of the ground networks we obtain a significant result already at the 3 % decade^{-1} level between 10–50 hPa. Our result is consistent with the earlier reports, although a direct comparison is not straightforward due to the different timespan and vertical coordinate (we come back to this in Sect. 6). From Fig. 5 we infer that the middle stratospheric drift of HALOE relative to SAGE II must range between 0 and −5 % decade^{-1}, which is comparable in sign and in magnitude with the −(0-10) % decade^{-1} reported by Morris et al. (2002, Fig. 4a) and −(2-4) % decade^{-1} by Nazaryan et al. (2005, Fig. 8).

4.3.4 UARS MLS

The UARS MLS record is somewhat short. This limits the drift study, especially at low altitudes and relative to the lidar instruments. Between 5–50 hPa (20–35 km) the single station drift uncertainty is 5 % decade^{-1} at best, but typically twice as large. When the results are averaged over the ground network the 2σ detection threshold is 4–12 % decade^{-1}. At other altitudes the threshold increases rapidly, by a factor of at least two.

Below the 10 hPa level the ozonesonde comparisons show a positive drift of up to +5 % decade^{-1}, which is not significant (Fig. 5). At these altitudes the average drift relative to lidar is mostly negative on the other hand. The sonde and lidar results are not discrepant however, due to the large uncertainties for the lidar comparisons. In fact, it is difficult to conclude anything from the lidar results; the results at different sites tend to be somewhat discrepant, especially above the 10 hPa level. While the upper stratospheric drift of UARS MLS goes up to +10 % decade^{-1} relative to the OHP and Table Mountain lidars, it goes down to −10 % decade^{-1} relative to the Mauna Loa and Lauder lidars. The error bars are about 10 % decade^{-1}. We conclude that between 10–50 hPa the UARS MLS instrument is stable within about ±5–10 % decade^{-1}, perhaps slightly worse. In the upper stratosphere the discrepancy between the lidar results is too large to assess the stability. We report later on that the UARS MLS ozone drift results depend

AMTD
8, 6661–6757, 2015

on profile representation due to the ascending drift of the accompanying GPH profiles (Fig. 9). More details and mitigation follow in Sect. 6.

4.3.5 Aura MLS

The stability of the Aura MLS instrument can be studied in great detail, thanks to its excellent temporal and spatial sampling. Single site drift uncertainty is at best 0.6 % decade^{-1} and 2 % decade^{-1} on average. The regression uncertainties are substantially smaller than the observed standard deviation of the drifts over the network, which is about 4–6 % decade^{-1} above the 50–100 hPa level (Fig. 3, bottom). This leads to a considerable χ^2-adjustment (Eq. (4)) of $\kappa \simeq 2.5$ in the middle stratosphere and $\kappa \simeq 3$ in the upper stratosphere. The resulting 2σ detection limit for drift averages is 2–3 % decade^{-1} and it increases rapidly in the lower stratosphere.

In the upper and middle stratosphere the average drift is slightly positive, but not more than +(1.5–2) % decade^{-1} (Fig. 5). The ozonesonde and lidar derived results are also very consistent. Below the 100 hPa level a more important negative drift seems to develop, although the uncertainty is quite large. We therefore conclude that Aura MLS v3.3 is stable in the entire stratosphere, certainly within 1.5 % decade^{-1} (MS) and 2 % decade^{-1} (US). Due to an overall descending drift of the accompanying Aura MLS GPH profiles, there is a dependence of the Aura MLS ozone drift results on profile representation (Fig. 9). More details and mitigation follow in Sect. 6.

4.3.6 POAM II

The analysis of POAM II is extremely limited due to its low sampling and short record, merely three years. The regression requirement of at least ten data points was met at just seven polar ozonesonde stations. There were not enough collocations with lidar instruments to study the upper stratosphere. Drift uncertainty is about 30 % decade^{-1} at most sites and 20 % decade^{-1} in the best case. The resulting 2σ detection threshold for the network average is 25–40 % decade^{-1} in the middle stratosphere. This is much

AMTD
8, 6661–6757, 2015

larger than the observed drifts, which range from −15 % decade^{-1} at 20 km and 30 km to +15 % decade^{-1} at 25 km (Fig. 5). We therefore conclude that the stability of POAM II is better than ±(20–25) % decade^{-1} in the middle stratosphere.

4.3.7 POAM III

The POAM III data record spans 7.5 years and can therefore be studied in greater detail than that of its predecessor. In addition to the seven polar stations in the POAM II drift analysis, five ozonesonde sites at northern mid-latitudes provide a sufficiently sampled time series. Again, the regression was not feasible for lidar comparisons, limiting the altitude of our analysis to 30 km. The single station uncertainty is 4 % decade^{-1} at best and about 6 % decade^{-1} on average. When the results are averaged, the 2σ drift uncertainty becomes 4–8 % decade^{-1} in the middle stratosphere and rapidly grows to 10 % decade^{-1} at 15 km. Overall, POAM III tends to drift to lower ozone values between 20–30 km, at a rate of −(0-10) % decade^{-1} (Fig. 5). At lower altitudes the drift changes sign. None of these results are significant, which leads to the conclusion that POAM III is stable within about ±(10-15) % decade^{-1} in the middle and lower stratosphere.

4.3.8 OSIRIS

The OSIRIS time series are densely sampled at many ground stations. In the middle and upper stratosphere the minimal drift uncertainty is 1.3 % decade^{-1} and typically amounts to 3–4 % decade^{-1}. The regression errors do not fully explain the observed variability of 5–6 % decade^{-1} between stations above 20 km. The corresponding χ^2-adjustment factor increases to $\kappa \simeq 2$ above 30 km. The 2σ detection limit for the network average becomes 3 % decade^{-1} at 15 km, 1.6 % decade^{-1} between 20–30 km and 5 % decade^{-1} at 45 km.

In the lower stratosphere the OSIRIS drift relative to the correlative measurements is negative, at most −5 % decade^{-1} and not significant (Fig. 5). Around 20–25 km the drift turns positive and remains about +(2–3) % decade^{-1} up to 35 km, close to the

AMTD
8, 6661–6757, 2015

2σ threshold. Ozonesonde and lidar results provide a very consistent picture. Above 37 km, the drift becomes more pronounced and very significant. Its presence is easily visible in the comparison time series, e.g. at the Observatoire de Haute-Provence lidar (Fig. 2). At three of the four lidar stations we find a $>2\sigma$ drift of up to +8 % decade^{-1} at 42 km (Fig. 4, blue curves). This is consistent with findings by Adams et al. (2014) who reported a +(3–6) % decade^{-1} drift of OSIRIS relative to Aura MLS (which drifts slightly to higher ozone relative to ground-based measurements) around these altitudes, depending on how the Aura MLS data (pressure-VMR) are converted to the native OSIRIS system (altitude-number density). In summary, OSIRIS ozone drifts very likely to higher values above 25 km. The drift is quite small up to 35 km and barely significant. In the upper stratosphere the presence of a +(5–8) % decade^{-1} drift is evident.

4.3.9 SMR

Even though the SMR record spans 12 years and has good sampling properties, the ability to assess its stability is limited by the noise of the profiles. In Sect. 5 we show that the single SMR profile noise exceeds 20 % in the Tropics and 30 % at higher latitudes, which is substantially larger than for any other satellite record in this study. As a result, the drift uncertainty is at best 5–6 % decade^{-1} and typically $\sim$ 10 % decade^{-1} at individual ground sites. The regression errors cover the observed drift variability across the ground network, so the χ^2-adjustment is close to one. In the end, the 2σ threshold to detect averaged drifts ranges from 3–10 % decade^{-1} between 25–40 km.

The SMR profile drifts slightly to higher values in the middle stratosphere, although by no more than +5 % decade^{-1} which is insignificant (Fig. 5). Above 30 km the drift changes sign and increases rapidly in magnitude, topping out at −12 % decade^{-1} and more above 40 km. The drift in this altitude region has highest significance, crossing the detection threshold above 40 km. We therefore conclude that SMR is stable within about ±(5–10) % decade^{-1} over most of the stratosphere. Trend results in the uppermost stratosphere, however, should be interpreted cautiously.

AMTD
8, 6661–6757, 2015

4.3.10 GOMOS

The constraints on the stability of GOMOS are weaker than for its contemporary limb sounders, due to sparser sampling and, below ∼20 km, the larger noise of GOMOS. This is, for instance, clear from the comparison time series at the Payerne ozonesonde station (Fig. 2). In the middle stratosphere, the drift variability between stations is about 10 % decade^{-1}, which is larger than the error bars at individual sites, about 3 % decade^{-1} at best and 7 % decade^{-1} in general. The χ^2-adjustment increases the uncertainty on the network averages by $\kappa \simeq 1.5$. The resulting 2σ detection threshold is 4–6 % decade^{-1} above 20 km and ∼20 % decade^{-1} at 15 km.

Below 25–30 km a pronounced negative drift develops with decreasing altitude, from −2 % decade^{-1} at 30 km to −5 % decade^{-1} at 20 km (Fig. 5). The drift is significant between 18–25 km, and the ozonesonde and lidar-derived estimates are consistent. At lower altitudes, the negative drift is not significant due to marked increased noise. These observations are qualitatively consistent with Tegtmeier et al. (2013, Fig. 8). In the upper stratosphere the lidar results are scattered, but point on average to a positive drift above 35 km (Fig. 4, green line). The maximal drift is +3 % decade^{-1} at 45 km, which remains much below the 2σ threshold.

4.3.11 MIPAS

The MIPAS FR data (2002–2004) are not considered here, due to an altitude-dependent bias versus the OR data taken during 2005–2012. Nevertheless, the loss of a quarter of the MIPAS record does not affect the sensitivity of our analysis too much, since the sampling is sufficiently high. In the middle and upper stratosphere, the smallest single-site regression uncertainty is 1.5 % decade^{-1} and typically ∼3 % decade^{-1}. These errors do not fully cover the observed variability between sonde stations. They are therefore scaled by a factor of $\kappa \simeq 2$ between 20–50 hPa and $\simeq$1 below 100 hPa. The resulting 2σ detection limit for the network average is 3–4 % decade^{-1} between 10–100 hPa and 4–8 % decade^{-1} in the upper stratosphere.

AMTD
8, 6661–6757, 2015

The stability of the MIPAS OR record relative to ground measurements is better than ±2 % decade^{-1} in the middle and upper stratosphere, and better than ±5 % decade^{-1} at lower altitudes (Fig. 5). None of the results are significant and we therefore conclude that the MIPAS OR record is stable up to 1 hPa. If the user wants to include FR data as well for trend analyses, the additional (altitude-dependent) drift can be avoided as long as the FR-OR bias is accommodated for as a free parameter in the regression analysis (Eckert et al., 2014).

4.3.12 SCIAMACHY

The excellent sampling of SCIAMACHY allows us to probe its stability down to 0.8 % decade^{-1} at some ground sites, and on average down to ∼ 2 % decade^{-1}. Again, these statistical uncertainties do not cover the variability of 6 % decade^{-1} observed between the stations. The χ^2-adjustment factor is therefore higher than for other instruments, $\kappa \simeq 2$ over most of the middle stratosphere. The drift averages become significant when they cross the 3–6 % decade^{-1} bar in the middle and upper stratosphere.

Below 30 km the SCIAMACHY profiles drift to higher values relative to ozonesondes and lidars (Fig. 5). The drift is nearly independent of altitude and amounts to about +(2-3) % decade^{-1}, right at the detection threshold. The sign changes around 30 km, above which a highly significant negative drift develops relative to all lidar stations (Fig. 4, purple line). The drift reaches maximal significance, more than 4σ, around 38 km with a magnitude of −8 % decade^{-1}. Figure 2 illustrates the negative drift at 38 km in the comparison time series for the Mauna Loa lidar station (19° N, 156° W).

The large drift of SCIAMACHY ozone explains, at least partially, the more negative trends derived from the IUP Bremen v2.5 data set than those found from other records like Aura MLS and OSIRIS (Gebhardt et al., 2014). It should be noted that the latter study considers a different Level-2 processor than here, but there have been reports of a negative drift at 30–40 km of −5 % decade^{-1} for the IUP Bremen processor as well (Tegtmeier et al., 2013; Lambert et al., 2014). Since both processors digest the

AMTD
8, 6661–6757, 2015

same Level-1 data, it is plausible that the degradation of the instrument is not entirely corrected for by the Level-1 calibration scheme (Krijger et al., 2014).

4.3.13 ACE-FTS

The solar occultation instruments aboard SCISAT sample mainly high latitudes where the analysis lacks a lidar site. Our study is therefore limited to the lower and middle stratosphere. The best single-site drift uncertainty is 3 % decade^{-1}, whereas it amounts to about 10 % decade^{-1} in general, close to the observed variability between stations. The observed drifts remain within ±5 % decade^{-1}, consistent with the no-drift hypothesis (Fig. 5). The ACE-FTS data record can therefore be considered stable to within at least ±8 % decade^{-1}. A more precise analysis down to the ±5 % decade^{-1} level should be possible when the ACE-FTS profiles taken after September 2010 are included in the analysis.

4.3.14 MAESTRO

The uncertainty on the stability of the MAESTRO record is slightly poorer than that of ACE-FTS. The larger single-station uncertainties, at least 5 % decade^{-1} and typically 12–14 % decade^{-1}, lead to a 2σ detection at a level of 8–14 % decade^{-1} and 8–40 % decade^{-1} in the middle and lower stratosphere. The results never cross these thresholds: below 20 km we find a drift between −5 and +10 % decade^{-1}, above 20 km the drift is mainly positive around +4 % decade^{-1} (Fig. 5). So the MAESTRO record is stable within ±10–15 % decade^{-1}. Again, like for ACE-FTS, the uncertainty will decrease once the post-September 2010 profiles can be added to the analysis.

5 Bias and short-term variability

After studying decadal stability, we address the overall bias and short-term variability and search for patterns in altitude, latitude and season. As in the previous section,

AMTD
8, 6661–6757, 2015

we focus here on the individual satellite records in their native profile representation. Later on we expand the discussion to the consistency between profile representations (Sect. 6) and between satellite records (Sect. 7).

5.1 Methodology

Again, robust statistics are adopted that protect against outliers. We define the bias $b(l)$ as the median of the difference distribution at grid level l

$$b(l) = \underset{i}{\mathrm{q}_{50}}(\Delta x_i(l)), \tag{5}$$

where i runs over the pairs in the comparison sample. The 68 % interpercentile of the difference distribution

$$s(l) = \frac{1}{2}\left[\underset{i}{\mathrm{q}_{84}}(\Delta x_i(l)) - \underset{i}{\mathrm{q}_{16}}(\Delta x_i(l))\right] \tag{6}$$

is referred to as comparison spread s. We stress that s should not be confused with an estimate of the precision of the satellite data, as other, non-negligible terms enter the comparison error budget. These include the precision of the ground-based data and random metrological uncertainties related to the difference in sampled air masses (Sect. 3). In principle a similar remark is valid for the bias b as well, but systematic metrological uncertainties are expected to play a smaller role, except perhaps in the UTLS.

5.2 Results

The vertical and meridional structure of bias and comparison spread, relative to ozonesonde measurements, is shown in Figs. 6 and 7. Since there is more resemblance between the instruments, we only showed a few typical cases for the comparison spread. Table 5 summarises the sonde and lidar-derived bias estimates in four layers of the atmosphere.

AMTD
8, 6661–6757, 2015

5.2.1 SAGE II

Between 20–40 km SAGE II ozone remains mostly within ±3 % of the correlative measurements. Above 30–35 km, however, the sunrise profiles have a ∼ 5 % more negative bias relative to lidars than the sunset profiles (not shown here). Hence, SAGE II sunrise ozone is lower than sunset ozone in the middle and upper stratosphere, confirming earlier studies (Randall et al., 2003; Kyrölä et al., 2013). In the lowermost stratosphere, and below, ozone is underestimated by up to −(10–15) %. The spread in the comparisons is lowest between 25 and 40 km and shows poleward increases, 5 % at the Equator and ∼ 10 % at the high latitudes. Below 20 km the observed spread increases rapidly to 20–30 %, and especially under Antarctic ozone hole conditions.

5.2.2 SAGE III

The stratospheric bias of SAGE III is mostly less than ±3 %, like its predecessor. The bias is negative below ∼ 15 km and in the Arctic between 10–35 km. Elsewhere, ozone is slightly overestimated. The difference in bias relative to northern and southern lidars is small, so any sunrise (30–50° S) versus sunset (50–80° N) bias can not be large above 30 km. The short-term variability seems a few percent better than that of SAGE II, about 5 % at mid-latitudes and 8 % in the Arctic. Below 20 km and above 35/40 km the variability in the comparisons increases markedly.

5.2.3 HALOE

In the upper stratosphere and tropical middle stratosphere HALOE overestimates ozone by up to +3 %. In contrast, a negative bias is noted over the rest of the atmosphere. In the middle stratosphere it is not more than −5 %, and decreases rapidly below the 50 hPa level, reaching at least −25 % at 200 hPa. The variability in the comparisons is similar to that from the SAGE instruments, ranging from 5–10 % in the middle and upper stratosphere, and peaking at 30–40 % around the tropopause. Dur-

AMTD
8, 6661–6757, 2015

ing the Antarctic ozone hole season, the volume mixing ratios are overestimated by +25 % and the spread increases to 35 %.

5.2.4 UARS MLS

Our findings corroborate most of those by Livesey et al. (2003): (a) above 50 hPa UARS MLS slightly overestimates ozone by up to +5 %, (b) the 68 hPa and 100 hPa levels exhibit larger biases up to +10 %, and (c) the bias peaks at 68 hPa. However, negative biases up to −5 % are seen between 10–50 hPa relative to southern ozonesondes and above 5 hPa relative to northern lidars. The short-term variability is similar to the previous records, but reaches the 5–10 % range somewhat higher up in the middle stratosphere, around 20 hPa. At lower altitudes the comparison spread increases fast, topping at more than 40–50 % at the tropopause. We noted furthermore that the UARS MLS bias depends on the profile representation if one uses the GPH and temperature data included in the MLS product to perform conversions. This will be discussed in more detail in Sect. 6.

5.2.5 Aura MLS

Aura MLS ozone remains within ±3 % of correlative measurements between 5–50 hPa, except in the Arctic where a clear negative bias of −5 % is noted. The most striking bias characteristics are the stationary vertical oscillations found in the finer vertical retrieval grid results (version 3.3/3.4 data, see Livesey et al. (2013a)). They are very pronounced in the tropical UTLS where the amplitude reaches 10–15 %, but extend to higher latitudes and altitudes as well, with amplitudes of 3–5 %. The previous data release, v2.2, has a coarser grid in the UTLS and displays fewer oscillations. The recent new release of Aura MLS data (version 4.2) mitigates these oscillations to some extent (Livesey et al., 2015). The comparison spread shows that the single-profile noise is better than 4–7 % in the middle and upper stratosphere, and starts to degrade below the 50 hPa level. The Aura MLS bias furthermore depends on the profile representa-

AMTD
8, 6661–6757, 2015

tion if one uses the GPH and temperature data included in the MLS product to perform conversions. We discuss this in Sect. 6.

5.2.6 POAM II

We observe a negative bias of about 5–10 % between 20–30 km, which becomes rapidly more pronounced in the Antarctic lower stratosphere. In the northern lower stratosphere, however, the bias remains between 0 and −5 %. Above 30 km, there is a positive bias of 5–10 % or more relative to the polar lidars. The small lidar comparison sample did not allow us to study sunrise versus sunset results. The comparison spread is 5–10 % in the middle and upper stratosphere, and increases below 20 km.

5.2.7 POAM III

The POAM III bias is less than 5 % in MS and US, and has a negative sign between 18–30 km and positive elsewhere. Here, the spread in the comparisons is also similar to its predecessor, ranging between 5 and 10 %. In the LS the overestimation becomes at least +10 %, and, again, the spread is more pronounced. The small lidar comparison sample did not allow us to confirm the 5 % sunrise/sunset bias reported by Randall et al. (2003).

5.2.8 OSIRIS

Our bias results are very consistent with those of a comprehensive comparison of OSIRIS to GOMOS and Aura MLS (Adams et al., 2013). OSIRIS ozone mostly remains within ±4 % from correlative measurements above 20 km. Two features stand out. First and foremost, a marked peak in bias around 22 km is observed at all latitudes, which is possibly related to biases in the aerosol retrieval preceding the ozone retrieval (Adams et al., 2013). The comparison to southern lidars shows a second jump towards a persistent +5 % high bias, occurring between 30–35 km. Such a feature is not seen in the Northern Hemisphere. In the lower stratosphere, below 20 km, OSIRIS

AMTD
8, 6661–6757, 2015

underestimates ozone by 5–10 % at mid and high latitudes and by more than 20 % in the Tropics. Comparison spreads range from 6 to 11 % between 20 and 35/40 km. In the UTLS these increase to 20–40 % at 15 km, depending on latitude.

5.2.9 SMR

Our analysis confirms earlier reports of a systematic underestimation by 5 to 10 % relative to sonde and lidars in the upper and (most of the) middle stratosphere. The bias changes sign at lower altitudes and peaks at +5 to +10 % around 20 km. The most notable characteristics is the high comparison spread. It ranges between 20 % to 30 % from the tropical to the polar middle stratosphere, and is larger at other altitudes (Fig. 7). The large single-profile variability is caused by the weak 501.8 GHz ozone line used for the retrievals. Another SMR ozone profile data set, retrieved from the stronger 544.6 Ghz line, is less noisy, but this comes at the cost of a more important negative bias.

5.2.10 GOMOS

The GOMOS ozone bias above 20 km is generally less than ±3 %. The exception is the Arctic where a −7 % bias is found relative to ozonesonde and lidar at 25 km and again at 40 km. Another notable feature is that the sign of the bias in the extratropical UTLS is opposite in both hemispheres. It reaches −20 % in the North and +20 % in the South at 10 km. In the middle and upper stratosphere the comparison spread ranges from 7 % in the Tropics to 11 % at high latitudes (Fig. 7). Below 20–25 km GOMOS data becomes notably more noisy, at 15 km the observed spreads amount to 25–50 %, as a result of the increasing opaqueness of the atmosphere.

5.2.11 MIPAS

MIPAS OR data overestimate ozone relative to sondes and lidars over most of the stratosphere, except in the Arctic. At mid and low latitudes there are two bias peaks of

AMTD
8, 6661–6757, 2015

+(5–10) % around 50 hPa and 5 hPa. At other altitudes the bias remains below 5 %. At the bottom of the profile, below 200 hPa, ozone is overestimated by at least 20 %. In the Tropics the bias briefly flips sign between 50–200 hPa, here a very negative bias is found. From 20–40 km, the observed spread ranges from 4 % in the Tropics to 8 % at higher latitudes. Again, in the UTLS a sharp increase is observed (Fig. 7). We also noted a dependence of the MIPAS bias on the profile representation, if one uses the pressure and temperature data included in the MIPAS product to perform conversions. More details follow in Sect. 6.

5.2.12 SCIAMACHY

The SCIAMACHY bias is considerably positive over most of the atmosphere and manifests a very rich structure in altitude, latitude and season. The agreement with ozonesonde and lidar is better than 10 %, and best at mid-latitudes (between 0 and +5 % over 15–40 km). However, the bias easily reaches +15 % over a large part of the stratosphere, stretching from 30° N–60° S (> 25 km) to 60° S–90° S (> 35 km). The Arctic data are peculiar as well. There is a clear vertical dependence of the bias, peaking at +10 % around 20 km and at −10 % at 15 and 30 km. Also, both bias and comparison spread vary strongly with season. The bias at 20 km reaches a maximum of +25 % during local winter and a minimum of 0 % in summer. Similarly, the mean comparison spread is about 20 %, but it peaks at 30 % in winter and shrinks to 10 % in summer. At other latitudes the observed spread is never below about 10 %. Furthermore, in Sect. 6 we will show that the SCIAMACHY bias depends on the profile representation, if one uses the pressure and temperature data included in the SCIAMACHY product to perform conversions.

5.2.13 ACE-FTS

ACE-FTS ozone remains generally within about ±3 % from ground-based measurements over the entire stratosphere. The bias is only negative relative to Arctic ozone

AMTD
8, 6661–6757, 2015

sondes, everywhere else the ozone mixing ratios are overestimated. Above 30 km, the comparisons to mid-northern and high-southern lidars indicate a slightly larger positive bias, but not more than +5 %. The relative bias exceeds the 5 % very low in the atmosphere, just around the tropopause. The record performs also well in terms of short-term variability. The comparison spread is at most 7 % (10 %) at high latitudes in the middle (upper) stratosphere. It increases strongly below 20 km.

5.2.14 MAESTRO

The MAESTRO data set is typically biased low in the Northern Hemisphere (3 to 5 %) and high in the Southern Hemisphere (0 to 10 %). Ozone is clearly underestimated below 15 km, by at least 20 % at all latitudes. The observed comparison spreads range from 7 % at mid-latitudes to 10 % in the Arctic.

6 Impact of auxiliary data on non-native representations

So far, we have considered the quality of the satellite records in their native profile representation. But a user may actually desire another representation depending on his/her application (e.g. model comparisons, merging or assimilation of different records). In this case coincident altitude, pressure and/or temperature profile data are necessary for the conversion between ozone VMR and number density or between altitude and pressure. Users may prefer measurements, climatologies or reanalysis fields, all of which bring along uncertainties (e.g. Thorne et al., 2011; Seidel et al., 2011; Stauffer et al., 2014; Simmons et al., 2014). These ultimately add uncertainty $\mathbf{S}_{\text{auxiliary}}$ to the transformed ozone profile, which may have structure in space (altitude, latitude) and time (season, decades). Moreover, the currently observed negative trend of 1 K decade^{-1} in upper stratospheric temperature data already leads to representation-dependent differences of up to 1 % decade^{-1} in the ozone trends (McLinden and Fioletov, 2011). It is important to realise that a drift in temperature data will, in a simi-

AMTD
8, 6661–6757, 2015

lar fashion, introduce extra (altitude-dependent) drift in non-native ozone representations. Here, we consider the auxiliary data provided in the satellite data files, see Table 3. The auxiliary profiles for ground-based data are taken from actual measurements (ozonesonde: coupled radiosonde) and reanalysis fields (lidar: ERA-Interim).

Below about 35 km (∼ 5 hPa) there is generally no clear change in bias or comparison spread (both < 0.5–1 %) and drift (< 1 % decade^{-1}) after the conversion to another representation. There is hence no considerable difference in bias, short-term variability or long-term stability of the auxiliary data for most satellite and ground-based profiles. Examples are shown for SAGE II bias (Fig. 8) and HALOE and OSIRIS drift (Fig. 9). Temperature records are generally less consistent in the upper stratosphere (Simmons et al., 2014), so it is not surprising to observe considerable changes in ozone bias (up to ∼ 5 %) or drift (up to ∼ 5 % decade^{-1}) around 45 km (∼ 1 hPa).

For a few records we find clear indications that the accompanying auxiliary data have a more important impact. MIPAS bias changes by about 3 % when switching between VMR and number density, except between 25–30 km (∼ 10–20 hPa), see Fig. 8. The effect is slightly more pronounced in the Tropics and slightly less in the polar regions (not shown). It is not impossible that this dependence is actually a result of the conversion of the averaging kernels used to smooth the correlative profiles. Interestingly, transforming the vertical coordinate does not influence the MIPAS bias. The SCIAMACHY bias depends on both vertical and ozone coordinate over the entire stratosphere, by 3–5 % (Fig. 8), likely as a result of P/T uncertainties in the McLinden climatology. Both MLS records on the other hand exhibit clear representation dependences of the drift (Fig. 9), and to a lesser extent of the bias as well (although the ozonesonde and lidar results are somewhat discrepant, Fig. 8). The Aura MLS drift changes by about 3 % decade^{-1}, similar to earlier reports (Adams et al., 2014). The dependence has opposite sign for UARS MLS and is more pronounced, up to ∼ 10 % decade^{-1}. These observations are consistent with the known drifts in absolute pointing of the MLS records. Whereas UARS MLS geopotential height profiles drift upwards, by ∼ 1000 m decade^{-1} (Livesey et al., 2003), the Aura MLS v3.3/v2.2 GPH data drift downwards by ∼ 120 m decade^{-1},

AMTD
8, 6661–6757, 2015

especially between 2005–2009 (Livesey et al., 2015). Obviously, if a more stable and less biased source of auxiliary data were used for the conversion, the reported issues for MIPAS, SCIAMACHY and the MLS records could be easily avoided. Our results suggest that radiosonde data, reanalysis fields by ERA-interim (lidar) and MERRA (SAGE II), and ECMWF operational data (SMR, OSIRIS, GOMOS) allow for consistent conversions.

7 Consistency between satellite records

Until now we discussed each satellite Level-2 data set individually. Here, we take advantage of the specific design of the analysis to compare the satellite records directly. What follows is an evaluation of their mutual consistency in terms of bias (Fig. 10), short-term variability (Fig. 11), decadal stability (Fig. 12) and auxiliary data.

7.1 Bias

Figure 10 shows a superimposed view of the vertical structure of satellite bias in five latitude bands, in the native representation of each satellite record. The smallest biases and best mutual consistency are found between 20–40 km (~2–50 hPa). Here, satellite and ground-based measurements agree within 5 % or better (grey shaded area). Furthermore, the inter-satellite bias is not more than about 5 %. This illustrates the excellent consistency of all satellite and ground records in this part of the atmosphere. The consistency appears slightly poorer in the uppermost stratosphere (above 40 km/~2 hPa), perhaps due to the lower ozone abundances or due to larger systematic uncertainties in the lidar measurements. In the lower stratosphere and below the tropopause (grey horizontal line) there is a clear degradation of the percentage bias and consistency, due to declining ozone levels and increasing interference by clouds and aerosols (Wang et al., 1999, 2002; Randall et al., 2003). The bias relative to sondes easily reaches 15 % and more, and the inter-satellite biases can be more than

AMTD
8, 6661–6757, 2015

twice as large. Exceptions to this general picture are POAM II (dashed green) and especially SCIAMACHY (solid yellow). POAM II ozone is systematically low by about −5 to −10 % in the middle stratosphere, except in the Arctic. The SCIAMACHY bias reaches 10 % and more over a large part of the stratosphere, with a peculiar meridional structure, and a seasonal dependence that is very pronounced in the Arctic (not shown). Section 5 presented noteworthy bias features for other records as well, but of smaller magnitude and at smaller atmospheric scales: SMR (crosses the −5 % threshold above 35 km or ∼ 5 hPa), Aura MLS (distinctive vertical oscillations in the UTLS), MIPAS (persistent positive bias of ∼ 5 %), OSIRIS (sudden bump in bias around 22 km) and GOMOS (larger negative bias in Arctic).

Some bias features in Fig. 10 are common to the satellite measurements, and, hence, possibly related to the ground-based data quality. Perhaps the most striking, and not understood at the moment, is that the Arctic middle stratospheric bias is negative for most satellite records, relative to both sonde and lidar. This may indicate that the Arctic ground-based ozone values are too high. Secondly, there is a systematic positive upper stratospheric bias at tropical and southern mid-latitudes, possibly caused by a small negative bias of the dominating lidar record (Mauna Loa and Lauder). These ∼ 3 % biases remain within the systematic uncertainty due to uncertainties on the absorption cross sections used for the lidar retrievals. Thirdly, a ∼ 10–15 % negative bias seems present in the early Dumont d'Urville lidar record (1991–1998) (Godin et al., 2001). Indeed, all satellite records that started before 1999 (dashed lines) are biased high with similar magnitude in the Antarctic middle and upper stratosphere, while that is not the case for the more recent comparisons at this lidar site (2008–2013, solid). And finally, we systematically note a curved vertical structure of the bias relative to ozonesondes: sonde ozone values are decreasing by up to 5 % between 25 km and the top of the profile. This may be related to an incomplete sonde correction scheme for the reduced pump performance or the increased bias in pressure readings. Apart from these differences, the ozonesonde and lidar results are highly consistent, highlighting the suitability of these ground-based networks as a transfer standard.

AMTD
8, 6661–6757, 2015

7.2 Comparison spread

The comparison spread results in Fig. 11 are more straightforward than those of the bias. There is a consistent dependence on latitude and altitude for all records. Between 20–35/40 km the spread ranges between 5–12 % and increases slightly from the Tropics towards the poles, qualitatively consistent with a larger co-location mismatch uncertainty due to higher natural variability at high latitudes. Above 35–40 km the precision of the lidar measurements degrades, leading to a ∼ 5 % and more increase in comparison spread. Below 20 km the spread increases rapidly, easily more than 40 % at the tropopause, due to the higher natural variability. But here the lower signal to noise ratio (clouds, aerosols) plays a role as well and differences between records become obvious. GOMOS and UARS MLS appear less sensitive to ozone in the lower stratosphere. The most precise measurements over the entire stratosphere, on the other hand, are done by ACE-FTS, Aura MLS and MIPAS, although the comparison spread results for the latter two records may include a smaller co-location mismatch component due to the tighter time window (6 h instead of 12 h). SCIAMACHY and SMR are clearly different. The single-profile variability in the SMR comparisons is more elevated over the entire stratosphere (20–30 %). For SCIAMACHY this is seen in the upper stratosphere (10–15 %), and particularly in the Arctic (25–40 %) where a clear bump is discerned around 25 km, together with a very strong seasonal dependence (10 % in boreal summer, more than 30 % in winter). During the Antarctic ozone hole season, the extremely low ozone conditions blow the comparison spread of all records up to 40 % or more around 20 km (not shown here in detail). The low signal to noise ratios thus pose a real challenge for all limb/occultation sounders.

AMTD
8, 6661–6757, 2015

7.3 Decadal stability

Figure 12 presents a superimposed view of the vertical structure of the ground-network averaged decadal stability of all satellite records[2], in their native representation. The drift relative to ground observations is generally not significant and less than 5 % decade^{-1} in the middle and upper stratosphere, for some records even better than 3 % decade^{-1} over a large part of the stratosphere. The relative drift between satellite records can be twice as large however. A few records deviate from this general tendency. Either seemingly so because of large drift uncertainty (UARS MLS, SAGE III, POAM II), or because of the presence of a significant drift (HALOE between 20–35 km, SCIAMACHY between 30–45 km, OSIRIS between 35–45 km). The GOMOS (below 25 km) and SMR (above 35 km) records may drift as well, although the results are close to the detection threshold. Another peculiarity is the possible presence of a common, weak vertical dependence of the drifts in the middle stratosphere. These tend to become gradually more positive with increasing altitude, by 1–2 % decade^{-1} between 20–30 km, see Fig. 5 (e.g. SAGE II, Aura MLS, OSIRIS, GOMOS). This unexplained feature is observed independently of satellite record, or of type of correlative instrument, and deserves further study.

7.4 Impact of auxiliary data

Satellite ozone profile data quality is generally not affected by the conversion to another representation with the help of the accompanying pressure and temperature profiles. Bias, spread or decadal stability typically change by less than 1 % or 1 % decade^{-1} respectively in the lower and middle stratosphere, and somewhat more in the upper stratosphere. This demonstrates the good mutual consistency of the meteorological data by ozonesonde, MERRA and ERA-Interim. The exceptions are MIPAS and SCIA-

[2]To avoid clutter in Fig. 12 SAGE III and POAM II are not shown in the panels on the left. These can be seen in Fig. 5.

AMTD
8, 6661–6757, 2015

MACHY (3–5 % change in bias) and the Aura MLS and UARS MLS records (respectively 3 and 10 % decade^{-1} change in drift). Obviously, the introduction of these artificial effects can be avoided by using less biased or more stable sources of auxiliary data.

8 Discussion

The patterns in bias, short-term variability and decadal stability of the Level-2 ozone profile records identified in the preceding sections will affect higher-level products if not properly accounted for. Many studies within the community are based on gridded Level-3 data (e.g. monthly zonal means from single or a series of instruments) or assimilated Level-4 fields. In this section we discuss the relevance of our Level-2 assessment for the construction and analysis of these derived records, and focus in particular on implications for recent ozone profile trend assessments.

8.1 Can end-user requirements be verified?

We start the discussion by reflecting on the requirements of end users. Naturally, these depend on the envisaged application, so various sets of requirements have been drafted by the community[3]. We focus here on climate applications which rely on stable data sets spanning multiple decades on a global scale. The Global Climate Observing System (GCOS), for instance, requests an accuracy (Joint Committee for Guides in Metrology, 2012) better than 10 % in the UTLS and 5–20 % above, and a stability better than 1 % decade^{-1} (GCOS, 2011). Within ESA's Climate Change Initiative program (Ozone_cci) similar requirements were set for accuracy (< 8–15 %) and somewhat looser targets for stability (< 1–3 % decade^{-1}) (van der A et al., 2011).

In practice, the accuracy and stability of a particular record can of course only be tested to a level determined by the accuracy and stability of the reference data and by metrological constraints. From Figs. 10 and 11 we conclude that ground-based

[3]For an overview, see http://www.wmo-sat.info/oscar/variables/view/108.

AMTD
8, 6661–6757, 2015

studies are indeed able to verify an accuracy of 5–10 %, and resolve altitude-latitude-season patterns, in the middle and upper stratosphere. This is much more challenging in the UTLS, where metrological uncertainties become important due to increased natural variability and imperfect co-locations or differences in smoothing. Model data can help to reduce these, e.g. Verhoelst et al. (2015) showed recently that MACC (IFS-MOZART) and MERRA reanalysis fields allow them to close the error budget for total ozone column validation studies. However, further work is needed in the context of vertical profile validation.

It is even more challenging to verify the GCOS requirements for stability. Figure 12 (right panels) shows that the verification of the 1 % decade^{-1} targets with 95 % confidence is possible for just a few records (SAGE II, Aura MLS) and only in the middle stratosphere. In general, the analysis is not sensitive to network-averaged drifts below 2–3 % decade^{-1} in the middle stratosphere. In the upper and lower stratosphere, focal regions for current trend studies, the 2σ uncertainty on the drift is 3–4 % decade^{-1} or worse. In addition, a ground-based assessment of the meridional structure of satellite drift is currently infeasible. This is due to a lack of stations in certain latitude bands and the considerable observed scatter in the single-station drift estimates. However, there is some room for improvement. The best sampled comparison time series yield 1σ drift uncertainties as low as 0.7 % decade^{-1} at individual sites. But the dominant contribution to the network-averaged drift uncertainty of some recent satellite records comes from the scatter in the drift estimates at individual sites (Fig. 3). More homogeneity across the network will surely be beneficial, and this is one of the aims of the Ozonesonde Data Quality Assessment initiative (O3S-DQA). New correction schemes are being developed for the few percent biases introduced by (station- and time-dependent) changes in instrumental and post-processing set-ups, which may, ultimately, lead to more homogeneous sonde time series in time and space (Smit et al., 2012). When successful, this may perhaps allow an exploration of meridional drift structure as well. Also longer time series will help, but not to the full extent of what is actually desired. And finally, with the help of current models part of the comparison spread

AMTD
8, 6661–6757, 2015

could be removed statistically, which should, at least in the UTLS, lead to reduced drift uncertainties. Nevertheless, we consider it improbable that in the next few years sufficient progress can be made to demonstrate that single satellite records are stable within 1 % decade^{-1} relative to ground-based network observations. At the moment, 2–3 % decade^{-1} seems a more realistic target.

8.2 Implications for merging schemes

Space-based instruments are rarely operational for much more than a decade. Various groups have therefore produced multi-decade data sets from a series of individual records. Those based on limb/occultation instruments include SAGE-GOMOS (Kyrölä et al., 2013; Penckwitt et al., 2015), SAGE-OSIRIS (Bourassa et al., 2014; Sioris et al., 2014), GOZCARDS (Froidevaux et al., 2015), SWOOSH (Davis et al., 2015) and Ozone_cci (Sofieva et al., 2013), all listed in Table 6. These Level-3 data are typically reported as monthly averaged ozone over 5–10° latitude bins. A recent intercomparison by Tummon et al. (2015) showed that the differences between the merged data sets are dominated by the differences between the underlying data sets and to a lesser extent by differences between the merging algorithms. This shows the importance of a detailed understanding of the consistency between the Level-2 records in order to understand the merged product. In addition, comprehensive intercomparison studies (such as Jones et al. (2009); Dupuy et al. (2009); Nair et al. (2012); Tegtmeier et al. (2013); Adams et al. (2014); Laeng et al. (2014) and this work), can guide the design of the merging algorithms so as to reduce the impact of unfavorable Level-2 characteristics.

Although it is well known that the bias correction scheme should be altitude-latitude dependent, further improvements could be made. The inclusion of a diurnal and seasonal component may be pertinent, as we found sunrise-sunset bias differences for a few solar occultation instruments and a pronounced seasonal dependence of the bias and short-term variability of e.g. Arctic SCIAMACHY data. We also reported that the single Level-2 profile noise of SMR and SCIAMACHY is considerably higher than that

AMTD
8, 6661–6757, 2015

of other records. Averaged profiles will be sufficiently precise over large bins (monthly, 5° latitude) since both instruments are dense samplers, but this may not be the case at finer spatio-temporal resolutions. Our assessment of stability furthermore demonstrates the potential of drift correction schemes, especially when HALOE, OSIRIS or SCIAMACHY data are involved (and likely GOMOS and SMR as well). Eckert et al. (2014) have recently explored this approach, by correcting MIPAS trends for a drift relative to Aura MLS. In practice, however, the uncertainty of the drift estimate between two satellite records is often larger than the actual drift, especially for a short overlap period, which makes it very challenging to obtain robust corrections. Finally, the impact of the auxiliary data should not be forgotten, since profile representation conversions are typically required. We observed considerable changes in bias (MIPAS, SCIAMACHY and UARS/Aura MLS to a lesser extent) and stability (UARS/Aura MLS) due to the auxiliary data provided along with the ozone data sets. The use of a common source of stable auxiliary profiles eliminates additional discrepancies between the contributing records. Our results suggest that ECMWF (operational and ERA-interim) and MERRA fields impact ozone trends in a consistent way over the entire stratosphere.

8.3 Are observed trend differences due to drift?

Recently, a number of regression analyses were carried out on gridded ozone profile data from a variety of limb and occultation instruments. A few studies considered single records (Eckert et al., 2014; Gebhardt et al., 2014), others a combination of two (Kyrölä et al., 2013; Laine et al., 2014; Bourassa et al., 2014; Sioris et al., 2014) or more data sets (Tummon et al., 2015; WMO, 2014; Harris et al., 2015). The resulting profile trends are generally in reasonable agreement, but notable differences are observed in some parts of the stratosphere. The scientific SCIAMACHY data set by IUP Bremen, for instance, suggests a 2004–2012 trend in the Tropics around 35 km that is 4–6 % decade^{-1} more negative than OSIRIS and Aura MLS data (Gebhardt et al., 2014). A combined SAGE-OSIRIS record, on the other hand, produces more positive post-1998 trends in the uppermost stratosphere, by 3–4 % decade^{-1} at mid northern

AMTD
8, 6661–6757, 2015

latitudes (Bourassa et al., 2014; Tummon et al., 2015; Harris et al., 2015). Two records that combine SAGE and GOMOS data lead to considerably more negative trends than other data sets in the lower stratosphere (Tummon et al., 2015; Harris et al., 2015).

Many of the ozone trend differences can not be explained by statistical uncertainty. Our ground-based assessment of decadal stability suggests that these may be interpreted, at least for the better part, in terms of instrumental drift. Indeed, we noted a $+8\,\%\,\text{decade}^{-1}$ drift above 40 km for OSIRIS and a $-8\,\%\,\text{decade}^{-1}$ drift for SCIAMACHY[4] around 35 km. Additionally, we found indications of a $-5\,\%\,\text{decade}^{-1}$ drift of GOMOS below 20 km. These quite successful interpretations of some recent ozone trend differences builds additional confidence in our single-instrument drift estimates, which could therefore be employed as 1σ systematic uncertainty for long-term trend results for the corresponding records.

No studies have been performed so far of the decadal stability of the merged data sets. Yet, there is a clear need for realistic drift estimates for such data sets as well (Harris et al., 2015). We therefore make a first attempt to provide these for the merged records used by the recent WMO and SI2N assessments (WMO, 2014; Harris et al., 2015). Table 6 presents drift estimates for three stratospheric layers and for two time periods typically differentiated in trend analyses. Before 1997 all merged records rely on SAGE II observations, which are stable to within $1–1.5\,\%\,\text{decade}^{-1}$ depending on altitude. Since GOZCARDS and SWOOSH include HALOE data, the drift is possibly somewhat larger in the middle stratosphere. Producing post-1998 estimates is a more intricate problem, due to the increasing number of contributing instruments, and due to the fact that none of these cover the entire period. We are inclined to a conservative approach, giving figures that should be considered upper limits to the actual drift. The SAGE-GOMOS record will be impacted by negative GOMOS drifts in the lower stratosphere. The SAGE-OSIRIS trends should be considered more uncertain in the upper stratosphere due to drifting OSIRIS data. Records that use Aura MLS as back-

[4]Our analysis is based on the operational SCIAMACHY SGP v5.02 data set. The scientific v2.9 data set by IUP Bremen drifts by $-5\,\%\,\text{decade}^{-1}$ between 30–35 km (Lambert et al., 2014)

AMTD
8, 6661–6757, 2015

bone (GOZCARDS, SWOOSH) should not be more unstable than about 2 % decade^{-1} in the stratosphere. A merged Ozone_cci data set is likely also prone to larger uncertainty in the upper stratosphere (drifting OSIRIS, SCIAMACHY, SMR) and to some extent in the lower stratosphere as well (GOMOS). We stress that a more rigorous assessment is needed, since the estimates in Table 6 may well overestimate the actual drift. This work is currently on-going, following an approach similar to that by Frith et al. (2014).

9 Conclusions

Ground-based network observations by ozonesonde and stratospheric lidar instruments allowed us to assess the quality of fourteen records of the vertical distribution of ozone, collected by limb and occultation instruments over the past three decades. We considered three aspects of satellite data quality: the stability at decadal time scale (or drift), the overall bias, and the short-term variability. Further investigation of the vertical, meridional and seasonal structure of these parameters, together with their dependence on auxiliary data, revealed common and distinguishing features between satellite instruments. Such a comprehensive analysis serves two main objectives. First, to verify whether the spatio-temporal patterns of atmospheric ozone are correctly reproduced by the individual instruments at different scales. Second, to assess the consistency between satellite records, which is vital for their synergistic exploitation, a topic that has received increased interest in recent years.

We start our concluding remarks by distilling the general tendencies, saving some prominent exceptions for the folowing paragraph. Typically, we found a satellite-based bias better than ±5 % between 20–40 km (∼ 2–50 hPa), increasing slowly towards the stratopause (±10 %) and quite rapidly towards the tropopause (±15 % and more). A similar vertical dependence was observed for the comparison spread. It generally ranges from 5 to 12 % between 20–40 km and increases towards the stratopause (15–20 %) and tropopause (40 % and more). The precision of the records is actually

AMTD
8, 6661–6757, 2015

better than suggested by the observed spread in the comparisons, since the latter also includes the precision of the ground-based record and, especially in the UTLS, the random uncertainties due to differences in co-location and horizontal smoothing. Nevertheless, the altitude at which the quality of UTLS observations starts to degrade rapidly is clearly not only determined by the tropopause. It depends on the measurement technique and instrument as well (e.g. UTLS observations of UV-visible star occultations being less sensitive than those of infrared emissions at the limb). There were furthermore no evident signs of seasonal patterns, except for the Arctic SCIAMACHY data which exhibit a 10 % increase (decrease) in bias and spread in boreal winter (summer). We found no significant drifts at decadal time scales, most records are stable within about ±5 % decade^{-1} in the middle and upper stratosphere, and, for some records, even within ±3 % decade^{-1} (SAGE II, Aura MLS and MIPAS). However, the drift uncertainty should not be neglected, as our analysis is typically not sensitive (at 2σ) to drifts smaller than 2–3 % decade^{-1} in the middle stratosphere and 3–4 % decade^{-1} at lower and higher altitudes. The auxiliary data that accompany the satellite ozone data sets are generally well suited for the conversion between ozone and vertical coordinates. Bias, spread and drift of ozone do not change by more than about 1 % or 1 % decade^{-1} due to quality features in the pressure and/or temperature information, and somewhat more in the uppermost stratosphere.

There are of course exceptions to previous general observations. We noted more important biases (∼ 10 %) over much of the stratosphere for POAM II and SCIAMACHY, the latter also exhibits a clear hemispheric asymmetry. Two records show markedly poorer single-profile precision: SMR (entire atmosphere) and SCIAMACHY (upper stratosphere and Arctic). And three records drift significantly: HALOE in the middle stratosphere (−5 % decade^{-1}), and, in the upper stratosphere, SCIAMACHY (−8 % decade^{-1}) and OSIRIS (+8 % decade^{-1}). There are also indications of a −5 % decade^{-1} or more negative drift in the lower stratosphere for GOMOS, and in the upper stratosphere for SMR. Further confirmation is needed however, in the meantime, we advise caution when using these two data sets at these altitudes. Finally, we

AMTD
8, 6661–6757, 2015

observed for a few records a considerable impact of the accompanying auxiliary data (e.g. GPH retrievals) on ozone quality in non-native profile representations. The ozone bias changes by 3–5 % for MIPAS and SCIAMACHY; both MLS records (UARS and Aura) show a coordinate dependence of the drift (by 3 % decade^{-1} or more) and perhaps of the overall bias as well. We stress that these representation-dependent quality issues are unrelated to the satellite ozone retrievals themselves, and can be avoided by using another, external source of auxiliary information for any necessary conversions.

Overall, the observing system of limb and occultation instruments produces ozone profiles that meet the ∼ 10–15 % accuracy requirements by climate users, most certainly over 20–40 km, and perhaps in the lower stratosphere as well. However, it remains unclear whether the current Level-2 records comply with the 1–3 % decade^{-1} target on decadal stability. The combination of different data sets has received widespread interest in recent years, but also poses several challenges. Our results show that the merging schemes should be sufficiently refined to temper additional artefacts in the Level-3 data sets. Even then, the characteristics of merged records remain mostly defined by those of their contributors (Tummon et al., 2015). Multi-instrument comparison studies are therefore prerequisite to establish observational evidence. Indeed, we could relate the most notable differences between recent ozone profile trend studies to instrumental drift (WMO, 2014; Harris et al., 2015). This led us to a conservative estimate of the decadal stability of several merged records, which, until more rigourous analyses are performed, provides essential information for the recent trend assessments by WMO and SI2N.

Covering most limb and occultation ozone profilers of the past three decades, the ground-based networks of sonde and lidar instruments, and all major data quality indicators, this assessment is arguably the most comprehensive ground-based analysis so far. While bias and short-term variability of satellite records are well documented in the literature, this is much less the case for their long-term stability, the impact of auxiliary data and their mutual consistency. We therefore believe that this work will contribute to an improved interpretation of observation-based studies of the long-term evolution

AMTD
8, 6661–6757, 2015

of ozone and its link to climate change. However, our results represent a snapshot of the current versions of the data sets. In the near future, improved (and for some instruments longer) ozone profile time series will be released by the satellite teams and by the ground-based observers. Their efforts may lead to more stable records, which, in turn, would increase the sensitivity to even smaller drifts. In addition, the inclusion of microwave radiometer measurements and model data should help to evaluate the stability in the mesosphere and improve current estimates in the UTLS, especially in the Tropics.

Acknowledgements. Part of this work was funded by ESA projects Multi-TASTE and VALID, by the EU under FP6 project GEOmon (FP6-2005-Global-4-036677), and by ESA's CCI Ozone project. D. Hubert, A. Keppens and T. Verhoelst acknowledge national funding from the Belgian Science Policy Office (BELSPO) and ProDEx projects SECPEA and A3C. K. Stebel acknowledges funding from the ESA/ProDEx projects Ex Val (C90190,-CN1-4, 2005-2011). A. van Gijsel and D. Swart acknowledge support from the Dutch Ministry of Infrastructure and Environment. Work performed at the Jet Propulsion Laboratory was done under contract with the National Aeronautics and Space Administration. We are also grateful to C. De Clercq, D. Pieroux and S. Vandenbussche for their valuable input. The ozonesonde and lidar data used in this publication were obtained as part of WMO's Global Atmosphere Watch (GAW) and two of its main contributors, namely, the Network for the Detection of Atmospheric Composition Change (NDACC) and Southern Hemisphere ADditional OZonesondes programme (SHADOZ). The authors acknowledge the meticulous and sustained work of the PIs and staff at ozonesonde and lidar stations to acquire and maintain long-term ozone data records of high quality. The data records are publicly available via the NDACC Data Host Facility (http://www.ndacc.org), the SHADOZ archive (http://croc.gsfc.nasa.gov/shadoz) and the World Ozone and Ultraviolet Data Centre (http://www.woudc.org). NDACC and SHADOZ are supported by meteorological services and space agencies from many countries, with archives funded by NASA and NOAA. The authors also thank the satellite science and processing teams and the contributing space agencies. The Atmospheric Chemistry Experiment (ACE), also known as SCISAT, is a Canadian-led mission mainly supported by the Canadian Space Agency and the Natural Sciences and Engineering Research Council of Canada. SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY (SCIAMACHY) is a joint contribution of Germany, The Netherlands and Belgium to ESA's environmental satellite Envisat and is funded by the German

AMTD
8, 6661–6757, 2015

(DLR) and Dutch (NIVR) space agencies with Belgian contribution via BIRA-IASB. Sweden's Odin satellite carries the atmospheric and astronomical missions OSIRIS and SMR, developed and funded jointly by the space agencies of Sweden, Canada, Finland and France. This work is dedicated to our much appreciated colleague J. Urban, who regrettably passed away.

References

Adams, C., Bourassa, A. E., Bathgate, A. F., McLinden, C. A., Lloyd, N. D., Roth, C. Z., Llewellyn, E. J., Zawodny, J. M., Flittner, D. E., Manney, G. L., Daffer, W. H., and Degenstein, D. A.: Characterization of Odin-OSIRIS ozone profiles with the SAGE II dataset, Atmos. Meas. Tech., 6, 1447–1459, doi:10.5194/amt-6-1447-2013, 2013. 6677, 6702

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, doi:10.5194/amt-7-49-2014, 2014. 6666, 6674, 6677, 6695, 6706, 6713

Barath, F. T., Chavez, M. C., Cofield, R. E., Flower, D. A., Frerking, M. A., Gram, M. B., Harris, W. M., Holden, J. R., Jarnot, R. F., Kloezeman, W. G., Klose, G. J., Lau, G. K., Loo, M. S., Maddison, B. J., Mattauch, R. J., McKlnney, R. P., Peckham, G. E., Pickett, H. M., Siebes, G., Soltis, F. S., Suttie, R. A., Tarsala, J. A., Waters, J. W., and Wilson, W. J.: The Upper Atmosphere Research Satellite Microwave Limb Sounder Instrument, J. Geophys. Res., 98, 10751–10762, doi:10.1029/93JD00798, 1993. 6673

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., De Mazière, 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, 1–5, doi:10.1029/2005GL022386, 2005. 6681

Bertaux, J. L., Kyrölä, E., Fussen, D., Hauchecorne, A., Dalaudier, F., Sofieva, V., Tamminen, J., Vanhellemont, F., Fanton d'Andon, O., Barrot, G., Mangin, A., Blanot, L., Lebrun, J. C., Pérot,

AMTD
8, 6661–6757, 2015

K., Fehr, T., Saavedra, L., Leppelmeier, G. W., and Fraisse, R.: Global ozone monitoring by occultation of stars: an overview of GOMOS measurements on ENVISAT, Atmos. Chem. Phys., 10, 12091–12148, doi:10.5194/acp-10-12091-2010, 2010. 6678

Boone, C. D., Walker, K. A., and Bernath, P. F.: The Atmospheric Chemistry Experiment ACE at 10: A Solar Occultation Anthology, chap. Version 3 Retrievals for the Atmospheric Chemistry Experiment Fourier Transform Spectrometer (ACE-FTS), 103–127, A. Deepak Publishing, Hampton, Virginia, USA, http://acebox2.uwaterloo.ca/publications/2013/Version3.5retreivals2013.pdf (last access: 23 June 2015), 2013. 6681

Bourassa, A. E., Degenstein, D. A., Randel, W. J., Zawodny, J. M., Kyrölä, E., McLinden, C. A., Sioris, C. E., and Roth, C. Z.: Trends in stratospheric ozone derived from merged SAGE II and Odin-OSIRIS satellite observations, Atmos. Chem. Phys., 14, 6983–6994, doi:10.5194/acp-14-6983-2014, 2014. 6713, 6714, 6715

Bovensmann, H., Burrows, J., Buchwitz, M., Frerick, J., Noël, S., Rozanov, V., Chance, K., and Goede, A.: SCIAMACHY: Mission objectives and measurement modes, J. Atmos. Sci., 56, 127–150, 1999. 6680

Brewer, A. W. and Milford, J. R.: The Oxford-Kew Ozone Sonde, P. R. Soc. London, 256, 470–495, doi:10.1098/rspa.1960.0120, 1960. 6668

Calisesi, Y., Soebijanta, V. T., and van Oss, R.: Regridding of remote soundings: Formulation and application to ozone profile comparison, J. Geophys. Res., 110, 1–8, doi:10.1029/2005JD006122, 2005. 6686

Ceccherini, S., Carli, B., and Raspollini, P.: Quality of MIPAS operational products, J. Quant. Spectrosc. Ra., 121, 45–55, doi:10.1016/j.jqsrt.2013.01.021, 2013. 6679

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, doi:10.5194/acp-7-4807-2007, 2007. 6685

Croux, C., Dhaene, G., and Hoorelbeke, D.: Robust standard errors for robust estimators, http://www.econ.kuleuven.ac.be/eng/ew/discussionpapers/Dps03/Dps0316.pdf (last access: 23 June 2015), Discussion Paper Series (DPS) 03.16, 2004. 6687

AMTD
8, 6661–6757, 2015

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, doi:10.5194/amt-6-3539-2013, 2013. 6671

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, doi:10.5194/acp-14-13455-2014, 2014. 6671

Danilin, M. Y., Ko, M. K. W., Froidevaux, L., Santee, M. L., Lyjak, L. V., Bevilacqua, R. M., Zawodny, J. M., Sasano, Y., Irie, H., Kondo, Y., Russell III, J. M., Scott, C. J., and Read, W. G.: Trajectory hunting as an effective technique to validate multiplatform measurements: Analysis of the MLS, HALOE, SAGE-II, ILAS, and POAM-II data in October-November 1996, J. Geophys. Res., 107, 4420, doi:10.1029/2001JD002012, 2002. 6675

Davis, S. M. et al.: The Stratospheric Water and Ozone Satellite Homogenized (SWOOSH) database: A long-term database for climate studies, Atmos. Meas. Tech. Disc., in preparation, 2015. 6713

De Clercq, C.: Implémentation du système global pour la mesure de la composition atmosphérique: concepts et méthodes pour une analyse intégrée des données satellitales, Ph.D. thesis, BIRA-IASB and Université Libre de Bruxelles, 2009. 6685

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C. an d Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. Roy. Meteor. Soc., 137, 553–597, doi:10.1002/qj.828, 2011. 6670

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. 6676

Deniel, C., Dalaudier, F., Chassefiere, E., Bevilacqua, R., Shettle, E., Hoppel, K., Hornstein, J., Lumpe, J., Rusch, D., and Randall, C.: A comparative study of POAM II and electrochemical concentration cell ozonesonde measurements obtained over northern Europe, J. Geophys. Res., 102, 23629–23642, doi:10.1029/97JD01665, 1997. 6675

AMTD
8, 6661–6757, 2015

Deshler, T., Mercer, J. L., Smit, H. G. J., Stübi, R., Levrat, G., Johnson, B. J., Oltmans, S. J., Kivi, R., Thompson, A. M., Witte, J., Davies, J., Schmidlin, F. J., Brothers, G., and Sasaki, T.: Atmospheric comparison of electrochemical cell ozonesondes from different manufacturers, and with different cathode solution strengths: The Balloon Experiment on Standards for Ozonesondes, J. Geophys. Res., 113, 1–17, doi:10.1029/2007JD008975, 2008. 6669

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., Godin-Beekmann, 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., T{'etard, 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. 6666, 6681, 6682, 6713, 6742

Eckert, E., von Clarmann, T., Kiefer, M., Stiller, G. P., Lossow, S., Glatthor, N., Degenstein, D. A., Froidevaux, L., Godin-Beekmann, S., Leblanc, T., McDermid, S., Pastel, M., Steinbrecht, W., Swart, D. P. J., Walker, K. A., and Bernath, P. F.: Drift-corrected trends and periodic variations in MIPAS IMK/IAA ozone measurements, Atmos. Chem. Phys., 14, 2571–2589, doi:10.5194/acp-14-2571-2014, 2014. 6666, 6674, 6680, 6681, 6697, 6714

Efron, B. and Tibshirani, R.: Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy, Statist. Sci., 1, 54–75, doi:10.1214/ss/1177013815, 1986. 6689

AMTD
8, 6661–6757, 2015

ESA: GOMOS Level 2 processor version GOMOS/6.01 Readme, Tech. rep. ENVI-GSOP-EOGD-QD-12-0117, https://earth.esa.int/documents/700255/708000/RMF_0117_GOM_NL__2P_Disclaimers.pdf (last access: 23 June 2015), 2012a. 6679, 6742

ESA: MIPAS Level 2 ML2PP Version 6 Readme, Tech. rep. ENVI-GSOP-EOGD-QD-12-0116, https://earth.esa.int/documents/700255/707722/MIP_NL_2P_README_V6.0.pdf (last access: 23 June 2015), 2012b. 6680, 6742

ESA: Readme file for SCIAMACHY Level 2 version 5.02 products, Tech. rep. ENVI-GSOP-EOGD-QD-13-0118, https://earth.esa.int/documents/700255/708683/SCIAMACHY+L2+Quality+Readme+file (last access: 23 June 2015), 2013. 6680, 6742

Fischer, H., Birk, M., Blom, C., Carli, B., Carlotti, M., von Clarmann, T., Delbouille, L., Dudhia, A., Ehhalt, D., Endemann, M., Flaud, J. M., Gessner, R., Kleinert, A., Koopman, R., Langen, J., López-Puertas, M., Mosner, P., Nett, H., Oelhaf, H., Perron, G., Remedios, J., Ridolfi, M., Stiller, G., and Zander, R.: MIPAS: an instrument for atmospheric and climate research, Atmos. Chem. Phys., 8, 2151–2188, doi:10.5194/acp-8-2151-2008, 2008. 6679

Frisk, U., Hagström, M., Ala-Laurinaho, J., Andersson, A., Berges, J.-C., Chabaud, J.-P., Dahlgren, M., Emrich, A., Florén, H.-G., Florin, G., Fredrixon, M., Gaier, T., Haas, R., Hirvonen, T., Hjalmarsson, Å., Jakobsson, B., Jukkala, P., Kildal, P. S., Kollberg, E., Lassing, J., Lecacheux, A., Lehikoinen, P., Lehto, A., Mallat, J., Marty, C., Michet, D., Narbonne, J., Nexon, M., Olberg, M., Olofsson, A. O. H., Olofsson, G., Origné, A., Petersson, M., Piironen, P., Pons, R., Pouliquen, D., Ristorcelli, I., Rosolen, C., Rouaix, G., Räisänen, A. V., Serra, G., Sjöberg, F., Stenmark, L., Torchinsky, S., Tuovinen, J., Ullberg, C., Vinterhav, E., Wadefalk, N., Zirath, H., Zimmermann, P., and Zimmermann, R.: The Odin satellite – I. Radiometer design and test, Astron. Astrophys., 402, L27–L34, doi:10.1051/0004-6361:20030335, 2003. 6677

Frith, S. M., Kramarova, N. A., Stolarski, R. S., McPeters, R. D., Bhartia, P. K., and Labow, G. J.: Recent changes in total column ozone based on the SBUV Version 8.6 Merged Ozone Data Set, J. Geophys. Res., 119, 9735–9751, doi:10.1002/2014JD021889, 2014. 6716

Froidevaux, L., Jiang, Y. B., Lambert, A., Livesey, N., Read, W., Waters, J., Browell, E., Hair, J., Avery, M., McGee, T., Twigg, L., Sumnicht, G., Jucks, K., Margitan, J., Sen, B., Stachnik, R., Toon, G., Bernath, P., Boone, C., Walker, K., Filipiak, M., Harwood, R., Fuller, R., Manney, G., Schwartz, M., Daffer, W., Drouin, B., Cofield, R., Cuddy, D., Jarnot, R., Knosp, B., Perun, V., Snyder, W., Stek, P., Thurstans, R., and Wagner, P.: Validation of Aura Mi-

AMTD
8, 6661–6757, 2015

crowave Limb Sounder stratospheric ozone measurements, J. Geophys. Res., 113, 1–24, doi:10.1029/2007JD008771, 2008. 6674

Froidevaux, L., Anderson, J., Wang, H.-J., Fuller, R. A., Schwartz, M. J., Santee, M. L., Livesey, N. J., Pumphrey, H. C., Bernath, P. F., Russell III, J. M., and McCormick, M. P.: Global OZone Chemistry And Related Datasets for the Stratosphere (GOZCARDS): methodology and sample results with a focus on HCl, H_2O, and O_3, Atmos. Chem. Phys. Disc., 15, 5849–5957, doi:10.5194/acpd-15-5849-2015, 2015. 6713

GCOS: Systematic Observation Requirements for Satellite-based Products for Climate, 2011 update, p. 127, http://www.wmo.int/pages/prog/gcos/Publications/gcos-154.pdf (last access: 23 June 2015), 2011. 6711

Gebhardt, C., Rozanov, A., Hommel, R., Weber, M., Bovensmann, H., Burrows, J. P., Degenstein, D., Froidevaux, L., and Thompson, A. M.: Stratospheric ozone trends and variability as seen by SCIAMACHY from 2002 to 2012, Atmos. Chem. Phys., 14, 831–846, doi:10.5194/acp-14-831-2014, 2014. 6697, 6714

Glaccum, W., Lucke, R. L., Bevilacqua, R. M., Shettle, E. P., Hornstein, J. S., Chen, D. T., Lumpe, J. D., Krigman, S. S., Debrestian, D. J., Fromm, M. D., Dalaudier, F., Chassefière, E., Deniel, C., Randall, C. E., Rusch, D. W., Olivero, J. J., Brogniez, C., Lenoble, J., and Kremer, R.: The Polar Ozone and Aerosol Measurement instrument, J. Geophys. Res., 101, 14479–14487, doi:10.1029/96JD00576, 1996. 6675

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. 6670

Godin, S., Bergeret, V., Bekki, S., David, C., and Mégie, G.: Study of the interannual ozone loss and the permeability of the Antarctic polar vortex from aerosol and ozone lidar measurements in Dumont d'Urville (66.4S, 140E), J. Geophys. Res. Atmos., 106, 1311–1330, doi:10.1029/2000JD900459, 2001. 6708

Godin-Beekmann, S., Porteneuve, J., and Garnier, A.: Systematic DIAL lidar monitoring of the stratospheric ozone vertical distribution at Observatoire de Haute-Provence (43.92N, 5.71E), J. Environ. Monit., 5, 57–67, doi:10.1039/B205880D, 2003. 6670

Harris, N. R. P., Hassler, B., Tummon, F., Bodeker, G. E., Hubert, D., Petropavlovskikh, I., Steinbrecht, W., Anderson, J., Bhartia, P. K., Boone, C. D., Bourassa, A., Davis, S. M., Degenstein, D., Delcloo, A., Frith, S. M., Froidevaux, L., Godin-Beekmann, S., Jones, N., Kurylo,

AMTD
8, 6661–6757, 2015

M. J., Kyrölä, E., Laine, M., Leblanc, S. T., Lambert, J.-C., Liley, B., Mahieu, E., Maycock, A., de Mazière, M., Parrish, A., Querel, R., Rosenlof, K. H., Roth, C., Sioris, C., Staehelin, J., Stolarski, R. S., Stübi, R., Tamminen, J., Vigouroux, C., Walker, K., Wang, H. J., Wild, J., and Zawodny, J. M.: Past changes in the vertical distribution of ozone, Part 3: Analysis and interpretation of trends, Atmos. Chem. Phys. Disc., 15, 8565–8608, doi:10.5194/acpd-15-8565-2015, 2015. 6665, 6666, 6714, 6715, 6718

Hassler, B., Petropavlovskikh, I., Staehelin, J., August, T., Bhartia, P. K., Clerbaux, C., Degenstein, D., Mazière, M. D., Dinelli, B. M., Dudhia, A., Dufour, G., Frith, S. M., Froidevaux, L., Godin-Beekmann, S., Granville, J., Harris, N. R. P., Hoppel, K., Hubert, D., Kasai, Y., Kurylo, M. J., Kyrölä, E., Lambert, J.-C., Levelt, P. F., McElroy, C. T., McPeters, R. D., Munro, R., Nakajima, H., Parrish, A., Raspollini, P., Remsberg, E. E., Rosenlof, K. H., Rozanov, A., Sano, T., Sasano, Y., Shiotani, M., Smit, H. G. J., Stiller, G., Tamminen, J., Tarasick, D. W., Urban, J., van der A, R. J., Veefkind, J. P., Vigouroux, C., von Clarmann, T., von Savigny, C., Walker, K. A., Weber, M., Wild, J., and Zawodny, J. M.: Past changes in the vertical distribution of ozone – Part 1: Measurement techniques, uncertainties and availability, Atmos. Meas. Tech., 7, 1395–1427, doi:10.5194/amt-7-1395-2014, 2014. 6665, 6668

Hervig, M. and McHugh, M.: Cirrus detection using HALOE measurements, Geophys. Res. Lett., 26, 719–722, doi:10.1029/1999GL900069, 1999. 6673, 6742

Inai, Y., Shiotani, M., Fujiwara, M., Hasebe, F., and Vömel, H.: Altitude misestimation caused by the Vaisala RS80 pressure bias and its impact on meteorological profiles, Atmos. Meas. Tech. Disc., 8, 2191–2222, doi:10.5194/amtd-8-2191-2015, 2015. 6686

Jégou, F., Urban, J., de La Noë, J., Ricaud, P., Le Flochmoën, E., Murtagh, D. P., Eriksson, P., Jones, A., Petelina, S., Llewellyn, E. J., Lloyd, N. D., Haley, C., Lumpe, J., Randall, C., Bevilacqua, R. M., Catoire, V., Huret, N., Berthet, G., Renard, J. B., Strong, K., Davies, J., Mc Elroy, C. T., Goutail, F., and Pommereau, J. P.: Technical Note: Validation of Odin/SMR limb observations of ozone, comparisons with OSIRIS, POAM III, ground-based and balloon-borne instruments, Atmos. Chem. Phys., 8, 3385–3409, doi:10.5194/acp-8-3385-2008, 2008. 6677

Jiang, Y. et al.: Validation of Aura Microwave Limb Sounder Ozone by ozonesonde and lidar measurements, J. Geophys. Res., 112, 1–20, doi:10.1029/2007JD008776, 2007. 6674

Joint Committee for Guides in Metrology: International Vocabulary of Metrology – Basic and General Concepts and Associated Terms, Tech. rep., http://www.bipm.org/utils/common/documents/jcgm/JCGM_200_2012.pdf (last access: 23 June 2015), 2012. 6711

AMTD
8, 6661–6757, 2015

Jones, A., Murtagh, D., Urban, J., Eriksson, P., and Rösevall, J.: Intercomparison of Odin/SMR ozone measurements with MIPAS and balloon sonde data, Can. J. Phys., 85, 1111–1123, doi:10.1139/p07-118, 2007. 6677

Jones, A., Urban, J., Murtagh, D. P., Eriksson, P., Brohede, S., Haley, C., Degenstein, D., Bourassa, A., von Savigny, C., Sonkaew, T., Rozanov, A., Bovensmann, H., and Burrows, J.: Evolution of stratospheric ozone and water vapour time series studied with satellite measurements, Atmos. Chem. Phys., 9, 6055–6075, doi:10.5194/acp-9-6055-2009, 2009. 6666, 6678, 6713, 6742

Kar, J., McElroy, C., Drummond, J., Zou, J., Nichitiu, F., Walker, K., Randall, C., Nowlan, C., Dufour, D., Boone, C., Bernath, P., Trepte, C., Thomason, L., and McLinden, C.: Initial comparison of ozone and NO_2 profiles from ACE-MAESTRO with balloon and satellite data, J. Geophys. Res., 112, 1–14, doi:10.1029/2006JD008242, 2007. 6682, 6742

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. 6670

Kobayashi, J. and Toyama, Y.: On various methods of measuring the vertical distribution of atmospheric ozone (III) – Carbon iodine type chemical ozonesonde, Pap. Met. Geophys., 17, 113–126, https://www.jstage.jst.go.jp/article/mripapers1950/17/2/17_97/_article/references, 1966. 6669

Komhyr, W. D.: Electrochemical concentration cells for gas analysis, Ann. Geophys., 25, 203–210, 1969, http://www.ann-geophys.net/25/203/1969/. 6668

Krijger, J. M., Snel, R., van Harten, G., Rietjens, J. H. H., and Aben, I.: Mirror contamination in space I: mirror modelling, Atmos. Meas. Tech., 7, 3387–3398, doi:10.5194/amt-7-3387-2014, 2014. 6698

Kyrölä, E., Tamminen, J., Sofieva, V., Bertaux, J. L., Hauchecorne, A., Dalaudier, F., Fussen, D., Vanhellemont, F., Fanton d'Andon, O., Barrot, G., Guirlet, M., Mangin, A., Blanot, L., Fehr, T., Saavedra de Miguel, L., and Fraisse, R.: Retrieval of atmospheric parameters from GOMOS data, Atmos. Chem. Phys., 10, 11881–11903, doi:10.5194/acp-10-11881-2010, 2010. 6678

AMTD
8, 6661–6757, 2015

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. 6671, 6700, 6713, 6714

Laeng, A., Grabowski, U., von Clarmann, T., Stiller, G., Glatthor, N., Höpfner, M., Kellmann, S., Kiefer, M., Linden, A., Lossow, S., Sofieva, V., Petropavlovskikh, I., Hubert, D., Bathgate, T., Bernath, P., Boone, C. D., Clerbaux, C., Coheur, P., Damadeo, R., Degenstein, D., Frith, S., Froidevaux, L., Gille, J., Hoppel, K., McHugh, M., Kasai, Y., Lumpe, J., Rahpoe, N., Toon, G., Sano, T., Suzuki, M., Tamminen, J., Urban, J., Walker, K., Weber, M., and Zawodny, J.: Validation of MIPAS IMK/IAA V5R_O3_224 ozone profiles, Atmos. Meas. Tech., 7, 3971–3987, doi:10.5194/amt-7-3971-2014, 2014. 6666, 6713

Laeng, A., Hubert, D., Verhoelst, T., von Clarmann, T., Dinelli, B., Dudhia, A., Raspollini, P., Stiller, G., Grabowski, U., Keppens, A., Kiefer, M., Sofieva, V., Froidevaux, L., Walker, K., Lambert, J.-C., and Zehner, C.: The ozone climate change initiative: Comparison of four Level-2 processors for the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS), Remote Sens. Environ., 162, 316–343, doi:10.1016/j.rse.2014.12.013, 2015. 6679

Laine, M., Latva-Pukkila, N., and Kyrölä, E.: Analysing time-varying trends in stratospheric ozone time series using the state space approach, Atmos. Chem. Phys., 14, 9707–9725, doi:10.5194/acp-14-9707-2014, 2014. 6714

Lambert, J.-C., Balis, D., Delcloo, A., Goutail, F., Granville, J., Hubert, D., Keppens, A., Kivi, R., Koukouli, ML., Pommereau, J.-P., Stübi, R., Verhoelst, T., Loyola, D., Siddans, R., van der A, R. J., Laeng, A., Sofieva, V., Weber, W., Van Roozendael, M., and Zehner, C.: Ozone_cci – Product Validation and Intercomparison Report (PVIR), Tech. rep., BIRA-IASB, http://www.esa-ozone-cci.org/?q=webfm_send/148 (last access: 23 June 2015), 2014. 6697, 6715

Lambert, J.-C. et al.: Past changes in the Vertical Distribution of Ozone, Part II: Measurement Intercomparisons, Atmos. Meas. Tech. Disc., in preparation, 2015. 6666

Lichtenberg, G.: SCIAMACHY Offline Processor Level1b-2 ATBD – Algorithm Theoretical Baseline Document (SGP OL Version 5), Tech. rep. ENV-ATB-QWG-SCIA-0085, DLR-IMF, 2011. 6680

Livesey, N. J., Read, W. G., Froidevaux, L., Waters, J. W., Santee, M. L., Pumphrey, H. C., Wu, D. L., Shippony, Z., and Jarnot, R. F.: The UARS Microwave Limb Sounder version 5 data set: Theory, characterization, and validation, J. Geophys. Res., 108, 4378, doi:10.1029/2002JD002273, 2003. 6673, 6674, 6701, 6706, 6742

AMTD
8, 6661–6757, 2015

Livesey, N. J., Read, W. G., Froidevaux, L., Lambert, A., Manney, G. L., Pumphrey, H. C., Santee, M. L., Schwartz, M. J., Wang, S., Cofield, R. E., Cuddy, D. T., Fuller, R. A., Jarnot, R. F., Jiang, J. H., Knosp, B. W., Stek, P. C., Wagner, P. A., and Wu., D. L.: Aura Microwave Limb Sounder (MLS). Version 3.3 Level 2 data quality and description document, Tech. rep., Jet Propulsion Laboratory, 2011. 6742

Livesey, N. J., Logan, J. A., Santee, M. L., Waters, J. W., Doherty, R. M., Read, W. G., Froidevaux, L., and Jiang, J. H.: Interrelated variations of O_3, CO and deep convection in the tropical/subtropical upper troposphere observed by the Aura Microwave Limb Sounder (MLS) during 2004-2011, Atmos. Chem. Phys., 13, 579–598, doi:10.5194/acp-13-579-2013, 2013a. 6674, 6701

Livesey, N. J., Read, W. G., Froidevaux, L., Lambert, A., Manney, G. L., Pumphrey, H. C., Santee, M. L., Schwartz, M. J., Wang, S., Cofield, R. E., Cuddy, D. T., Fuller, R. A., Jarnot, R. F., Jiang, J. H., Knosp, B. W., Stek, P. C., Wagner, P. A., and Wu., D. L.: Aura Microwave Limb Sounder (MLS). Version 3.3 and 3.4 Level 2 data quality and description document, Tech. rep., Jet Propulsion Laboratory, http://mls.jpl.nasa.gov/data/v3_data_quality_document.pdf (last access: 23 June 2015), 2013b. 6674, 6675

Livesey, N. J., Read, W. G., Wagner, P. A., Froidevaux, L., Lambert, A., Manney, G. L., Mill'an Valle, L. F., Pumphrey, H. C., Santee, M. L., Schwartz, M. J., Wang, S., Fuller, R. A., Jarnot, R. F., Knosp, B. W., and Martin, E.: Aura Microwave Limb Sounder (MLS). Version 4.2x Level 2 data quality and description document, Tech. rep., Jet Propulsion Laboratory, http://mls.jpl.nasa.gov/data/v4-2_data_quality_document.pdf (last access: 23 June 2015), 2015. 6674, 6701, 6707

Lucke, R. L., Korwan, D. R., Bevilacqua, R. M., Hornstein, J. S., Shettle, E. P., Chen, D. T., Daehler, M., Lumpe, J. D., Fromm, M. D., Debrestian, D., Neff, B., Squire, M., König-Langlo, G., and Davies, J.: The Polar Ozone and Aerosol Measurement (POAM) III instrument and early validation results, J. Geophys. Res., 104, 18785–18800, doi:10.1029/1999JD900235, 1999. 6675

Lumpe, J. D., Bevilacqua, R. M., Hoppel, K. W., Krigman, S. S., Kriebel, D. L., Debrestian, D. J., Randall, C. E., Rusch, D. W., Brogniez, C., Ramananahérisoa, R., Shettle, E. P., Olivero, J. J., Lenoble, J., and Pruvost, P.: POAM II retrieval algorithm and error analysis, J. Geophys. Res., 102, 23593–23614, doi:10.1029/97JD00906, 1997. 6675

Lumpe, J. D., Bevilacqua, R., Hoppel, K., and Randall, C.: POAM III retrieval algorithm and error analysis, J. Geophys. Res., 107, 4575, doi:10.1029/2002JD002137, 2002. 6676

AMTD
8, 6661–6757, 2015

Mauldin, L., Zaun, N., McCormick, M., Guy, J. H., and W.R., V.: Stratospheric Aerosol and Gas Experiment II instrument: a functional description, Opt. Eng., 24, 307–312, 1985. 6671
Mauldin, L. E., Salikhov, R., Habib, S., Vladimirov, A. G., Carraway, D., Petrenko, G., and Comella, J.: Meteor-3M(1)/Stratospheric Aerosol and Gas Experiment III (SAGE III) jointly sponsored by the National Aeronautics and Space Administration and the Russian Space Agency, in: Asia-Pacific Symposium on Remote Sensing of the Atmosphere, Environment, and Space, 3501, 355–365, doi:10.1117/12.317767, 1998. 6672
McDermid, I. S., Godin, S. M., and Lindqvist, L. O.: Ground-based laser DIAL system for long-term measurements of stratospheric ozone, Appl. Opt., 29, 3603–3612, doi:10.1364/AO.29.003603, 1990. 6689
McDermid, I. S., Walsh, T. D., Deslis, A., and White, M. L.: Optical systems design for a stratospheric lidar system, Appl. Opt., 34, 6201–6210, doi:10.1364/AO.34.006201, 1995. 6690
McElroy, C., Nowlan, C., Drummond, J., Bernath, P., Barton, D., Dufour, D., Midwinter, C., Hall, R., Ogyu, A., Ullberg, A., Wardle, D., Kar, J., Zou, J., Nichitiu, F., Boone, C., Walker, K., and Rowlands, N.: The ACE-MAESTRO instrument on SCISAT: description, performance, and preliminary results, Appl. Opt., 46, 4341–4356, doi:10.1364/AO.46.004341, 2007. 6682
McGee, T. J., Whiteman, D. N., Ferrare, R. A., Butler, J. J., and Burris, J. F.: STROZ LITE: Stratospheric Ozone Lidar Trailer Experiment, Opt. Eng., 30, 31–39, doi:10.1117/12.55771, 1991. 6670
McLinden, C. A. and Fioletov, V.: Quantifying stratospheric ozone trends: Complications due to stratospheric cooling, Geophys. Res. Lett., 38, L03808, doi:10.1029/2010GL046012, 2011. 6705
McLinden, C. A., McConnell, J. C., Griffioen, E., and McElroy, C. T.: A vector radiative-transfer model for the Odin/OSIRIS project, Can. J. Phys., 80, 375–393, doi:10.1139/p01-156, 2002. 6680
McLinden, C. A., Bourassa, A. E., Brohede, S., Cooper, M., Degenstein, D. A., Evans, W. F. J., Gattinger, R. L., Hal ey, C. S., Llewell yn, 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. 6676
Mégie, G., Allain, J. Y., Chanin, M. L., and Blamont, J. E.: Vertical profile of stratospheric ozone by lidar sounding from the ground, Nature, 270, 329–331, doi:10.1038/270329a0, 1977. 6669

AMTD
8, 6661–6757, 2015

Mieruch, S., Weber, M., von Savigny, C., Rozanov, A., Bovensmann, H., Burrows, J. P., Bernath, P. F., Boone, C. D., Froidevaux, L., Gordley, L. L., Mlynczak, M. G., Russell III, J. M., Thomason, L. W., Walker, K. A., and Zawodny, J. M.: Global and long-term comparison of SCIAMACHY limb ozone profiles with correlative satellite data (2002-2008), Atmos. Meas. Tech., 5, 771–788, doi:10.5194/amt-5-771-2012, 2012. 6666

Morris, G., Gleason, J., Russell III, J., Schoeberl, M., and McCormick, M.: A comparison of HALOE V19 with SAGE II V6.00 ozone observations using trajectory mapping, J. Geophys. Res., 107, 4177, doi:10.1029/2001JD000847, 2002. 6673, 6691, 6692

Muhlbauer, A., Spichtinger, P., and Lohmann, U.: Application and Comparison of Robust Linear Regression Methods for Trend Estimation, J. Appl. Meteorol. Clim., 48, 1961–1970, doi:10.1175/2009JAMC1851.1, 2009. 6687

Nair, P. J., Godin-Beekmann, S., Pazmiño, A., Hauchecorne, A., Ancellet, G., Petropavlovskikh, I., Flynn, L. E., and Froidevaux, L.: Coherence of long-term stratospheric ozone vertical distribution time series used for the study of ozone recovery at a northern mid-latitude station, Atmos. Chem. Phys., 11, 4957–4975, doi:10.5194/acp-11-4957-2011, 2011. 6678

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. 6666, 6670, 6689, 6713

NASA LaRC: SAGE III Data Products User's Guide, Tech. rep. LaRC 475-03-060, NASA Langley Research Center, https://eosweb.larc.nasa.gov/sites/default/files/project/sage3/guide/Data_Product_User_Guide.pdf (last access: 23 June 2015), 2004. 6672

Naval Research Lab: Description of POAM III Version 4 Retrievals, https://eosweb.larc.nasa.gov/sites/default/files/project/poam3/readme/poam3_ver4_documentation.pdf (last access: 23 June 2015), 2005. 6676, 6742

Nazaryan, H., McCormick, M. P., and Russell III, J. M.: New studies of SAGE II and HALOE ozone profile and long-term change comparisons, J. Geophys. Res., 110, 1–13, doi:10.1029/2004JD005425, 2005. 6673, 6691, 6692

Oman, L. D., Plummer, D. A., Waugh, D. W., Austin, J., Scinocca, J. F., Douglass, A. R., Salawitch, R. J., Canty, T., Akiyoshi, H., Bekki, S., Braesicke, P., Butchart, N., Chipperfield, M. P., Cugnet, D., Dhomse, S., Eyring, V., Frith, S., Hardiman, S. C., Kinnison, D. E., Lamarque,

AMTD
8, 6661–6757, 2015

J.-F., Mancini, E., Marchand, M., Michou, M., Morgenstern, O., Nakamura, T., Nielsen, J. E., Olivié, D., Pitari, G., Pyle, J., Rozanov, E., Shepherd, T. G., Shibata, K., Stolarski, R. S., Teyssèdre, H., Tian, W., Yamashita, Y., and Ziemke, J. R.: Multimodel assessment of the factors driving stratospheric ozone evolution over the 21st century, J. Geophys. Res. Atmos., 115, n/a–n/a, doi:10.1029/2010JD014362, d24306, 2010. 6665

Penckwitt, A. A., Bodeker, G. E., Revell, L. E., Richter, L., Kyrölä, E., and Young, P. J.: Construction and analysis of a new merged SAGE II-GOMOS ozone profile data set for 1984–2012, Earth Syst. Sci. Data, in preparation, 2015. 6713

Rahpoe, N., von Savigny, C., Weber, M., Rozanov, A., Bovensmann, H., and Burrows, J. P.: Error budget analysis of SCIAMACHY limb ozone profile retrievals using the SCIATRAN model, Atmos. Meas. Tech., 6, 2825–2837, doi:10.5194/amt-6-2825-2013, 2013. 6680

Randall, C. E., Rusch, D. W., Bevilacqua, R. M., Hoppel, K. W., Lumpe, J. D., Shettle, E., Thompson, E., Deaver, L., Zawodny, J., Kyrö, E., Johnson, B., Kelder, H., Dorokhov, V. M., König-Langlo, G., and Gil, M.: Validation of POAM III ozone: Comparison with ozonesonde and satellite data, J. Geophys. Res., 108, 4367, doi:10.1029/2002JD002944, 2003. 6676, 6700, 6702, 6707

Raspollini, P., Carli, B., Carlotti, M., Ceccherini, S., Dehn, A., Dinelli, B. M., Dudhia, A., Flaud, J.-M., López-Puertas, M., Niro, F., Remedios, J. J., Ridolfi, M., Sembhi, H., Sgheri, L., and von Clarmann, T.: Ten years of MIPAS measurements with ESA Level 2 processor V6 – Part 1: Retrieval algorithm and diagnostics of the products, Atmos. Meas. Tech., 6, 2419–2439, doi:10.5194/amt-6-2419-2013, 2013. 6679, 6680

Rozanov, A., Eichmann, K.-U., von Savigny, C., Bovensmann, H., Burrows, J. P., von Bargen, A., Doicu, A., Hilgers, S., Godin-Beekmann, S., Leblanc, T., and McDermid, I. S.: Comparison of the inversion algorithms applied to the ozone vertical profile retrieval from SCIAMACHY limb measurements, Atmos. Chem. Phys., 7, 4763–4779, doi:10.5194/acp-7-4763-2007, 2007. 6680

Rusch, D., Bevilacqua, R., Randall, C., Lumpe, J., Hoppel, K., Fromm, M., Debrestian, D., Olivero, J., Hornstein, J., Guo, F., and Shettle, E.: Validation of POAM II Ozone Measurements with Coincident MLS, HALOE, and SAGE II Observations, J. Geophys. Res., 102, 23615–23627, doi:10.1029/97JD00458, 1997. 6675

Russell III, J. M., Gordley, L. L., Park, J. H., Drayson, S. R., Hesketh, W. D., Cicerone, R. J., Tuck, A. F., Frederick, J. E., Harries, J. E., and Crutzen, P. J.: The Halogen Occultation Experiment, J. Geophys. Res., 98, 10 777–10 797, doi:10.1029/93JD00799, 1993. 6672

AMTD
8, 6661–6757, 2015

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, doi:10.5194/acp-15-829-2015, 2015. 6671, 6673, 6681

Schwartz, M. J., Lambert, A., Manney, G. L., Read, W. G., Livesey, N. J., Froidevaux, L., Ao, C. O., Bernath, P. F., Boone, C. D., Cofield, R. E., Daffer, W. H., Drouin, B. J., Fetzer, E. J., Fuller, R. A., Jarnot, R. F., Jiang, J. H., Jiang, Y. B., Knosp, B. W., Krüger, K., Li, J.-L. F., Mlynczak, M. G., Pawson, S., Russell, J. M., Santee, M. L., Snyder, W. V., Stek, P. C., Thurstans, R. P., Tompkins, A. M., Wagner, P. A., Walker, K. A., Waters, J. W., and Wu, D. L.: Validation of the Aura Microwave Limb Sounder temperature and geopotential height measurements, J. Geophys. Res., 113, D15S11, doi:10.1029/2007JD008783, 2008. 6674

Seidel, D. J., Gillett, N. P., Lanzante, J. R., Shine, K. P., and Thorne, P. W.: Stratospheric temperature trends: our evolving understanding, WIREs Climate Change, 2, 592–616, doi:10.1002/wcc.125, 2011. 6705

Simmons, A. J., Poli, P., Dee, D. P., Berrisford, P., Hersbach, H., Kobayashi, S., and Peubey, C.: Estimating low-frequency variability and trends in atmospheric temperature using ERA-Interim, Q. J. Roy. Meteor. Soc., 140, 329–353, doi:10.1002/qj.2317, 2014. 6686, 6705, 6706

Sioris, C. E., McLinden, C. A., Fioletov, V. E., Adams, C., Zawodny, J. M., Bourassa, A. E., Roth, C. Z., and Degenstein, D. A.: Trend and variability in ozone in the tropical lower stratosphere over 2.5 solar cycles observed by SAGE II and OSIRIS, Atmos. Chem. Phys., 14, 3479–3496, doi:10.5194/acp-14-3479-2014, 2014. 6713, 6714

Smit, H., Straeter, W., Johnson, B., Oltmans, S., Davies, J., Tarasick, D., Hoegger, B., Stübi, R., Schmidlin, F., Northam, T., Thompson, A., Witte, J.-C., Boyd, I., and Posny, F.: Assessment of the performance of ECC-ozonesondes under quasi-flight conditions in the environmental simulation chamber: Insights from the Juelich Ozone Sonde Intercomparison Experiment (JOSIE), J. Geophys. Res., 112, 1–18, doi:10.1029/2006JD007308, 2007. 6669

Smit, H., Oltmans, S., Deshler, T., Tarasick, D., Johnson, B., Schmidlin, F., Stübi, R., and Davies, J.: SI2N/O3S-DQA Activity: Guide Lines for Homogenization of Ozone Sonde Data, Tech. rep. 2.0, http://www-das.uwyo.edu/~deshler/NDACC_O3Sondes/O3s_DQA/O3S-DQA-GuidelinesHomogenization-V2-19November2012.pdf (last access: 23 June 2015), 2012. 6712

AMTD
8, 6661–6757, 2015

Smit, H. G. J. and ASOPOS-panel: Quality Assurance and Quality Control for Ozonesonde Measurements in GAW, WMO Global Atmosphere Watch report series 201, World Meteorological Organization, http://www.wmo.int/pages/prog/arep/gaw/documents/FINAL_GAW_201_Oct_2014.pdf (last access: 23 June 2015), 2014. 6668, 6669

Sofieva, V. F., Rahpoe, N., Tamminen, J., Kyrölä, E., Kalakoski, N., Weber, M., Rozanov, A., von Savigny, C., Laeng, A., von Clarmann, T., Stiller, G., Lossow, S., Degenstein, D., Bourassa, A., Adams, C., Roth, C., Lloyd, N., Bernath, P., Hargreaves, R. J., Urban, J., Murtagh, D., Hauchecorne, A., Dalaudier, F., van Roozendael, M., Kalb, N., and Zehner, C.: Harmonized dataset of ozone profiles from satellite limb and occultation measurements, Earth System Science Data, 5, 349–363, doi:10.5194/essd-5-349-2013, 2013. 6713

Stauffer, R. M., Morris, G. A., Thompson, A. M., Joseph, E., Coetzee, G. J. R., and Nalli, N. R.: Propagation of radiosonde pressure sensor errors to ozonesonde measurements, Atmos. Meas. Tech., 7, 65–79, doi:10.5194/amt-7-65-2014, 2014. 6686, 6705

Street, J. O., Carroll, R. J., and Ruppert, D.: A Note on Computing Robust Regression Estimates Via Iteratively Reweighted Least Squares, The American Statistician, 42, 152–154, doi:10.1080/00031305.1988.10475548, 1988. 6687, 6688

Sun, B., Reale, A., Schroeder, S., Seidel, D. J., and Ballish, B.: Toward improved corrections for radiation-induced biases in radiosonde temperature observations, J. Geophys. Res., 118, 4231–4243, doi:10.1002/jgrd.50369, 2013. 6686

Tamminen, J., Kyrölä, E., Sofieva, V. F., Laine, M., Bertaux, J.-L., Hauchecorne, A., Dalaudier, F., Fussen, D., Vanhellemont, F., Fanton-d'Andon, O., Barrot, G., Mangin, A., Guirlet, M., Blanot, L., Fehr, T., Saavedra de Miguel, L., and Fraisse, R.: GOMOS data characterisation and error estimation, Atmos. Chem. Phys., 10, 9505–9519, doi:10.5194/acp-10-9505-2010, 2010. 6678

Tegtmeier, S., Hegglin, M. I., Anderson, J., Bourassa, A., Brohede, S., Degenstein, D., Froidevaux, L., Fuller, R., Funke, B., Gille, J., Jones, A., Kasai, Y., Krüger, K., Kyrölä, E., Lingenfelser, G., Lumpe, J., Nardi, B., Neu, J., Pendlebury, D., Remsberg, E., Rozanov, A., Smith, L., Toohey, M., Urban, J., von Clarmann, T., Walker, K. A., and Wang, H. J.: The SPARC Data Initiative: A comparison of ozone climatologies from international satellite limb sounders, J. Geophys. Res., 118, 12229–12247, doi:10.1002/2013JD019877, 2013. 6666, 6696, 6697, 6713

AMTD
8, 6661–6757, 2015

Terao, Y. and Logan, J.: Consistency of time series and trends of stratospheric ozone as seen by ozonesonde, SAGE II, HALOE, and SBUV(/2), J. Geophys. Res., 112, 1–23, doi:10.1029/2006JD007667, 2007. 6666

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, doi:10.5194/acp-10-2159-2010, 2010. 6672

Thompson, A. M., Witte, J. C., Smit, H. G. J., Oltmans, S. J., Johnson, B. J., Kirchhoff, V. W. J. H., and Schmidlin, F. J.: Southern Hemisphere Additional OZonesondes (SHADOZ) 1998-2004 tropical ozone climatology: 3. Instrumentation, station-to-station variability, and evaluation with simulated flight profiles, J. Geophys. Res., 112, 1–18, doi:10.1029/2005JD007042, 2007. 6669

Thompson, A. M. et al.: Southern Hemisphere Additional Ozonesondes (SHADOZ) ozone climatology (2005-2009): Tropospheric and tropical tropopause layer (TTL) profiles with comparisons to OMI-based ozone products, J. Geophys. Res., 117, D23 301, doi:10.1029/2011JD016911, 2012. 6669

Thorne, P. W., Lanzante, J. R., Peterson, T. C., Seidel, D. J., and Shine, K. P.: Tropospheric temperature trends: history of an ongoing controversy, WIREs Climate Change, 2, 66–88, doi:10.1002/wcc.80, 2011. 6705

Toohey, M., Hegglin, M. I., Tegtmeier, S., Anderson, J., Añel, J. A., Bourassa, A., Brohede, S., Degenstein, D., Froidevaux, L., Fuller, R., Funke, B., Gille, J., Jones, A., Kasai, Y., Krüger, K., Kyrölä, E., Neu, J. L., Rozanov, A., Smith, L., Urban, J., von Clarmann, T., Walker, K. A., and Wang, R. H. J.: Characterizing sampling biases in the trace gas climatologies of the SPARC Data Initiative, J. Geophys. Res., 118, 11,847–11,862, doi:10.1002/jgrd.50874, 2013. 6684

Tummon, F., Hassler, B., Harris, N. R. P., Staehelin, J., Steinbrecht, W., Anderson, J., Bodeker, G. E., Bourassa, A., Davis, S. M., Degenstein, D., Frith, S. M., Froidevaux, L., Kyrölä, E., Laine, M., Long, C., Penckwitt, A. A., Sioris, C. E., Rosenlof, K. H., Roth, C., Wang, H.-J., and Wild, J.: Intercomparison of vertically resolved merged satellite ozone data sets: interannual variability and long-term trends, Atmos. Chem. Phys., 15, 3021–3043, doi:10.5194/acp-15-3021-2015, 2015. 6666, 6713, 6714, 6715, 6718

Urban, J., Lautié, N., Le Flochmoën, E., Jiménez, C., Eriksson, P., de La Noë, J., Dupuy, E., Ekström, M., El Amraoui, L., Frisk, U., Murtagh, D., Olberg, M., and Ricaud, P.:Odin/SMR limb observations of stratospheric trace gases: Level 2 processing of ClO, N_2O, HNO_3, and O_3, J. Geophys. Res., 110, doi:10.1029/2004JD005741, 2005. 6677

AMTD
8, 6661–6757, 2015

van der A, R. J. et al.: Ozone_cci – User Requirement Document (URD), Tech. rep. v2.1, KNMI, http://www.esa-ozone-cci.org/?q=webfm_send/37 (last access: 23 June 2015), 2011. 6711
van Gijsel, J. A. E., Swart, D. P. J., Baray, J.-L., Bencherif, H., Claude, H., Fehr, T., Godin-Beekmann, S., Hansen, G. H., Keckhut, P., Leblanc, T., McDermid, I. S., Meijer, Y. J., Nakane, H., Quel, E. J., Stebel, K., Steinbrecht, W., Strawbridge, K. B., Tatarov, B. I., and Wolfram, E. A.: GOMOS ozone profile validation using ground-based and balloon sonde measurements, Atmos. Chem. Phys., 10, 10473–10488, doi:10.5194/acp-10-10473-2010, 2010. 6678
Verhoelst, T., Granville, J., Hendrick, F., Köhler, U., Lerot, C., Pommereau, J.-P., Redondas, A., Van Roozendael, M., and Lambert, J.-C.: Metrology of ground-based satellite validation: Co-location mismatch and smoothing issues of total ozone comparisons, Atmos. Meas. Tech. Disc., submitted, 2015. 6712
von Clarmann, T.: Validation of remotely sensed profiles of atmospheric state variables: strategies and terminology, Atmos. Chem. Phys., 6, 4311–4320, doi:10.5194/acp-6-4311-2006, 2006. 6684
Wang, H. J., Cunnold, D. M., Froidevaux, L., and Russell, J. M.: A reference model for middle atmosphere ozone in 1992-1993, J. Geophys. Res., 104, 21 629–21 643, doi:10.1029/1999JD900412, 1999. 6707
Wang, H. J., Cunnold, D. M., Thomason, L. W., Zawodny, J. M., and Bodeker, G. E.: Assessment of SAGE version 6.1 ozone data quality, J. Geophys. Res., 107, 4691, doi:10.1029/2002JD002418, 2002. 6671, 6707, 6742
Wang, H.-J., Cunnold, D. M., Trepte, C., Thomason, L. W., and Zawodny, J. M.: SAGE III solar ozone measurements: Initial results, Geophys. Res. Lett., 33, doi:10.1029/2005GL025099, 2006. 6672
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., Gun-Shing, C., Chudasama, B. V., Dodge, R., Fuller, R. A., Girard, M. A., Jiang, J. H., Yibo, J., 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., Van Snyder, W., Tope, M. C., Wagner, P. A., and Walch, M. J.: The Earth Observing System Microwave

AMTD
8, 6661–6757, 2015

Limb Sounder (EOS MLS) on the Aura Satellite, IEEE T. Geosci. Remote., 44, 1075–1092, doi:10.1109/TGRS.2006.873771, 2006. 6674

Waugh, D. W., Oman, L., Kawa, S. R., Stolarski, R. S., Pawson, S., Douglass, A. R., Newman, P. A., and Nielsen, J. E.: Impacts of climate change on stratospheric ozone recovery, Geophys. Res. Lett., 36, L03805, doi:10.1029/2008GL036223, 2009. 6665

Waymark, C., Walker, K., et al.: Report on Current ACE-FTS Data Versions: Processing and Validation, in: Proceedings of the ACVE workshop, Frascati, Italy, March 2013, 2013. 6681

WMO: Scientific Assessment of Ozone Depletion: 2010, Global Ozone Research and Monitoring Project – Report No. 52, http://ozone.unep.org/Assessment_Panels/SAP/Scientific_Assessment_2010/00-SAP-2010-Assement-report.pdf (last access: 23 June 2015), 2011. 6665

WMO: Scientific Assessment of Ozone Depletion: 2014, Global Ozone Research and Monitoring Project – Report No. 55, http://ozone.unep.org/Assessment_Panels/SAP/Scientific_Assessment_2014/Scientific_Assessment_Report_2014.pdf (last access: 23 June 2015), 2014. 6665, 6666, 6714, 6715, 6718

AMTD
8, 6661–6757, 2015

Table 1. Overview of the 72 ozonesonde stations considered in this work, their location and the archive the data were taken from. Period and profile statistics reflect the total, screened sample straddling the analysis period (10/1984–5/2013), not the co-located sampled (which differs per satellite instrument). The mark in the last column indicates whether the station is used for the drift analysis.

Station	Lat. (° N)	Lon. (° E)	Data archive*	Period		$N_{profile}$	Included in drift analysis
Alert	82.5	−62.5	WOUDC	12/1987	12/2011	1244	✓
Eureka	80.0	−85.9	WOUDC	11/1992	9/2011	1318	✓
Ny-Ålesund	78.9	11.9	WOUDC	10/1990	5/2013	2224	✓
Thule	76.5	−68.7	NDACC	10/1991	1/2013	349	✓
Resolute	74.7	−95.0	WOUDC	10/1984	8/2011	1020	✓
Summit	72.3	−38.3	NDACC	2/2005	5/2013	427	✓
Scoresbysund	70.5	−21.9	NDACC	2/1989	5/2013	1169	✓
Sodankylä	67.4	26.6	NDACC	11/1991	12/2010	1085	✓
Edmonton	53.5	−114.1	WOUDC	10/1984	8/2011	1193	✓
Goose Bay	53.3	−60.4	WOUDC	10/1984	8/2011	1272	✓
Lindenberg	52.2	14.1	WOUDC	10/1984	5/2013	1660	✓
De Bilt	52.1	5.2	NDACC	11/1992	12/2012	1061	✓
Vanscoy	52.0	−107.0	WOUDC	8/1990	9/2004	60	
Valentia	51.9	−10.2	WOUDC	1/1994	12/2012	555	✓
Uccle	50.8	4.3	WOUDC	10/1984	6/2012	3712	✓
Gimli	50.6	−97.0	WOUDC	7/1985	8/1985	10	
Bratt's Lake	50.2	−104.7	WOUDC	12/2003	9/2011	402	✓
Praha	50.0	14.4	WOUDC	1/1985	4/2013	1210	✓
Kelowna	49.9	−119.4	WOUDC	11/2003	8/2011	432	✓
Hohenpeißenberg	47.8	11.0	WOUDC	10/1984	5/2013	3586	✓
Payerne	46.8	7.0	WOUDC	10/1984	12/2012	4052	✓
Pellston	45.6	−84.7	WOUDC	7/2004	8/2004	5	
Pietro Capofiume	44.6	11.6	WOUDC	3/1991	12/1993	95	
Egbert	44.2	−79.8	WOUDC	12/2003	8/2011	373	✓
Yarmouth	43.9	−66.1	WOUDC	10/2003	8/2011	394	✓
Sofia	42.8	23.4	WOUDC	11/1984	12/1991	145	
Trinidad Head	40.8	−124.2	WOUDC	1/1999	8/2006	197	✓
Madrid	40.5	−3.7	WOUDC	12/1994	5/2013	738	✓
Boulder	40.0	−105.2	NDACC	6/1991	5/2013	1097	✓
Beltsville	39.0	−76.5	WOUDC	8/2006	8/2006	12	
Huntsville	34.7	−86.6	WOUDC	4/1999	12/2007	574	✓
Table Mountain	34.4	−117.7	WOUDC	2/2006	8/2006	35	✓
Isfahan	32.5	51.4	WOUDC	7/1995	4/2011	151	✓
Palestine	31.8	−95.7	WOUDC	5/1985	6/1985	26	

AMTD
8, 6661–6757, 2015

Table 1. Continued.

Station	Lat. (°N)	Lon. (°E)	Data archive*	Period		$N_{profile}$	Included in drift analysis
Houston	29.7	−95.4	WOUDC	7/2004	8/2006	62	✓
Santa Cruz	28.5	−16.3	WOUDC	1/1996	5/2003	322	
Izaña	28.3	−16.5	NDACC	1/1995	3/2012	976	✓
Taipei	25.0	121.5	WOUDC	1/2000	8/2001	64	
Hilo	19.7	−155.1	NDACC	7/1991	6/2010	855	✓
Tecamec	19.3	−99.2	WOUDC	3/2006	9/2006	34	
Barbados	13.2	−59.4	WOUDC	7/2006	8/2006	27	
Cotonou	6.2	2.2	SHADOZ	1/2005	1/2007	97	
Paramaribo	5.8	−55.2	NDACC	9/1999	5/2013	534	✓
Kaashidhoo	5.0	73.5	WOUDC	1/1999	3/1999	54	
San Cristóbal	−0.9	−89.6	SHADOZ/WOUDC	3/1998	10/2008	708	✓
Nairobi	−1.3	36.8	SHADOZ/WOUDC	12/1996	12/2012	1058	✓
Malindi	−3.0	40.2	SHADOZ	3/1999	1/2006	191	✓
Brazzaville	−4.3	15.2	WOUDC	4/1990	10/1992	80	
Natal	−5.8	−35.2	SHADOZ/WOUDC	3/1990	12/2010	650	✓
Watukosek	−7.5	112.6	SHADOZ	8/1999	12/2011	573	✓
Ascension Island	−8.0	−14.4	SHADOZ/WOUDC	7/1990	8/2010	1112	✓
Porto Nacional	−10.8	−48.4	WOUDC	9/1992	10/1992	15	
Samoa	−14.2	−170.6	NDACC	8/1995	5/2013	663	✓
Cuiaba	−15.6	−56.1	WOUDC	9/1992	10/1992	22	
Papeete	−18.0	−149.0	SHADOZ/WOUDC	7/1995	12/1999	167	
Suva	−18.1	178.4	SHADOZ	2/1997	12/2011	727	✓
Etosha Pan	−19.2	15.9	WOUDC	9/1992	10/1992	15	
Réunion Island	−20.9	55.5	SHADOZ	1/1998	11/2012	810	✓
Irene	−25.9	28.2	SHADOZ/WOUDC	7/1990	10/2007	581	✓
Easter Island	−27.2	−109.4	WOUDC	8/1995	6/1997	71	
Broadmeadows	−37.7	144.9	WOUDC	2/1999	12/2012	623	✓
Laverton	−37.9	144.8	WOUDC	10/1984	2/1999	344	✓
Lauder	−45.0	169.7	NDACC	8/1986	5/2013	1609	✓
Macquarie	−54.5	158.9	WOUDC	3/1994	12/2012	712	✓
Marambio	−64.2	−56.6	WOUDC	11/1988	5/2013	891	✓
Mirny	−66.5	93.0	WOUDC	7/1989	12/1991	114	
Davis	−68.6	78.0	WOUDC	2/2003	12/2012	282	✓
Syowa	−69.0	39.6	WOUDC	12/1984	5/2013	1134	✓
Neumayer	−70.7	−8.3	WOUDC	3/1992	5/2013	1540	✓
Novolasarevskaya	−70.8	11.9	WOUDC	5/1985	2/1991	374	
Mac Murdo	−77.8	166.6	NDACC	8/1986	10/2010	817	✓
Amundsen-Scott	−90.0	−24.8	NDACC	11/1990	5/2013	1463	✓

*Sources: NDACC, http://www.ndacc.org; WOUDC, http://www.woudc.org; SHADOZ, http://croc.gsfc.nasa.gov/shadoz.

AMTD
8, 6661–6757, 2015

Table 2. Like Table 1, but for the 13 considered stratospheric ozone lidar stations.

Station	Lat. (° N)	Lon. (° E)	Data archive	Period		$N_{profile}$	Included in drift analysis
Eureka	80.0	−85.9	NDACC	2/1993	3/2009	513	
Ny-Ålesund	78.9	11.9	NDACC	11/1991	3/2011	791	
Andøya	69.3	16.0	NDACC	12/1994	4/2011	594	
Hohenpeißenberg	47.8	11.0	NDACC	9/1987	5/2013	2280	
Observatoire de Haute-Provence	43.9	5.7	NDACC	7/1985	5/2013	2776	✓
Toronto	43.8	−79.5	NDACC	5/1991	12/1997	235	
Tsukuba	36.0	140.1	NDACC	8/1988	2/2010	592	
Table Mountain	34.4	−117.7	NDACC	1/1989	5/2013	1758	✓
Mauna Loa	19.5	−155.6	NDACC	7/1993	5/2013	2401	✓
Réunion Island	−20.9	55.5	NDACC	5/2000	12/2006	85	
Lauder	−45.0	169.7	NDACC	11/1994	6/2011	1030	✓
Rio Gallegos	−51.6	−69.3	NDACC	8/2005	11/2010	140	
Dumont d'Urville	−66.7	140.0	NDACC	4/1991	2/2013	678	

AMTD
8, 6661–6757, 2015

Table 3. Overview of satellite ozone profile data records.

Instrument	L2 data version	Observation geometry	Spectral region O_3 retrieval	Analysis period		Latitude range
SAGE II	v7.0	solar occ.	VIS	10/1984	8/2005	80° N – 80° S
SAGE III	v4.0	solar occ.	UV-VIS	3/2002	11/2005	50–80° N (SS), 30–50° S (SR)
HALOE	v19	solar occ.	MIR	10/1991	11/2005	80° N – 80° S
UARS MLS	v5	limb	MW	9/1991	*6/1997	34° N – 80° S, 34° S – 80° N
Aura MLS	v3.3	limb	MW	8/2004	5/2013	82° N – 82° S
POAM II	v6	solar occ.	VIS	11/1993	11/1996	55–71° N (SR), 63–88° S (SS)
POAM III	v4	solar occ.	VIS	4/1998	12/2005	55–71° N (SR), 63–88° S (SS)
OSIRIS	v5.07	limb	UV-VIS	10/2001	5/2013	82° N – 82° S
SMR	v2.1 (501.8 GHz)	limb	MW	6/2001	5/2013	82° N – 82° S
GOMOS	IPF 6.01	stellar occ.	UV-VIS	7/2002	4/2012	90° N – 90° S
MIPAS	ML2PP 6.0 (OR)	limb	MIR	*1/2005	4/2012	80° N – 80° S
SCIAMACHY	SGP 5.02	limb	VIS	8/2002	4/2012	82° N – 80° S
ACE-FTS	v3.0	solar occ.	MIR	2/2004	*9/2010	85° N – 85° S
MAESTRO	v1.2 (VIS)	solar occ.	VIS	2/2004	*9/2010	85° N – 85° S

*Part of the record was not considered for this analysis.

AMTD
8, 6661–6757, 2015

Table 3. Continued.

Instrument	Vertical range [km]	Vertical resolution [km]	Native profile representation	Source *P/T*	Screening reference
SAGE II	CT–60	1	(z, n)	MERRA (+GRAM-95)	Wang et al. (2002)
SAGE III	6–85	1	(z, n)	NCEP (+GRAM-95)	Done by SAGE III team
HALOE	15–60	2.3	(p, VMR)	Retrieved/NCEP	Hervig and McHugh (1999)
UARS MLS	15–60	3–5	(p, VMR)	Retrieved	Livesey et al. (2003)
Aura MLS	10–75	2.5–4	(p, VMR)	Retrieved	Livesey et al. (2011)
POAM II	15–50	1	(z, n)	UKMO	–
POAM III	10–60	1–2	(z, n)	UKMO	Naval Research Lab (2005)
OSIRIS	CT–60	1–2	(z, n)	ECMWF	Done by OSIRIS team
SMR	18–60	3	(z, VMR)	ECMWF	Jones et al. (2009)
GOMOS	12–100	2–3	(z, n)	ECMWF	ESA (2012a)
MIPAS	6–68	3–4	(p, VMR)	Retrieved	ESA (2012b)
SCIAMACHY	15–40	3	(z, n)	McLinden clim.	ESA (2013)
ACE-FTS	CT–95	3–4	(z, VMR)	Retrieved/CMC	Dupuy et al. (2009)
MAESTRO	CT–100	1.5	(z, VMR)	ACE-FTS	Kar et al. (2007)

AMTD
8, 6661–6757, 2015

Table 4. Overview of the drift of satellite ozone profile records relative to ozonesonde and lidar, in the lower, middle and upper stratosphere. For each altitude region we present the range of the network average of the drift ($\bar{\alpha}$) and its adjusted one-sigma uncertainty ($\sigma^*_{\bar{\alpha}}$). Bold values indicate results with more than 2σ significance.

Drift SAT-GND [%/decade]	10–20 km		20–30 km		30–45 km		Remark
	$\bar{\alpha}$	$1\sigma^*_{\bar{\alpha}}$	$\bar{\alpha}$	$1\sigma^*_{\bar{\alpha}}$	$\bar{\alpha}$	$1\sigma^*_{\bar{\alpha}}$	
SAGE II	[−4, −1]	0.8–3.5	[−2, 0]	0.5–0.8	[−2, +2]	1–2.5	very stable
SAGE III	[−15, +5]	5–10	[−10, −2]	3–6	no results		record too short
HALOE	[−4, +4]	1.5–7	**[-7, -1]**	1–1.5	[−5, 0]	1.5–4	significant 20–35 km
UARS MLS	no UARS MLS data		[+1, +5]	2–6	[−3, +7]	6–12	record short
Aura MLS	[−5, −1]	0.8–2	[−1, +3]	0.6–1	[0, +3]	1.5–5	very stable
POAM II	no results		[−15, +15]	12–20	no results		record too short
POAM III	[−3, +10]	2–6	[−10, 0]	2–4	no results		record sparse
OSIRIS	[−4, +1]	1.3–3.5	[+1, +3]	0.8–1	**[+1, +8]**	1–2	significant 37–45 km
SMR	no SMR data		[−4, +5]	1.5–3	[−15, +3]	3–7	SMR profiles noisy
GOMOS	[−10, −4]	4–30	[−5, −1]	2–3	[−2, +3]	2–3	significant 18–25 km
MIPAS (OR)	[−2, +2]	2–6	[−2, +3]	1.5–2	[−4, +1]	2–4	
SCIAMACHY	[−5, +5]	1.5–8	[+1, +3]	1.5–3	**[-8, +1]**	1.5–3	significant 30–45 km, at threshold < 30 km
ACE-FTS	[−5, +2]	4–10	[−5, +5]	3–4	no results		record sparse
MAESTRO	[−5, +10]	4–20	[0, +6]	4–7	no results		record sparse

AMTD
8, 6661–6757, 2015

Table 5. Overview of the bias of satellite ozone profile records relative to ozonesonde and lidar, in the upper troposphere and stratosphere. For each altitude range we present the median relative difference (bias *b*), and whether there are any dependences on latitude and season that depart from the general tendency.

Bias SAT-GND [%]	< TP	TP–20 km	20–30 km	30–45 km	Remark
SAGE II	< −10	±4	[−5, 0]	[0, +3]	
SAGE III	< −20	±3	±5	[+3, +7]	
HALOE	< −20	[−15, 0]	[−5, +3]	±4	
UARS MLS	–	+10	[0, +5]	[−5, +2]	
Aura MLS	±15	±5	±4	±3	vertical oscillations in UTLS
POAM II	–	[−7, −5]	[−7, −5]	–	
POAM III	–	[−3, +10]	[−5, −3]	[0, +5]	
OSIRIS	−10	[−10, −3]	±3	[0, +5]	+5 % at 22 km
SMR	–	+5	−5	−5	negative bias above TP+10 km
GOMOS	−20 (N) +20 (S)	[−10, 0] (N) [+5, +15] (S)	±5	±5	N/S sign change troposphere
MIPAS OR	> +20	[0, +5]	[0, +5]	[+3, +7]	persistent positive bias
SCIAMACHY	> +20	±7	[−5, +20]	[0, +20]	(a) Arctic seasonality, (b) large + bias towards S hemisphere
ACE-FTS	> +10	±3	±3	[0, +5]	
MAESTRO	–	[−15, −3]	[−5, 0]	±3	sharp change at 15 km

AMTD
8, 6661–6757, 2015

Table 6. Decadal stability of merged ozone profile records (Level-3) estimated from our ground-based assessment of the stability of the contributing Level-2 records, in two time periods and three layers of the stratosphere. These drift values could serve as 1σ systematic uncertainty in trend studies. Nevertheless, in expectation of more rigourous analyses, the estimates below should be considered with care, as they may overestimate the actual drift.

Merged record	Pre-1997			Post-1998		
	LS	MS	US	LS	MS	US
SAGE II–GOMOS	1–1.5	1	1–1.5	3–5	1.5–2	1.5–2
SAGE II–OSIRIS	1–1.5	1	1–1.5	2	2	3–5
GOZCARDS	1–1.5	1–1.5	1–1.5	1.5–2	1.5–2	1.5–2
SWOOSH	1–1.5	1–1.5	1–1.5	1.5–2	1.5–2	1.5–2
Ozone_cci		no data		2–2.5	2	2–3.5

*LS: 10–20 km (~ 50–250 hPa), MS: 20–30 km (~ 10–50 hPa), US: 30–45 km (~ 1–10 hPa).

AMTD
8, 6661–6757, 2015

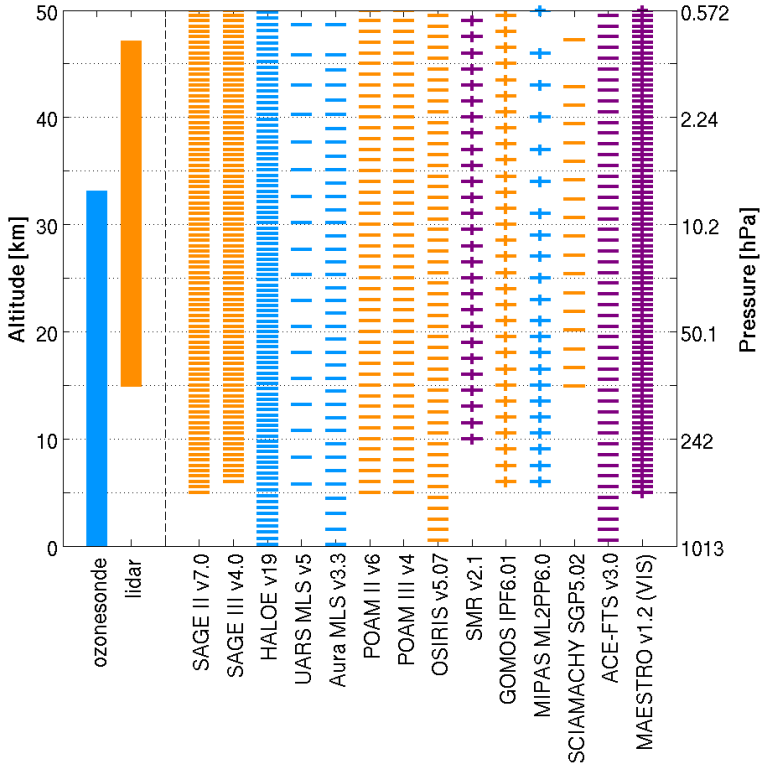


Figure 1. Overview of the native representation of the ozone profile records. A vertical band defines the approximate range for the ground-based data sets, while individual levels are shown for satellite profiles. Colours differentiate between altitude versus O_3 number density (orange), altitude versus ozone VMR (purple) and pressure versus ozone VMR (blue). Vertical grids that are profile-dependent are marked with small vertical bars. Mesospheric measurements are not considered in this work and hence not depicted.

AMTD
8, 6661–6757, 2015

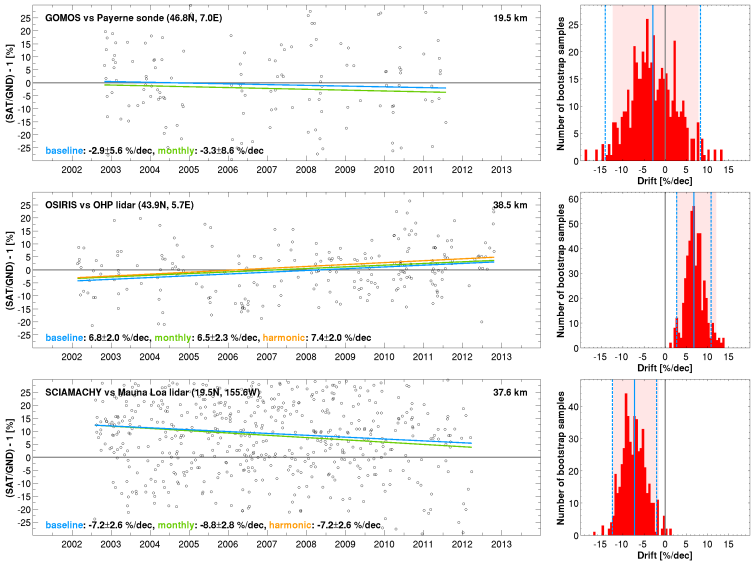


Figure 2. (Left) Time series of the ozone differences for GOMOS (top), OSIRIS (centre) and SCIAMACHY (bottom) relative to the measurements (at three altitudes) from three different ground-based instruments. The blue line depicts the baseline regression model, while the green and orange lines show cross-checks (see text). The estimated drift $\hat{\alpha}$ and its $1\hat{\sigma}_\alpha$ uncertainty is mentioned at the bottom of each panel. (Right) Distribution of the drift obtained from 500 bootstrapped samples of the time series on the left. The light red zone marks the 95 % inter-percentile of the drift distribution, which should be compared to the analytical derivation $\hat{\alpha} \pm 2\hat{\sigma}_\alpha$ (vertical blue lines).

AMTD
8, 6661–6757, 2015

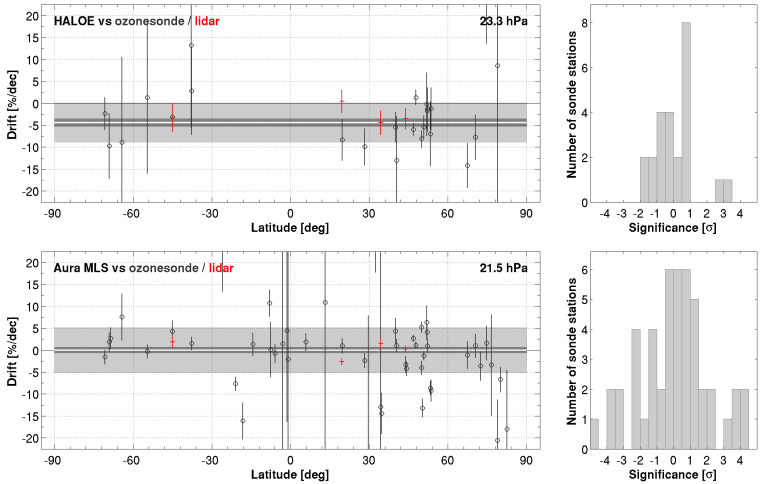


Figure 3. (Left) Each marker (and associated error bar) represents the drift $\hat{\alpha}_j$ (and one sigma uncertainty $\hat{\sigma}_{\alpha,j}$) of HALOE (top) and Aura MLS (bottom) around 25 km relative to a station in the NDACC/GAW/SHADOZ ozonesonde (black) and NDACC lidar (red) networks. The network weighted mean $\bar{\alpha}$ (horizontal white line), its adjusted uncertainty $\sigma^*_{\bar{\alpha}}$ (dark grey area) and the standard deviation of the single site drifts (light grey area) are shown for the ozonesonde network. (Right) Corresponding distribution of ν (defined in the text) for the ozonesonde network.

AMTD
8, 6661–6757, 2015

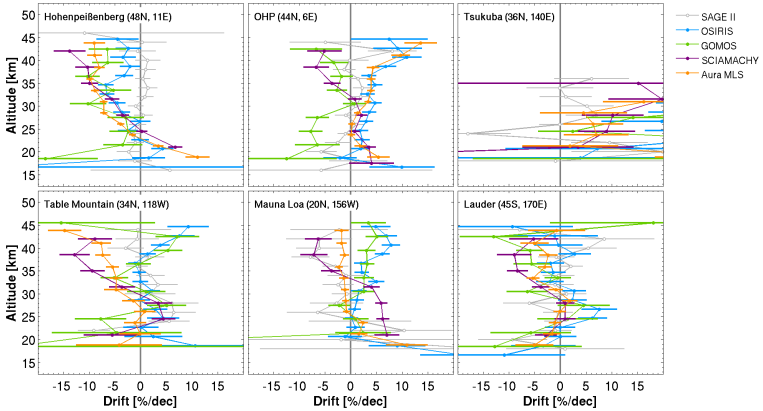


Figure 4. Comparison of the vertical structure of the drift and its 1σ uncertainty of SAGE II and four recent satellite records relative to six lidar stations. Only selected levels are shown to improve the readability.

AMTD
8, 6661–6757, 2015

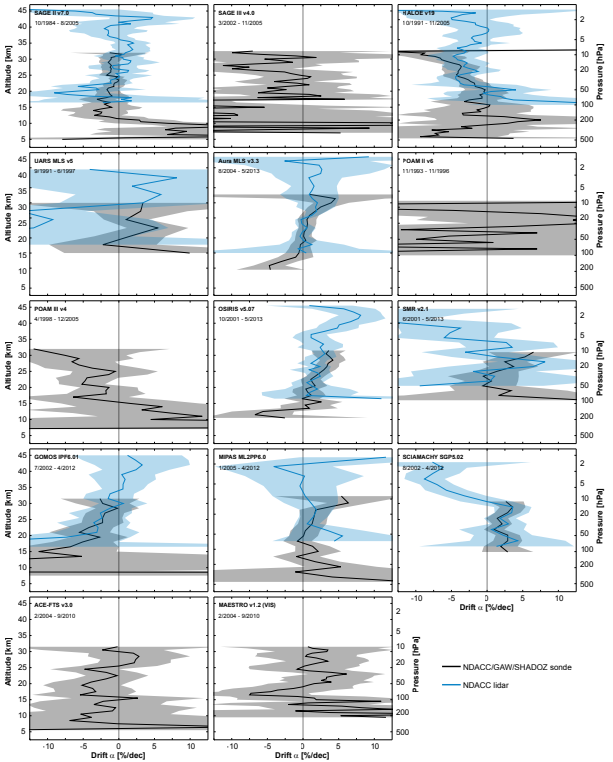


Figure 5. Average drift of each satellite record relative to the entire ozonesonde (black) and lidar network (blue). The shaded region indicates the χ^2-adjusted 95 % confidence interval for the drift. The analysis is performed in the native profile representation of each satellite record.

AMTD
8, 6661–6757, 2015

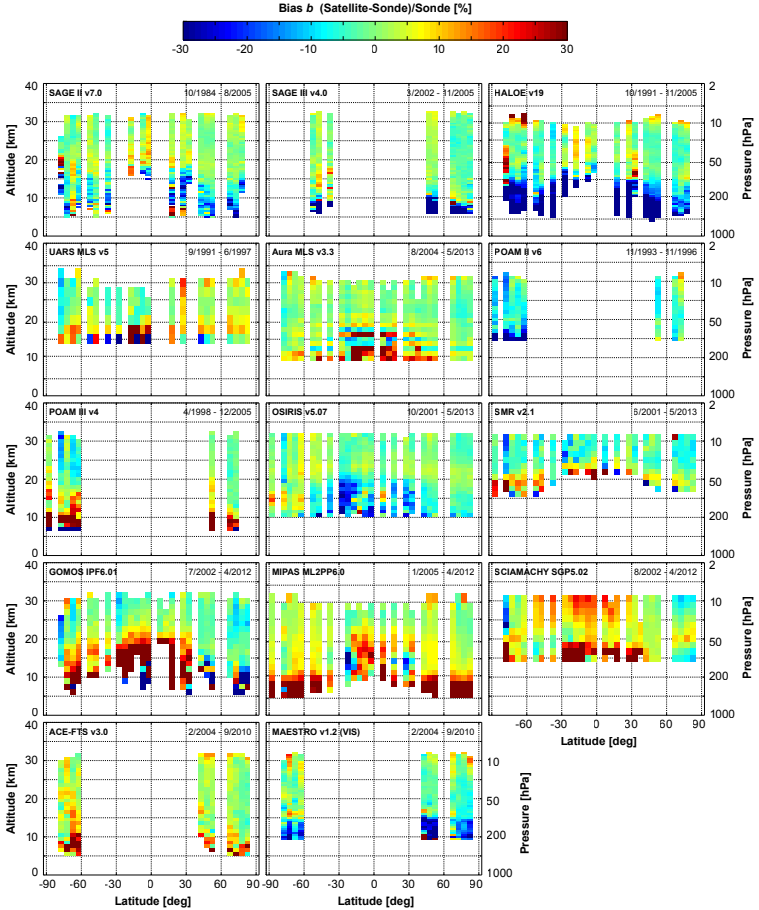


Figure 6. Meridional structure of the bias *b* relative to the ozonesonde network for all fourteen satellite ozone profile records. The analysis is performed in the native profile representation of each satellite record. Only bins with more than 10 comparison pairs are shown.

AMTD
8, 6661–6757, 2015

Ground-based assessment of limb and occultation ozone profile data records

D. Hubert et al.

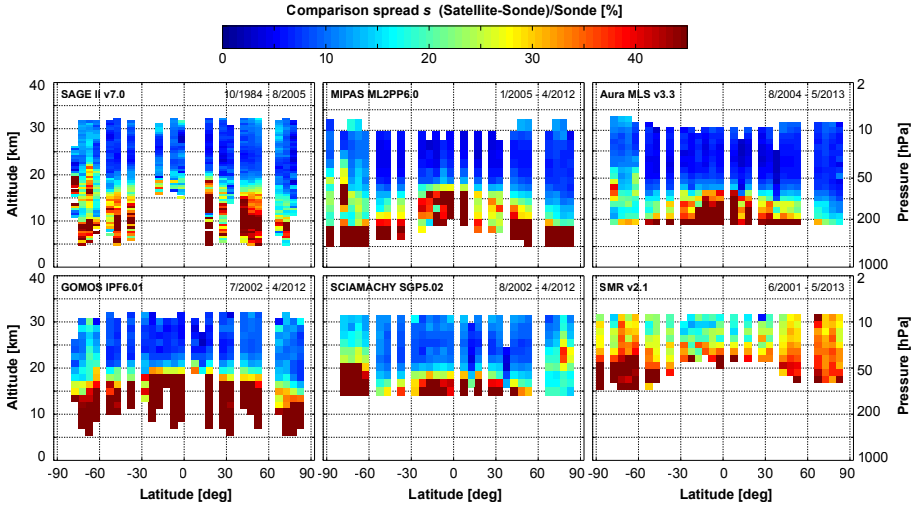


Figure 7. Meridional structure of the comparison spread *s* of six satellite records relative to the ozonesonde network. The analysis is performed in the native profile representation of each satellite record. Only bins with more than 10 comparison pairs are shown.

AMTD
8, 6661–6757, 2015

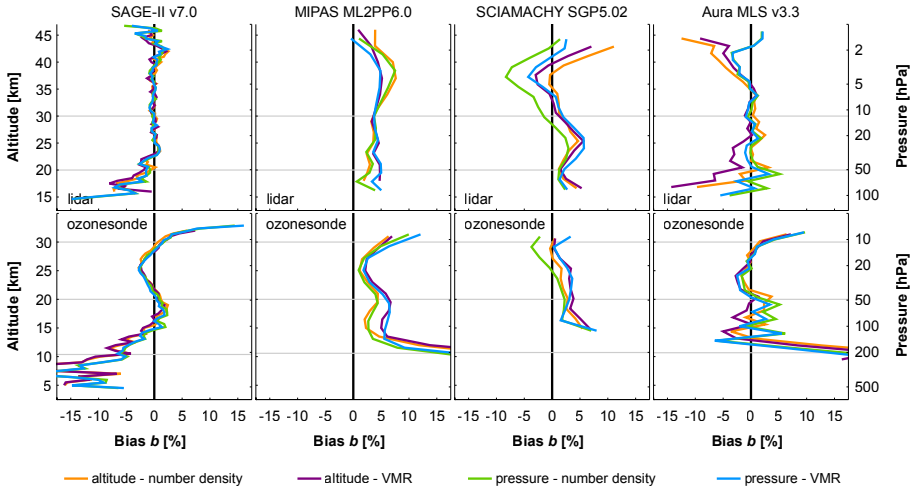
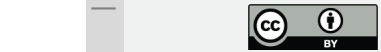


Figure 8. Dependence on profile representation of satellite bias relative to lidar (top row) and ozonesonde data (bottom) at northern mid-latitudes. The satellite profiles are converted using the auxiliary information provided in the respective satellite data files.

AMTD
8, 6661–6757, 2015



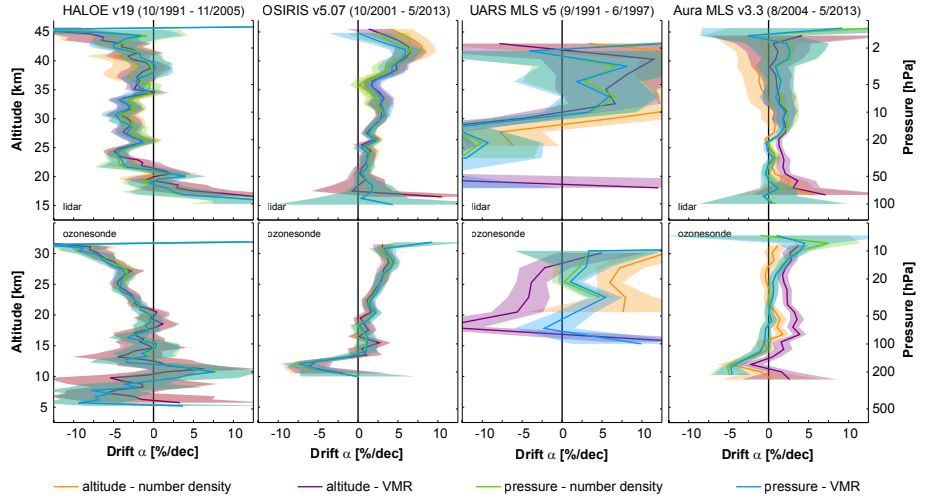
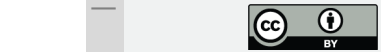


Figure 9. Dependence on profile representation of satellite drift relative to the lidar (top row) and ozonesonde network (bottom) network. The satellite profiles are converted using the auxiliary information provided in the respective satellite data files. Shaded area represents 1σ uncertainty.

AMTD
8, 6661–6757, 2015



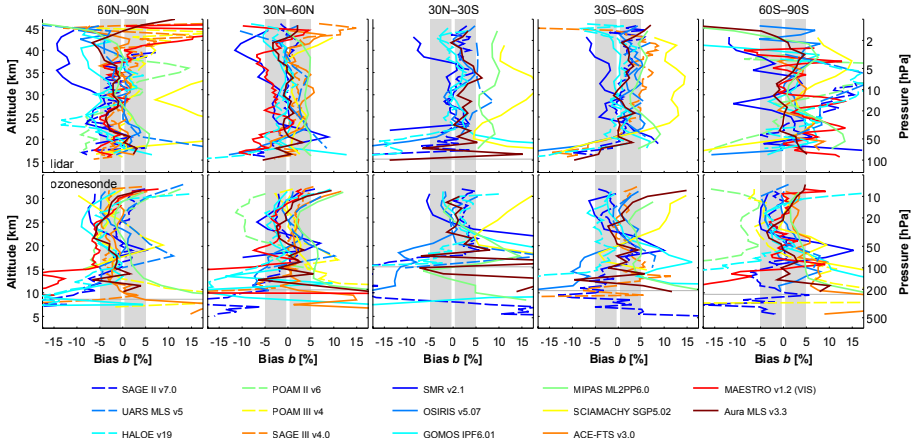


Figure 10. Overview of the bias *b* of all satellite ozone records relative to the lidar (top row) and ozonesonde (bottom) network in five latitude bands (columns). Dashed lines indicate instruments that ceased operations prior to 2006. Horizontal lines indicate the average altitude of the tropopause. The analysis is done in the native system of each satellite record. Most satellite records agree within ±5 % with the ground-based data in the middle and upper stratosphere.

AMTD
8, 6661–6757, 2015

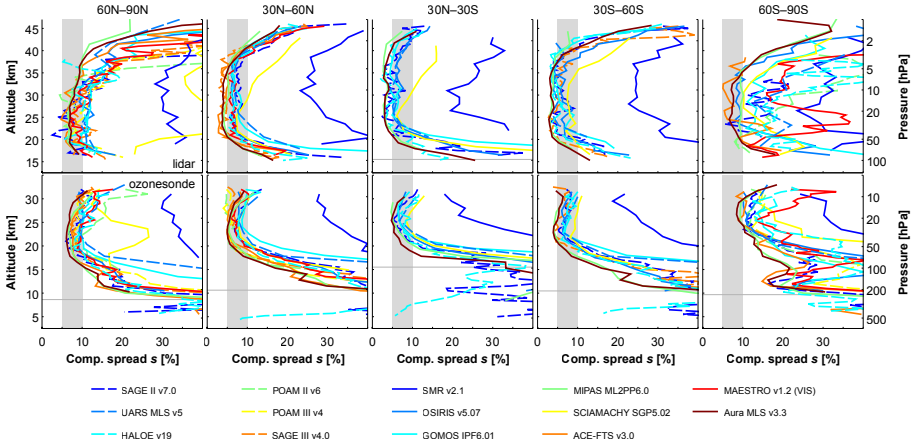


Figure 11. Similar as Fig. 10, but for the comparison spread *s*. The spread is mostly between 5–12 % in the middle and upper stratosphere.

AMTD
8, 6661–6757, 2015

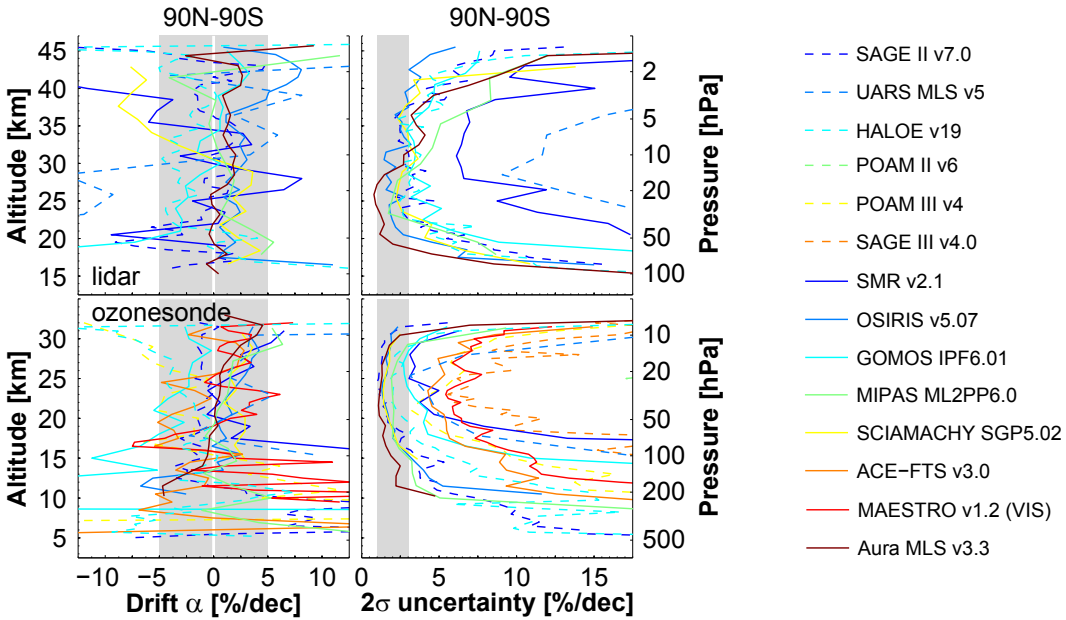


Figure 12. Overview of the drift α (left column) and its 2σ uncertainty (right) of all satellite ozone records relative to the lidar (top row) and ozonesonde (bottom) network. Dashed lines indicate instruments that ceased operations prior to 2006. The analysis is done in the native system of each satellite record. The decadal stability of most records remains within $\pm 5\,\%\,\mathrm{decade}^{-1}$ in the middle and upper stratosphere. POAM II and SAGE III drift results are not displayed in the lower left panel to avoid clutter.

AMTD
8, 6661–6757, 2015