Intercomparison and evaluation of ground- and satellite-based stratospheric ozone and temperature profiles above Observatoire de Haute-Provence during the Lidar Validation NDACC Experiment (LAVANDE)

Robin Wing, Wolfgang Steinbrecht, Sophie Godin-Beekmann, Thomas J. Mcgee, John T. Sullivan, Grant Sumnicht, Gérard Ancellet, Alain Hauchecorne, Sergey Khaykin, Philippe Keckhut

Atmos. Meas. Tech., 13, 5621–5642, 2020
https://doi.org/10.5194/amt-13-5621-2020

Atmospheric Measurement Techniques
Open Access
EGU

Intercomparison and evaluation of ground- and satellite-based stratospheric ozone and temperature profiles above Observatoire de Haute-Provence during the Lidar Validation NDACC Experiment (LAVANDE)

Robin Wing[1], Wolfgang Steinbrecht[2], Sophie Godin-Beekmann[1], Thomas J. McGee[3], John T. Sullivan[3], Grant Sumnicht[3], Gérard Ancellet[1], Alain Hauchecorne[1], Sergey Khaykin[1], and Philippe Keckhut[1]

[1]LATMOS/IPSL, OVSQ, Sorbonne Universités, CNRS, Paris, France
[2]Deutscher Wetterdienst, Met. Obs. Hohenpeissenberg, Hohenpeissenberg, Germany
[3]NASA Goddard Space Flight Center, Greenbelt, MD, USA

Correspondence: Robin Wing (robin.wing@latmos.ipsl.fr)

Received: 13 January 2020 – Discussion started: 14 April 2020
Revised: 8 July 2020 – Accepted: 11 September 2020 – Published: 22 October 2020

Abstract. A two-part intercomparison campaign was conducted at Observatoire de Haute-Provence (OHP) for the validation of lidar ozone and temperature profiles using the mobile NASA Stratospheric Ozone Lidar (NASA STROZ), satellite overpasses from the Microwave Limb Sounder (MLS), the Sounding of the Atmosphere using Broadband Emission Radiometry (SABER), meteorological radiosondes launched from Nîmes, and locally launched ozonesondes. All the data were submitted and compared "blind", before the group could see results from the other instruments. There was good agreement between all ozone measurements between 20 and 40 km, with differences of generally less than 5 % throughout this region. Below 20 km, SABER and MLS measured significantly more ozone than the lidars or ozonesondes. Temperatures for all lidars were in good agreement between 30 and 60 km, with differences on the order of ±1 to 3 K. Below 30 km, the OHP lidar operating at 532 nm has a significant cool bias due to contamination by aerosols. Systematic, altitude-varying bias up to ±5 K compared to the lidars was found for MLS at many altitudes. SABER temperature profiles are generally closer to the lidar profiles, with up 3 K negative bias near 50 km. Total uncertainty estimates for ozone and temperature appear to be realistic for nearly all systems. However, it does seem that the very low estimated uncertainties of lidars between 30 and 50 km, between 0.1 and 1 K, are not achieved during Lidar Validation Network for the Detection of Atmospheric Composition Change (NDACC) Experiment (LAVANDE). These estimates might have to be increased to 1 to 2 K.

1 Introduction

The international Network for the Detection of Atmospheric Composition Change (NDACC; http://www.ndacc.org, last access: 17 June 2018), formerly the Network for the Detection of Stratospheric Change (NDSC), is composed of more than 70 research stations worldwide (Kurylo et al., 2016; De Mazière et al., 2018). Ground-based remote sensing techniques measuring atmospheric parameters such as temperature and trace gas concentrations are used in NDACC to allow (1) early detection of long-term changes in the atmosphere; (2) validation of atmospheric measurements from satellites; (3) investigation of connections between atmospheric composition and climate change; and (4) support for testing and improving numerical computer models of the atmosphere.

Ground-based NDACC lidar stations have been providing routine long-term vertical profiles of stratospheric ozone and temperature since the mid-1980s (Steinbrecht et al., 2009a). One key lidar station is Observatoire de Haute-Provence (OHP) in southern France, situated at 43.94° N, 5.71° E, and 650 m above sea level (http://www.obs-hp.fr/geo/geo_ohp.shtml, last access: 12 June 2018). The first stratospheric

ozone measurements at OHP started in 1977 (Megie et al., 1977), with routine measurements since 1985 (Godin et al., 1989). Dedicated temperature lidars at OHP have been providing routine stratospheric and mesospheric temperature profiles since 1978 (Hauchecorne and Chanin, 1980). A lidar for tropospheric ozone has been operating routinely since 1990 (Ancellet and Beekmann, 1997).

NDACC requires standardised, consistent, high-quality, long-term measurements. Regular instrument and algorithm intercomparison campaigns are used to validate NDACC instruments and to track possible instrument biases. NDACC lidars, for example, have been intercompared in the 1989 Stratospheric Ozone Intercomparison Campaign in Table Mountain, California (STOIC; Margitan et al., 1995); the 1995 Ozone Profiler Assessment in Lauder, New Zealand (OPAL; McDermid et al., 1998); the 1997 OTOIC intercomparison in Haute-Provence (Braathen et al., 2004); the 1998 Ny-Ålesund Ozone Measurements Intercomparison in Spitsbergen, Norway (NAOMI; Steinbrecht et al., 1999); the 1999 DIfferential Absorption Lidar (DIAL) algorithm intercomparison campaign (Godin et al., 1999); the 2005 Hohenpeissenberg Ozone Profiling Experiment in Germany (HOPE; Steinbrecht et al., 2009b); and the 2009 Measurements of Humidity in the Atmosphere and Validation Experiments in Table Mountain, California (MOHAVE; Leblanc et al., 2011). Many of these campaigns have resulted in corrections and improvements for the involved lidar systems and their analysis software. A review of NDACC validation exercises was done by Keckhut et al. (2004). In general, the intercomparisons have shown that NDACC lidars can measure the stratospheric ozone profile with an accuracy better than 3 % between 12 and 35 km altitude and better than 10 % between 35 and 40 km. For temperature, NDACC lidars are typically precise to better than 1 K from 30 to 40 km altitude, with precision decreasing above to, e.g. 5 K near 70 km depending on the particular lidar station and integration time. These campaign findings are consistent with recent re-evaluations of theoretical uncertainty budgets by Leblanc et al. (2016a, b, c).

In addition to the NDACC campaigns which primarily focus on stratospheric ozone, there have been a few recent NDACC-like lidar intercomparisons for tropospheric ozone in the Tropospheric Ozone Lidar Network (TOLNet). The 2014 series of campaigns at five sites in the United States and Canada (DISCOVER-AQ and FRAPPÉ; Wang et al., 2017); the 2015 Langley Research Center (LaRC) Ozone Lidar intercomparison in Hampton, Virginia (LaRC; Sullivan et al., 2015); and the 2016 Southern California Ozone Observation Project (SCOOP; Leblanc et al., 2018). Tropospheric ozone concentrations from the ozonesondes regularly launched at OHP have been also frequently compared to the tropospheric ozone lidar data operated at the same site (Beekmann et al., 1995; Gaudel et al., 2015).

The purpose of the present paper is to report on the Lidar Validation NDACC Experiment (LAVANDE), which took place in July 2017 and March 2018 at Observatoire de Haute-Provence (OHP) in southern France. LAVANDE allows the comparison of the measured ozone profiles from the stationary differential absorption lidars for stratospheric (LiO_3S) and tropospheric ozone (LiO_3T) at OHP (Godin-Beekmann et al., 2003; Ancellet and Beekmann, 1997) with ozone profiles measured from the mobile trailer-based NDACC Stratospheric Ozone Lidar (NASA STROZ), operated by NASA's Goddard Space Flight Center (McGee et al., 1991). Additional comparisons are made with routine electrochemical cell (ECC) ozonesondes flown at OHP, and with satellite measurements by the Microwave Limb Sounder (MLS Aura; Waters et al., 2006) and the Sounding of the Atmosphere using Broadband Emission Radiometry instrument (SABER TIMED; Russell et al., 1999). Except for LiO_3T, all these instruments also provide temperature profiles over a substantial part of the stratosphere. The lidar temperature profiles taken during LAVANDE are derived from the non-absorbing 355 nm line of the two ozone lidars (LiO_3S and NASA STROZ) and from the dedicated stratospheric and mesospheric temperature Rayleigh lidar at OHP (Hauchecorne and Chanin, 1980), nowadays using a Nd:YAG laser at 532 nm. These temperature profiles are compared with the routine radiosondes from the nearby Météo-France station at Nîmes (43.86° N, 4.41° E; about 100 km west of the OHP station) and with routine stratospheric meteorological analyses from the US National Centers for Environmental Prediction (NCEP).

It is important to note that LAVANDE was a "blind" intercomparison. All the data were collected by an impartial referee (Wolfgang Steinbrecht), who was not involved in running the campaign. Data from each ground-based instrument were submitted "blind" to the referee, within days (or maximum weeks) after the measurement, and without seeing results from the other instruments. The referee also carried out all the comparison data analysis.

2 Instruments used for LAVANDE

Table 1 summarises all the different systems participating in the LAVANDE intercomparison. Ozone profiles taken by the Stratospheric Aerosol and Gases Experiment III (SAGE-III) satellite instrument aboard the International Space Station (ISS) (Mauldin et al., 1998) in solar or lunar occultation geometry were also considered for the LAVANDE intercomparison. However, the number of reasonably coincident SAGE-III profiles turned out to be too low for statistically meaningful results (only three or four profiles). Therefore, SAGE-III ISS profiles are not included here.

In addition to Table 1, each instrument in the intercomparison campaign is described briefly below. Key aspects are noted in each subsection. References to original or most recent instrument descriptions are given for those seeking further details.

Table 1. Instruments compared during the LAVANDE campaign in July 2017 and March 2018.

Instrument	Measurement of ozone	Altitude range	Measurement of temperature	Altitude range	Data source
NASA STROZ	DIAL (308 and 355 nm)	10 to 50 km	Rayleigh and Raman lidar (355 nm)	10 to 70 km	1
LiO_3S	DIAL (308 and 355 nm)	10 to 50 km	Rayleigh lidar (355 nm)	25 to 60 km	1
LiO_3T	DIAL (289 and 316 nm)	2.5 to 13 km	–	–	1
LTA	–	–	Rayleigh lidar (532 nm)	30 to 80 km	1
OHP ECC sondes (ENSCI-Z)	KI electrochemical cell	0 to 35 km	Thermistor (Modem M10)	0 to 35 km	1
Nîmes radiosondes	–	–	Thermistor (Modem M10)	0 to 35 km	2
NCEP analyses	–	–	Meteorological data assimilation	0 to 50 km	3
MLS satellite, version 4.23	μwave limb sounding (240 GHz)	10 to 80 km	μwave limb sounding (118 GHz)	15 to 90 km	4
SABER satellite, version 2.0	IR limb sounding (9.6, 1.27 µm)	15 to 90 km	IR limb sounding (4.3, 15 µm)	10 to 100 km	5

[1] ftp://ftp.cpc.ncep.noaa.gov/ndacc/station/ohp/ (last access: 17 June 2018), [2] http://weather.uwyo.edu/upperair/sounding.html (last access: 10 June 2018),
[3] ftp://ftp.cpc.ncep.noaa.gov/ndacc/ncep/temp/ (last access: 1 June 2018), [4] https://disc.gsfc.nasa.gov/datasets?keywords=MLS (last access: 17 June 2018), and
[5] SABER temperature and ozone profiles are available at http://saber.gats-inc.com/ (last access: 17 June 2018).

2.1 Lidars

2.1.1 OHP stratospheric lidar (LiO_3S)

The stratospheric ozone lidar (LiO_3S) is a differential absorption lidar which relies on the difference in the absorption cross-section for ozone at two different wavelengths. The DIAL technique infers the ozone number density by taking the derivative of the ratio between a strongly absorbed line (online) and a weakly absorbed line (non-absorbed) (Pelon et al., 1986). The system at OHP has two lasers emitting in the ultraviolet at 308 nm (online) and at 355 nm (offline), a constellation of four receiver telescopes, and a Horiba Jobin Yvon holographic grating for line selection, described in Godin-Beekmann et al. (2003). In addition to making measurements of ozone, the offline of a DIAL system (355 nm) can be used to calculate Rayleigh temperature (Hauchecorne and Chanin, 1980). The LAVANDE campaign represents the first attempt to validate LiO_3S temperature profiles within the framework of NDACC. The comparisons made during this campaign will prove vital for the assessment of the temperature combined uncertainty budget. Measurements with this instrument have been ongoing since 1985 and to date amount to 3678 nights of data. Further details can be found for ozone profile retrieval, error analysis, and vertical resolution determination in Godin-Beekmann et al. (2003) and for temperature profile retrieval in Wing et al. (2018a).

2.1.2 OHP tropospheric lidar (LiO_3T)

The tropospheric ozone lidar (LiO_3T) is also a DIAL system; however, it differs from its stratospheric counterpart in a few key ways. The tropospheric DIAL system does not rely on two separate lasers to generate the absorbed and non-absorbed wavelengths. The laser source is a Nd:YAG laser fourth harmonic emission at 266 nm. Two additional wavelengths are generated from the original 266 nm beam at 289 and 316 nm through a process known as stimulated Raman scattering in a high-pressure deuterium cell. Using this Raman technique allows for the tropospheric lidar to measure much lower tropospheric ozone concentrations (on the order of ppb rather than ppm) as compared to the stratospheric system. Further details of this technique can be found in Papayannis et al. (1990) and Milton et al. (1998). Both photocounting and analogue detection are applied to provide vertical profiles in the altitude range of 2.5–15 km (Ancellet and Beekmann, 1997). The tropospheric ozone lidar has made continuous twice-weekly measurements since 1990 (Gaudel et al., 2015).

2.1.3 OHP Lidar Température et Aérosols (LTA)

The Lidar Température et Aérosols (LTA) is a classic Rayleigh–Mie–Raman lidar operating at 532 nm (Keckhut et al., 1993). The absolute temperature profile is directly derived from the range-square corrected lidar return signal (Hauchecorne and Chanin, 1980). The system employs a high-powered laser transmitter and a constellation of four receiver telescopes. It has been making regular measurements since 1978. Further details about this instrument, algorithm details, and the most recent technical specifications can be found in Wing et al. (2018a).

2.1.4 NASA STROZ

NASA's Goddard Space Flight Center STROZ is a mobile validation lidar which is shipped across the world on a regular basis to run intercomparison and validation campaigns with ozone and temperature lidars in NDACC. NASA STROZ is a DIAL system similar to the LiO_3S, relying on an online wavelength of 308 nm and an offline wavelength of 355 nm. The system was originally constructed in 1988 (McGee et al., 1991) and has been used as a reference during campaigns for multiple lidar stations since then (McGee et al., 1995).

2.1.5 Radiosondes and ozonesondes (ECC)

ECC ozonesondes manufactured by ENSCI-Z filled with 1 % of potassium iodide (KI) and coupled to Meteomodem M10 radiosondes were launched every two nights during the first phase of the campaign in July 2017 and nightly during the second phase of the campaign in March 2018. The sondes and balloons were prepared and launched by the same OHP technicians responsible for the weekly ozonesonde launch. The OHP radiosonde programme is homogenised under the auspices of NDACC France ozone measurements. A new publication describing the full data treatment details, quality metrics, and uncertainty budget estimates is envisioned for 2021.

The campaign ECCs reached a median burst altitude of 32.7 km with only one balloon bursting early at 17 km. Below 21 km, in the first phase of the campaign, the sondes flew north at the beginning of July, west near the middle of the month, and south by the end of the month. Above 21 km, all the 2017 sondes were carried east by the prevailing summer stratospheric wind. During the second phase of the campaign, the sondes flew generally north with only slight westerly changes in trajectory as they ascended. ECC ozonesondes provide a precision of ±3 %–5 % and an accuracy of ±5 %–10 % (Smit, 2013; Tarasick et al., 2016). A known positive bias of the ENSCI ECC data in the troposphere when using 1 % KI concentration (Smit et al., 2007) is corrected by decreasing the ECC ozone concentration by 4 % below the tropopause (Gaudel et al., 2015). Weekly ECC launches have been conducted at OHP since 1991 and a quality control factor (qcf) is calculated using a normalisation of the total ozone from the sonde to the total ozone measured by a Système D'Analyse par Observations Zénithales (SAOZ) spectrophotometer at OHP (Nair et al., 2012; Gaudel et al., 2015). The ECC data are discarded if the calculated qcf is outside the range of 0.8–1.2. The control factor is not applied to the ECC data and the measured ozone partial pressures are not corrected above the tropopause. During the LAVANDE campaign, the control factor is always in the range of 0.92–1.05 except for on 20 March when qcf is 1.16.

In addition to the ECCs, we also used the Meteomodem M10 meteorological radiosondes launched twice daily from the nearby station at Nîmes.

2.2 Co-located satellite overpasses

The satellite-based MLS and SABER instruments provide stratospheric ozone and temperature profiles over most of the globe.

2.2.1 Microwave Limb Sounder (MLS)

MLS is a spectrometer aboard the Aura satellite which measures thermal microwave radiation from the atmosphere in limb geometry and allows retrieval of stratospheric ozone profiles with a vertical resolution of about 3 km and retrieval of stratospheric temperature profiles with a typical vertical resolution of 8 km at 30 km altitude, 9 km at 45 km altitude, and 14 km at 80 km (full width at half maximum (FWHM) of the averaging kernels; Schwartz et al., 2008). We have used version 4.0 MLS profiles of temperature, geopotential height, and ozone. For comparison with the ground-based lidars and ozonesondes, the geopotential altitude is converted to a geometric altitude. A more complete description of the instrument is given in Waters et al. (2006).

2.2.2 Sounding of the Atmosphere using Broadband Emission Radiometry (SABER)

The SABER instrument aboard the TIMED satellite makes ozone and temperature measurements from about 15 to 100 km. For temperature, it provides a vertical resolution of 2 km and temperature accuracy of 1 to 2 K between 15 and 60 km, decreasing to 5 K near 85 km, and to 10 K near 100 km (Rezac et al., 2015a, b). For ozone, SABER provides 1 % precision between 40 and 50 km altitude, decreasing to 2 % near 30 and 55 km and to 10 % near 15 and 80 km (Rong et al., 2009). We have used version 2.0 SABER profiles of temperature and ozone. A more complete description of the instrument is given in Mertens et al. (2001).

2.2.3 Co-locating satellite profiles and ground-based profiles

While all the lidars were measuring at the same location and the same time during LAVANDE, and the ECC sondes were quite close in time and space, satellite profiles almost never match the exact time and location of a ground-based measurement. For LAVANDE, we considered all satellite profiles with a tangent point within ±5° latitude and ±15° longitude of the OHP station (43.94° N, 5.71° E) and within ±12 h of 00:00 UTC (1 h after local midnight for the lidar measurements nights) (see also Wing et al., 2018b). This fairly large coincidence box is depicted in Fig. 1. It covers most of southern Europe, from Paris in the north to the southern tips of Spain or Sardinia in the south, and from Portugal in the west to Slovakia, Hungary, or Serbia in the east. The size of the chosen box size is a matter of compromise. On the one hand, a small coincidence box results in very few coinciding satellite profiles but also very close matches in time and space between satellite and ground-based profiles. On the other hand, a large box results in many coinciding satellite profiles but poor matches in space and time. The box size chosen here is similar to the compromise chosen in Wing et al. (2018b). It results in between 10 to 20 coincident profiles for MLS and SABER, which are generally divided between one or two satellite overpasses, for a given night during the LAVANDE campaign.

The question of which of these 10 to 20 profiles should be used for the intercomparison then arises. One choice would

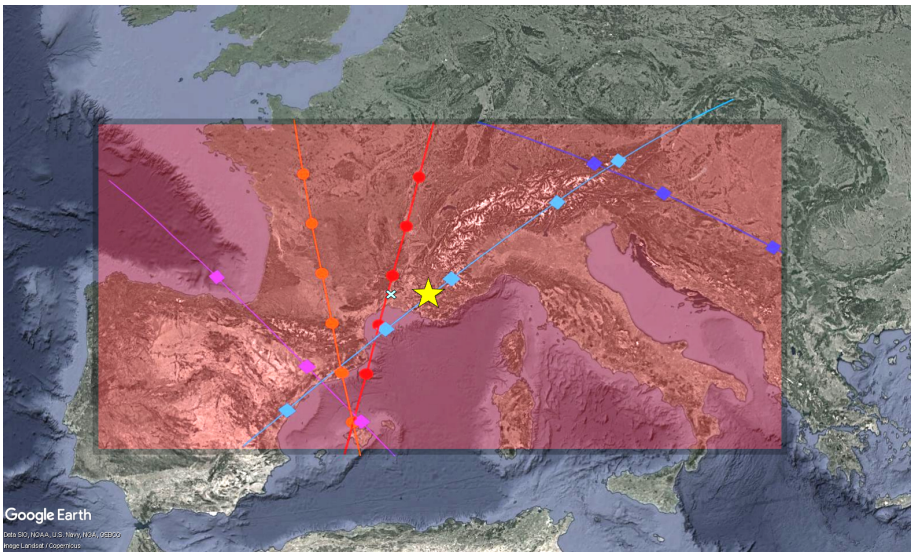

Figure 1. The area defined for coincident measurements during the LAVANDE campaign (39,−9) to (49,21). Observatoire de Haute-Provence is represented by the yellow star at (43.93,5.71) and Nîmes radiosonde launches by a cyan X at (43.86,4.41). Ascending (red) and descending (orange) orbits for MLS with tangent point locations of profiles for 17 July 2018. Ascending (light blue) and descending (purple and dark blue) orbits for SABER with tangent point locations of profiles for 17 July 2018 (data: © Google Earth Pro, 2019).

be to take the profile that matches most closely in space and time. Another choice would be to use the average profile obtained from all satellite profiles in the coincidence box. A third possibility is to use the weighted average profile, with lower weight given to satellite profiles that are further away in space or time. We used weights proportional to one over the $\sqrt{(\Delta r^2 + (v \cdot \Delta t)^2)}$, where Δr and Δt are the distance in space and time between the lidar profile and the satellite profile, and $v = 10\,\mathrm{m\,s^{-1}}$ is a wind speed typical for the mid-stratosphere. For the LAVANDE intercomparison, we tested these three possible profile choices. Generally, differences between all three choices were quite small. Overall, however, the weighted average profile gave slightly better results than the others. Therefore, the weighted average MLS and SABER profiles are used throughout most of this paper. The same three techniques were applied to the associated measurement uncertainty profiles to produce the nightly average measurement uncertainty profile (hereafter referred to simply as the "measurement uncertainty"). In practice, these three versions of the measurement uncertainty profiles were nearly identical, showing that the statistical uncertainty on the measurement uncertainty is extremely low.

3 Campaign overview

The LAVANDE campaign took place in two parts: the first period covered about 2 weeks in summer 2017, from 10 to 26 July, and the second period covered 10 d in early spring 2018, from 12 to 22 March. Table 2 shows which ground-based systems provided ozone and/or temperature profiles on each of the different nights of the campaign. Temperature profiles from NCEP reanalysis were included as well. Overall, LAVANDE covered about 4 weeks of measurements and provided ≈ 120 ground-based temperature profiles and ≈ 60 ground-based ozone profiles. Due to a laser failure in the NASA STROZ system, that system was not able to measure ozone profiles after 18 July in 2017. Temperature measurements, however, were still possible and were not affected. The NASA STROZ laser was repaired by March 2018 for the second phase of the campaign. All other systems were operating nominally throughout the campaign with no significant problems. The MLS and SABER satellite instruments provided ozone and temperature profiles during all campaign nights, in the spatial and temporal coincidence box introduced in Fig. 1.

3.1 Example comparisons

Two examples for both ozone and temperature profiles for a LAVANDE night in July 2017 and March 2018 are given in Figs. 2 and 3. We can see the high degree of fidelity in reproducing the ozone profile across all ground-based instruments. In particular, we see very good agreement of the small-scale features present below 15 km in the July example. In Fig. 2, we see that the ozone number density is fairly low throughout the troposphere, about $1 \times 10^{12}\,\mathrm{cm^{-3}}$, slightly declining up to the tropopause at about 13 to 15 km. Above the tropopause, ozone increases substantially up to the number density maximum, located at about 25 km altitude in July 2017 and about 19 km in March 2018. In the left-hand panel, above the ozone maximum, ozone decreases

Table 2. Measurement dates for the ground-based instruments during the LAVANDE campaign in July 2017 and March 2018. The lidar measurements require nighttime conditions and averaging over several hours. The dates give the beginning of these nights. X denotes a valid measurement for the given night. (x) denotes a measurement that appeared faulty and was not used in the later statistical analysis. Satellite profiles of ozone and temperature are available for all nights.

Start	Ozone		Ozone and temp.			Temperature only				
of night	OHP tropo.	NASA[a] STROZ	OHP[b] DIAL	ECC[b] sonde	Total O_3	NASA[a] temp.	OHP temp.	Nîmes sonde	NCEP[c] anal.	Total temp.
10 July	X		X		2		X	X	X	4
11 July		X	X	X	3	X	X	X	X	6
12 July	X	(x)	X	X	3	(x)	X	X	X	5
13 July		X	X		2	X	X	X	X	5
14 July	X	X	X	X	4	X		X	X	5
15 July		X	X		2	X	X	X	X	5
16 July	X	X	X	X	4	X	X	X	X	6
17 July	X	X	X		3	X	X	X	X	5
18 July	X		X		2		X		X	3
19 July					0			X	X	2
20 July	X		X	X	3		X	X	X	5
21 July			X		1	X	X	X	X	5
22 July	X		X	X	3	X	X	X	X	6
23 July					0			X	X	2
24 July			X		1	X	X	X	X	5
25 July	X		X		2	X	X	X	X	5
26 July	X		X	X	3	(x)	X	X	X	5
Subtotal	10	6	15	7	38	10	14	16	17	79
12 March			X		1		X	X	X	4
13 March			X	X	2			X	X	4
14 March					0			X	X	2
15 March					0			X	X	2
16 March					0			X	X	2
17 March		X	X		2	X	X	X	X	5
18 March					0			X	X	2
19 March	X	X	X	X	4	X	X	X	X	6
20 March	X	X	X[d]	(x)	3	X	X	X	X	5
21 March	X	X	X	X	4	X	X	X	X	6
22 March	X	X	X	X	4	X	X	X	X	5
Subtotal	4	5	7	4	20	5	6	11	11	43
Grand total	14	10	22	11	58	15	20	27	28	122

[a] Due to a laser failure on 18 July, the NASA STROZ system was not able to measure ozone profiles for the rest of July 2017. Temperature measurements were still possible and a separate column was included for temperature profiles from the NASA system. [b] The LiO_3S system and the ECC sondes measure both ozone and temperature profiles. [c] NCEP analyses usually provide data for 12:00 UTC. For comparison with the nightly mean lidar profiles (typically around 20:30 UTC), we used the average of the two 12:00 UTC analyses before and after each night. [d] The LiO_3S temperature profile was clearly faulty on that night, but the ozone profile appeared to be fine.

steadily with altitude, from about $4 \times 10^{12}\,\mathrm{cm}^{-3}$ near 25 km to less than $1 \times 10^{12}\,\mathrm{cm}^{-3}$ near 50 km. In the right-hand panel, we see much more variation in the upper troposphere and lower stratosphere (UTLS) which is consistent with the more dynamically variable spring at OHP. Additionally, the March ozone maximum is greater and lower in altitude, about $7 \times 10^{12}\,\mathrm{cm}^{-3}$ at 18 km. In general, the ozone profiles have less vertical structure and are smoother above 25 km. It is important to note that the lower stratospheric ozone is much more variable in the springtime (left panel) than in the summer in response to seasonal dynamics. This increased variability introduces an added layer of complexity to our analysis and must be accounted for carefully.

In order to compare the ozone profiles from the different systems, it is necessary to put the data on a common altitude grid. For LAVANDE, a vertical grid with 300 m spacing was chosen. Data with finer vertical spacing (lidars and sondes) were averaged to 300 m wide altitude bins centred around

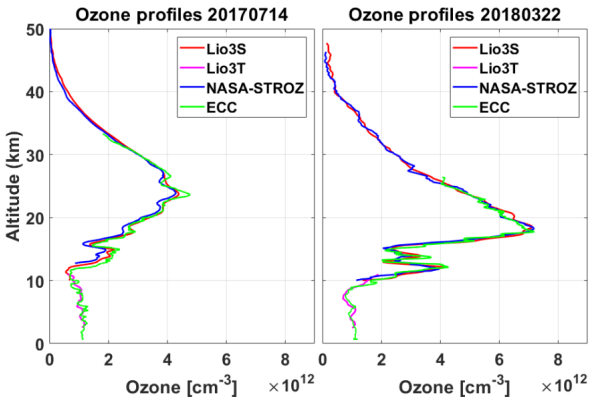

Figure 2. Ozone profiles measured by the different instruments at Observatoire de Haute-Provence on the nights of 14–15 July 2017 and 22–23 March 2018. Note the seasonal differences in the character of the ozone profiles in spring and summer.

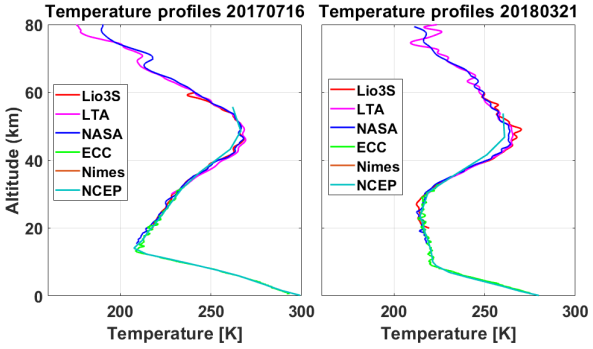

Figure 3. Temperature profiles measured by the different instruments at Observatoire de Haute-Provence on the nights of 14–15 July 2017 and 22–23 March 2018. Note the seasonal differences in the character of the ozone profiles in spring and summer.

the midpoints of this grid. Data with coarser vertical spacing (satellites and NCEP) were interpolated to the 300 m grid. In the troposphere and lower stratosphere up to about 25 km, the conversion to the 300 m vertical grid smooths out some of the finer structures present in the original lidar data, whereas at higher altitudes the differences between the original data and the data on the 300 m grid are small. For most instruments, the lack of finer structures above 30 km is due to limited vertical resolution of the original retrieved profiles.

The temperature profiles in Fig. 3 are for the same night from July 2017 and March 2018 and show the usual temperature decline throughout the troposphere. During nighttime in July, the tropopause is located at about 13 km altitude and around 10 km in March. Above the tropopause, the temperature increases with altitude up to the stratopause at 45 to 50 km. There is a distinct difference in the temperature lapse rate of the lower stratosphere in the spring (right panel) as the atmosphere is nearly isothermal until 30 km. The increased spring time variance in the lower stratospheric temperatures

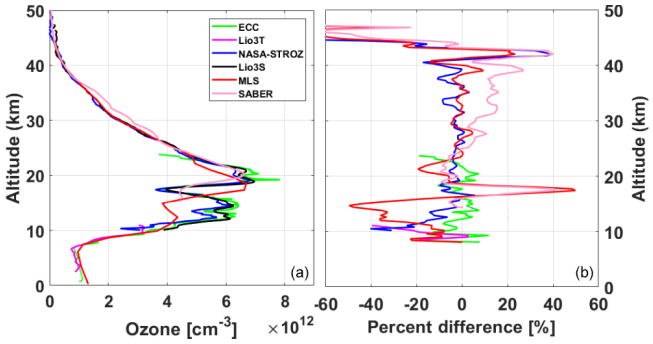

Figure 4. Satellite and lidar ozone profiles measured on the night of 19–20 March 2018 at or near Observatoire de Haute-Provence **(a)**. Percent differences for each profile with respect to the LiO_3S profile **(b)**. All profiles have been converted to the same 300 m vertical spacing altitude grid. For MLS and SABER, the weighted average profile is calculated based on the distance in time and space between the individual satellite profiles and the OHP station.

should be considered when conducting lidar validation studies. In the mesosphere, from 50 to 80 or 90 km, temperatures decrease again with altitude. Temperature profiles measured by all systems in Fig. 3 show these features with good consistency between systems over a wide altitude range. As with the ozone profiles in Fig. 2, conversion to the regular 300 m altitude grid smooths out finer structures at lower altitudes. For temperature, the highest vertical resolution data, down to a few metres, come from the radiosondes coupled to the ECC ozone sensors. Lidar temperatures have vertical resolution of 150 m in the lower stratosphere to greater than 1 km in the mesosphere. The other systems have vertical resolutions which are generally coarser than 1 km.

3.2 Comparisons with satellites

Figure 4 (ozone density) and Fig. 5 (temperature) give examples from the second part of LAVANDE in March 2018 and also include MLS and SABER satellite data. There is generally good agreement between all instruments for both ozone and temperature profiles; all instruments show similar ozone profiles with the ozone maximum occurring near 20 km. The ground-based measurements also reproduce the fine-scale ozone features as narrow as 150 m in vertical extent over a wide range of altitudes. All instruments correctly identify the tropopause and stratopause at same altitudes and amplitudes, to within 5 K.

In the left panel of Fig. 4, we present a case with less than 10 % difference (with the exception of MLS below 20 km) between ozone profiles measured by the lidars and the satellites. In the right panel, the percent difference for each profile is shown with respect to the LiO_3S profile. We can see that MLS and LiO_3T agree fairly well between 5 and 11 km, following the same trend of ozone increasing with altitude. The agreement between all measurements from 20 to 40 km

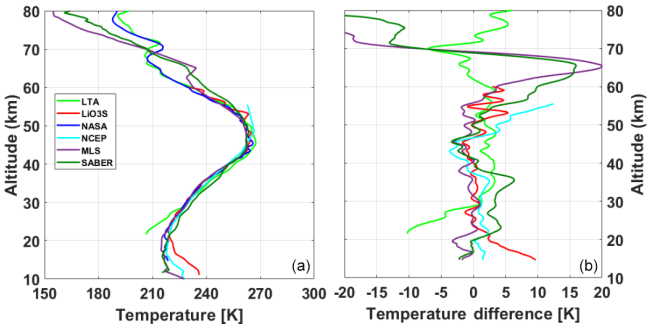

Figure 5. Satellite, NCEP, and lidar temperature profiles measured on the night of 24–25 July 2017 at or near Observatoire de Haute-Provence **(a)** and temperature difference profiles with respect to the NASA temperature profile **(b)**. All profiles have been converted to the same 300 m vertical spacing altitude grid. For MLS and SABER, the weighted average profile is calculated based on the distance in time and space between the individual satellite profiles and the OHP station.

is good, with percent differences less than 20 %. Of particular interest is the region of disagreement between 11 and 20 km, characterised by rapid variation and spikes in the percent difference plot, where differences in spatiotemporal coincidence and atmospheric variability can lead to the sampling of different air masses.

In most cases, SABER ozone does not agree with ozone measurements from the other instruments below 25 km as it is principally an instrument focused on the upper middle atmosphere. The extent of the disagreement can be an order of magnitude larger than the differences between the ozone concentration measured by the other instruments. We will revisit this topic later in the article when discussing the ensemble ozone differences in Fig. 7. Presented in Fig. 4 is our best SABER comparison where we can see good agreement between SABER and the lidar. SABER tends to report slightly higher ozone number densities above 30 km than other measurements.

One key point to keep in mind when interpreting the right panel of Fig. 4 is that in regions on either side of the ozone maximum, where ozone densities are low, the percentage differences can be quite large but only represent slight differences in the number density.

In the right panel of Fig. 5, the temperature differences are plotted for each profile with respect to the NASA lidar temperature. We can see that all instruments agree fairly well with the NASA lidar up to 60 km with disagreements in the mesosphere. The deviation of the LTA temperature profile from the NASA temperature profile below 30 km is a known cooling effect of the differential absorption of laser light by aerosols in the visible and UV. The 532 nm LTA system is more strongly influenced by stratospheric aerosols than the 355 nm NASA lidar and LiO_3S systems. There is a warm bias in LiO_3S below 20 km. As the primary purpose of LiO_3S is the measurement of stratospheric ozone, the temperature retrievals, particularly those in the troposphere, are a value-added product of this system. The temperature measurements in the stratosphere compare very well with those of the other instruments, and with the addition of a new Raman channel, and a new comprehensive temperature retrieval package, it is anticipated that the warm bias evident below 20 km in Fig. 5 will be reduced.

Of particular interest is a small developing mesospheric inversion layer present near 71 km which is seen by both the NASA and LTA lidars. MLS displays an evident kink in the temperature profile at 65 km which could be the signal of the inversion layer given that the satellite has an effective vertical resolution of nearly 15 km at those altitudes. SABER does not detect the layer on this night but does track the development of the feature over the next few nights.

4 Intercomparison results for ozone

Figure 6 shows the time series of ozone concentrations measured by the different systems for a number of selected levels. A clear separation can be seen between the two measurement periods in July 2017 and March 2018, due to the normal seasonal cycle. Ozone values in the lower stratosphere (below about 25 km) were higher in March 2018 than in July 2017. In the upper stratosphere (above 30 km), in contrast, ozone values were lower in March 2018. In addition, atmospheric conditions (and ozone values) were much more variable in March 2018. Generally, all instruments track ozone variations in a similar way. However, Fig. 6 does indicate some systematic deviations. For instance, the NASA STROZ lidar tends to report lower ozone values near 40 km, while LiO_3S reports higher ozone concentrations than MLS, and SABER tends to report more ozone at lower levels.

A closer look at the systematic differences in the ozone profiles produced by each instrument, as well as their statistical uncertainty, is given in Fig. 7. This figure shows the average relative difference profile between ozone from the various instruments and ozone from the LiO_3S. The LiO_3S was chosen here as a reference, because it had the most measurement nights of all ozone systems (due to the unfortunate laser failure of NASA STROZ in July 2017). Similar to the results of previous NDACC intercomparisons (see the introduction), the best agreement between the different ozone systems is found between 20 and 40 km altitude. During LAVANDE, agreement over most of this altitude range was better than ± 5 % between most systems, with no statistically significant differences at 2σ (95 % confidence level). SABER measured some larger and more significant differences up to ± 10 % at some altitudes. Above 30 km, the ECC sondes measured slightly lower ozone concentrations than the other instruments by up to -10 %.

Below 20 km and above 40 km, the ozone concentration profiles from the different systems show larger deviations.

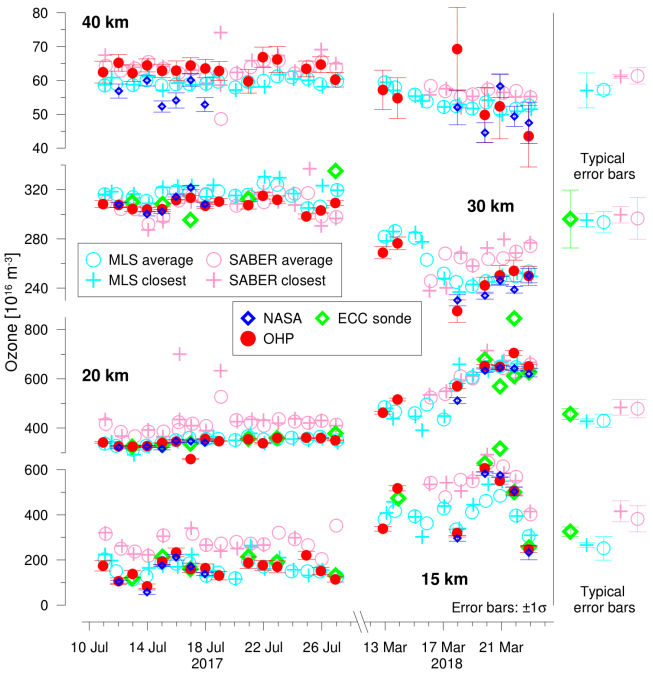

Figure 6. Time series of ozone concentrations measured at different altitude levels during LAVANDE.

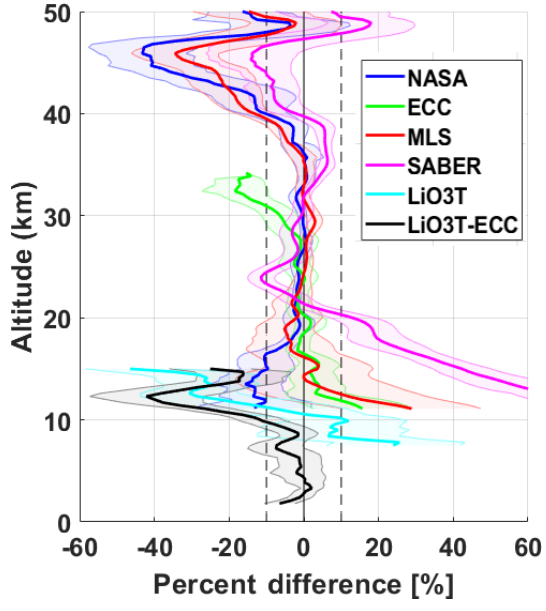

Figure 7. The average relative difference profile between the ozone profiles measured by the various LAVANDE instruments compared to the ozone profile measured by the LiO_3S. The shaded range gives the ± 2 standard deviations of the mean and indicates the statistical confidence interval at the 95 % uncertainty level. Results for MLS and SABER are reported using the weighted average profiles, but very similar results are obtained by using only the profile from the closest SABER or MLS overpass.

Around 45 km, for example, NASA STROZ, MLS, and SABER give 40 %, 30 %, and 15 % lower ozone values, respectively, than the LiO_3S system. These differences are statistically significant for at the 2σ level. Differences of this kind can be caused by the specific differential filter used at high altitudes above 40 km in the LiO_3S and NASA STROZ retrieval software (see also Godin et al., 1999). The heavier smoothing and integration is required above 40 km due to the drop in the lidar signal-to-noise ratio.

Below 20 km, SABER reports significantly higher ozone than the other systems. MLS also tends to report higher ozone, with differences up +20 % near 12 km, compared to the LiO_3S. However, this is not statistically significant at the 2σ level. The ECC sondes tend to report up +5 % higher ozone than the LiO_3S between 10 and 15 km, whereas NASA STROZ tends to report less ozone, −12 % on average near 10 km. These ECC and NASA STROZ differences are also not statistically significant at 2σ above 15 km. Finally, Fig. 7 indicates that the LiO_3T was in good agreement with the ECC sondes and the OHP stratospheric DIAL below 9 km, when the ECC sondes are corrected by the 4 % in the troposphere. This differences increase above 9 km to a maximum of −40 % near 14 km. The large percent difference between LiO_3T and LiO_3S between 10 and 15 km is unsurprising as both instruments are operating near their detection range limits (low signal-to-noise ratio and vertical averaging larger than 1 km for LiO_3T and large sensitivity to systematic errors for the LiO_3S near 10 km).

Another way of viewing the differences between the ozone profiles measured by the different instruments is to use scatter plots of ozone concentration as function of altitude (seen in Fig. 8). To plot the scatter between datasets, we further integrated the ozone profiles to 2 km resolution to reduce the high-frequency components. The three panels show generally good tracking of ozone measured by each of the different instruments against ozone measured by the LiO_3S, over a substantial range of ozone concentration values. Some of the systematic differences appearing in Fig. 4 can further examined in the scatter plots. One prominent example is the sharp onset of a high ozone concentration bias in SABER data below 20 km with respect to the other instruments. Looking at Fig. 8a, which represents the ozone concentration in the UTLS (0 to 20 km), we can see that the SABER (magenta) bias occurs most strongly at the lowest ozone concentrations. SABER profiles appear to have a lower ozone concentration limit of 2 to 3×10^{12} cm^{-3} and cannot match other instruments measuring below 2×10^{12} cm^{-3}. We can also examine the behaviour of the MLS bias in Fig. 7 which abruptly changed from positive below 25 km to negative below 15 km. Again, we can see in Fig. 8a that the sharp change occurs at very low ozone concentrations. For concentrations above 1×10^{12} cm^{-3}, MLS has a low bias with respect to all other instruments; however, below 1×10^{12} cm^{-3}, the variance abruptly increases with the majority of points exhibiting a high bias. These satellite–lidar biases in measured ozone

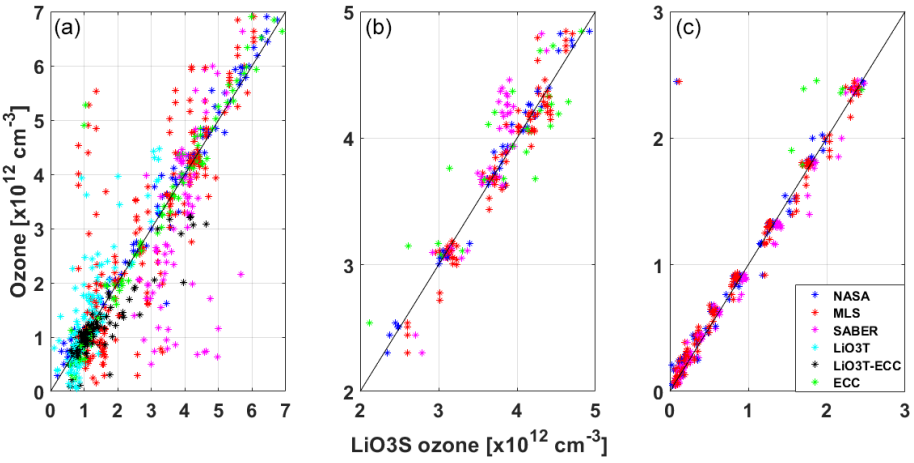

Figure 8. Scatter plots of ozone concentration as measured by the various LAVANDE instruments (along the vertical axis) and ozone measured by the LiO_3S (along the horizontal axis). **(a)** Ozone from 0 to 20 km altitude. **(b)** Ozone from 20 to 30 km altitude. **(c)** Ozone from 30 to 50 km altitude.

concentration are a convolution of an unknown real ozone bias, a bias arising from sampling different air, and a bias arising from the vertical resolution and smoothing of the satellites.

Figure 8b shows the scatter between ozone measurements in the region between 20 and 30 km (nominally near the altitude of the ozone maximum). We can see five tight clusters of data points which correspond to data points every 2 km. It is important to note that the real differences in the ozone concentration at these altitudes is low, so we have a very low variance associated with each cluster of points. Figure 8c shows the tracking of ozone concentrations from 30 to 50 km and much like Fig. 8b can be characterised by low variability and low variance. It is important to note that neither MLS nor SABER exhibit strong biases at these altitudes. Also, note that the comparison between the ECC and LiO_3T (black) is only present in Fig. 8a as the upper limit of the tropospheric lidar is around 12 to 15 km.

A complementary method for tracking the "goodness" of the match between the various LAVANDE instruments is presented in Fig. 9. It shows vertical profiles of the correlation between ozone from the each of the instruments and ozone from the LiO_3S. These correlations are taken using data from all LAVANDE nights (except outliers indicated in Table 2) which have been integrated to 2 km in an effort to filter out the high-frequency components. Figure 9 shows very high correlation between ozone concentration profiles measured by the LiO_3S and by NASA STROZ (blue line) and between LiO_3S and the ECC below 20 km (green line). Over much of the 10 to 35 km altitude range, the correlations exceed 0.95 between the two stratospheric ozone lidars. A slightly surprising feature in Fig. 9 is the marked drop in correlation around 25 km near the maximum of the ozone concentration. This drop is due to the relatively low variability of real ozone in both time and altitude, as was demonstrated in Fig. 8b. When the covariance of the data, arising from real differences in ozone concentration, drops faster than the variance of the data, in part arising from statistical scatter, we see a resulting drop in the correlation. As a result, the drop occurs at altitudes where the combined sampling and instrumental uncertainty of each instrument play a larger role in the correlation than true variations in ozone. Rather unsurprisingly, this effect is most noticeable in the comparisons between the lidars and the satellites where the sampling and resolutions are most different. By varying the size of the window (number of data points/altitude range) used when calculating the correlations, we can drastically increase or decrease the amplitude of this peak. As such, the drop in the correlations at the ozone maximum should be considered as an artefact and not a true measure of geophysical differences. At other altitudes, ozone concentration varies much more over time and with altitude, giving more meaningful estimates of correlation.

Ozone uncertainty analysis

Apart from the highlighted systematic differences and overall good tracking/correlation of the ozone concentration profiles, another important question we should ask is how realistic the combined measurement and statistical uncertainty estimates of the different systems are. In the case of the lidars, the small number of photons scattered back from the stratosphere and detected by the lidar receiver on the ground is generally the most important contributing factor to the measurement uncertainty (Godin-Beekmann et al., 2003; Leblanc et al., 2016a, b). Uncertainty sources for the ECC sondes include uncertain corrections for declining pump efficiency above 25 km, uncertain pressure/altitude registration, uncertain background current, evaporation of the sensing solution, and changing stoichiometry in the chemical cell (Tara-

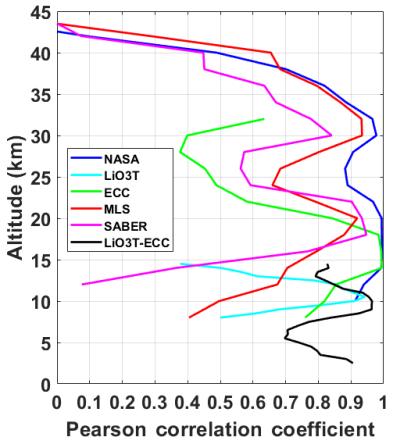

Figure 9. Vertical profiles of the correlation of ozone concentrations measured by the various LAVANDE instruments and ozone concentration measured by the LiO_3S (outliers were excluded). Correlation is taken over the 28 nights of the LAVANDE campaign and over 2 km in altitude. Results for MLS and SABER are calculated from the weighted average profiles. Slightly smaller correlations were obtained for the closest match SABER or MLS profiles (not shown).

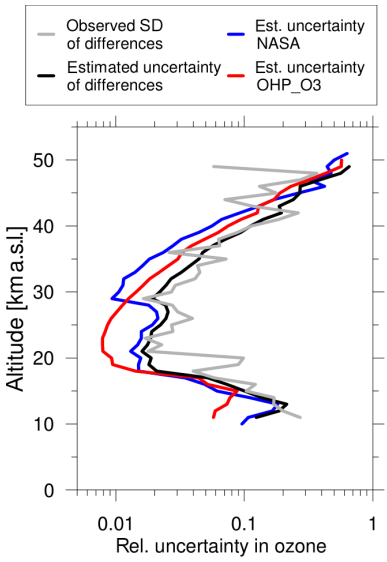

Figure 10. Vertical profiles of relative ozone uncertainties. Red indicates estimated by the LiO_3S retrieval. Blue indicates estimated by the NASA STROZ retrieval. Black indicates estimated for the relative ozone difference between NASA STROZ and LiO_3S (O_3(NASA)/O_3(OHP)−1). Grey indicates observed standard deviation for the relative ozone differences between NASA STROZ and LiO_3S during LAVANDE.

sick et al., 2016). The MLS and SABER satellite ozone retrievals also provide measurement uncertainty estimates (Waters et al., 2006; Froidevaux et al., 2008; Rezac et al., 2015a, b) associated with each individual 10 s profile. As was stated in Sect. 3, we use the same weighting technique on each of the associated measurement uncertainty profiles when calculating the "nightly average" measurement uncertainty profile for co-located satellite overpasses.

As previously mentioned, additional complications arise due to substantial variations in real ozone concentration between the OHP lidar measurement and a SABER or MLS ozone profile which can be measured many hundred kilometres and several hours away. In principle, such real differences can also occur for the ECC sondes. However, the ECC sondes during LAVANDE were fairly close to the lidar profiles, particularly in the troposphere. They were launched at OHP during the time of the lidar measurements and did not drift away by more than 100 km, even during the more variable weather and higher winds in the springtime part of the campaign.

Figure 10 shows the average of the total relative ozone uncertainty estimated by the LiO_3S and the NASA STROZ retrievals for nightly mean ozone profiles during LAVANDE. Both total uncertainty profiles are comparable and have a magnitude of less than 2 % between 20 and 35 km, with increasing measurement uncertainty towards higher and lower altitudes. Below 15 km, the combined uncertainty is in the range 5 % to 20 %, while above 35 km, the combined uncertainty increases to about 10 % near 40 km and to about 60 % near 50 km. Very similar combined ozone uncertainties are reported in the comprehensive NDACC lidar uncertainty budget analysis of Godin-Beekmann et al. (2003) and Braathen et al. (2004). Given that for comparisons between any two pairs of lidar measurements during the LAVANDE campaign, there is nearly perfect spatiotemporal coincidence, we can neglect geophysical variations in our uncertainty budget. This is not true for lidar comparisons with sondes, satellites, or NCEP. Assuming that there is no correlation between the average measurement noise of LiO_3S, σ_L (red), and NASA STROZ lidar, σ_N (blue), in Fig. 10, the relative standard deviation of the ozone difference, σ_{RSD}, between the two systems is given by Eq. 1 (grey), where L is the measurement of LiO_3S and N is the measurement of NASA STROZ; $\overline{L}$ and $\overline{N}$ are the respective average measurements of LiO_3S and NASA STROZ.

$$\sigma_{RSD} = \sqrt{\left(\frac{1}{N-1}\right)\Sigma\left(\left(\frac{N_i}{L_i}\right)-\left(\frac{\overline{N}}{\overline{L}}\right)\right)^2} \quad (1)$$

If the combined uncertainty estimates, expressed in Eq. (2) (black) are correct, it should be similar to the observed standard deviation of all the nightly mean ozone profile differences, σ_{RSD} (grey), expressed in Eq. (1) during LAVANDE.

$$\sigma_{combined} = \frac{\overline{N}}{\overline{L}}\sqrt{\left(\frac{\sigma_L}{\overline{L}}\right)^2+\left(\frac{\sigma_N}{\overline{N}}\right)^2} \quad (2)$$

Apart from some additional noise (especially near 20 km), agreement between the relative standard deviation of the ozone difference and observed standard deviation of all the

nightly mean ozone profile differences (black line and the grey line in Fig. 10) is quite good. From this agreement, we have a strong indication that the ozone uncertainties provided by the LiO_3S and NASA STROZ retrievals are realistic and we can proceed with our analysis.

Figure 11 shows similar results for the measurement uncertainty of ECC sondes (green line) and the OHP stratospheric and tropospheric DIALs (red and orange lines). In this case, the estimated combined uncertainty of the relative ozone difference (black line) is dominated at most altitudes by the larger measurement uncertainty of the ECC sondes (green line). Again, agreement between estimated combined ozone difference uncertainty (black line) and the corresponding observed standard deviation (grey line) is quite reasonable. However, to achieve this level of agreement, the estimate for ECC sonde ozone measurement uncertainty from Tarasick et al. (2016) had to be doubled (to about 5 % between 15 and 25 km, and to about 10 % below 10 km and above 30 km). This would indicate that, at least during LAVANDE, the ozone concentration uncertainty for ECC sondes might be larger than estimated by Tarasick et al. (2016); see also Smit (2013). It may improve once the homogenisation of the OHP dataset has been completed taking into account the use of 1 % KI concentration in the stratosphere data processing (3 %–10 %) and the humidification correction for the pump flow rate correction (1 %–4 %), which are not currently applied.

As previously discussed, the lidar measurements during LAVANDE were almost coincident in space and time, and the ECC sondes were very close. MLS and SABER satellite measurements, however, are usually taken for several hours and located several hundred kilometres away from the lidar measurement. Therefore, substantial additions to the total uncertainty in the relative ozone difference between MLS and the LiO_3S arise from geophysical ozone variations. This "sampling uncertainty" can be estimated by the standard deviation of all satellite profiles in the previously discussed coincidence box (see Fig. 1). Note that this standard deviation includes both sampling uncertainty due to true ozone variation over the box and measurement noise of the individual profiles.

The resulting uncertainties are shown in Fig. 12. At nearly all altitudes between 10 and 40 km, with the exception of 25 km where ozone variations are minimal (recall the discussion of the dip in the correlations in Fig. 9), the MLS sampling uncertainty (blue line) is clearly larger than the MLS measurement uncertainty for an individual profile (cyan line). From 37 to 47 km, MLS sampling uncertainty and measurement uncertainty for an individual profile are comparable, indicating that the estimate for measurement uncertainty is realistic and that geophysical ozone variability at these altitudes is small in comparison. However, above 47 km, sampling uncertainty is actually smaller than the estimated measurement uncertainty – indicating that the MLS measurement uncertainty estimate may be too conservative in this region.

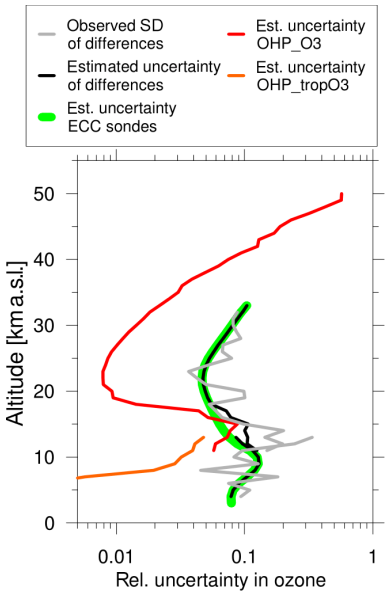

Figure 11. Vertical profiles of estimated relative ozone uncertainties for ECC sonde ozone profiles (green line, 2 times the estimate from Tarasick et al., 2016, excluding radiosonde pressure errors) and ozone measurement uncertainty estimated by the LiO_3S retrieval (red line), and the LiO_3T (orange line). Black indicates estimated combined uncertainty for the relative ozone difference between ECC sondes and the two LiO_3S (tropospheric system up to 13 km; stratospheric system above 10 km). Grey indicates corresponding observed standard deviation for the relative ozone differences during LAVANDE.

Comparing the grey and black lines in Fig. 12, it is obvious that MLS sampling uncertainty (blue line) plays a major role in this intercomparison. From 10 to 30 km, it is the dominant source of uncertainty and the major contributor to the observed standard deviation (grey line). Above 35 km, the estimated measurement uncertainty of LiO_3S (red line) is the dominant source of uncertainty – fully consistent with the observed standard deviation (grey line). From Fig. 12, it becomes clear that throughout most of the lower stratosphere, below 25 km, sampling uncertainty (spatial and temporal mismatches) is a major limitation for intercomparisons like LAVANDE. To narrow down uncertainties, closer matches and/or a much larger number of coincident events are needed.

Similar results can be seen for SABER ozone profiles in Fig. 13. Again, SABER sampling uncertainty (purple line) dominates the uncertainty budget in the relative ozone differences when compared to the LiO_3S between 20 and 35 km and needs to be considered to explain the observed standard deviation of the relative ozone differences (grey line). Above 35 km, the combined uncertainty in the ozone differences is again dominated by the measurement uncertainty of the LiO_3S ozone profiles (red line). Also above 35 km, estimated SABER measurement uncertainty (pink line) is much

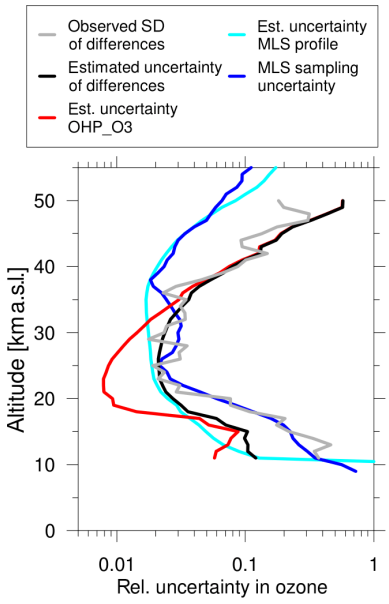

Figure 12. Vertical profiles of estimated relative ozone measurement uncertainties for individual MLS profiles (Froidevaux et al., 2008) from the MLS data files (cyan line), and MLS spatial variation/sampling uncertainty estimated from all profiles in the co-location box (blue line). LiO_3S measurement uncertainty is indicated by the red line. Estimated combined uncertainty for the relative ozone difference (MLS minus LiO_3S) based on MLS individual profile uncertainty is given by the black line. The grey line gives the observed standard deviation of the relative ozone differences between MLS and LiO_3S during LAVANDE.

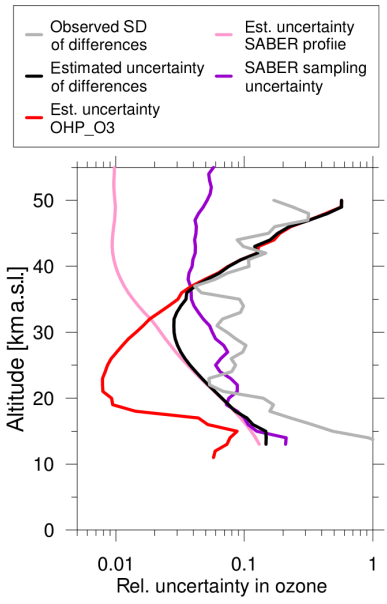

Figure 13. Vertical profiles of estimated relative ozone measurement uncertainties for individual SABER profiles (pink line, from Rong et al., 2009), SABER spatial variation/sampling uncertainty over the co-location box (purple line), and LiO_3S measurement uncertainty (red line). Estimated combined uncertainty for the relative ozone difference SABER minus LiO_3S based on SABER individual profile uncertainty is given by the black line. The grey line gives observed standard deviation of the relative ozone differences between SABER and LiO_3S.

smaller than the observed SABER sampling uncertainty (purple line). With the limited number of coincident measurements available during LAVANDE, it was, however, not possible to check if this small SABER measurement uncertainty estimate (pink line) is realistic or too optimistic.

5 Intercomparison results for temperature

Similar to the analysis done in Fig. 6 for ozone, Fig. 14 shows examples for the temperature time series recorded by the different systems during LAVANDE. As was the case for ozone, a seasonal variation is apparent in the temperature profiles between the two different periods of July 2017 and March 2018. In the upper stratosphere, above 30 km, temperatures were colder in March 2018 than in July 2017, whereas, in the mesosphere, above 70 km, temperatures were colder in July 2017. All the LAVANDE instruments track these expected seasonal variations. Shorter-term variations, such as the slight temperature oscillation appearing near 70 km during July 2017, are also tracked by all the instruments. Near 30 km, some of the NASA STROZ data points after 22 July seem to lie outside of the usual range, but the temperatures at higher altitudes are consistent with data points from the LTA.

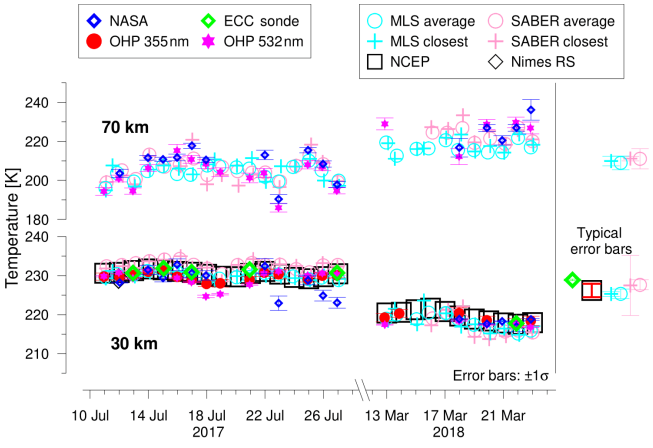

Figure 14. Time series of the temperatures measured by the different systems for selected altitude levels during LAVANDE.

This indicates that NASA STROZ generally provides correct temperature profiles but may have experienced a slight misalignment in a couple of nights.

The average temperature difference between the various systems and NASA STROZ is presented in Fig. 15. Unlike the ozone analysis, where the LiO_3S was chosen as the reference, NASA STROZ was chosen here as the reference for

temperature, because it had measurements in nearly all nights and covered a wider altitude range for temperature than either the LiO_3S or the LTA. For most altitudes between 25 km and $\approx$ 70 km, the agreement between the temperatures from the different LAVANDE systems and temperature from NASA STROZ is better than $\pm$2 K. Below about 35 km, temperatures from the LiO_3S (red), Nîmes radiosondes (yellow), the radiosondes coupled to the OHP ECC sondes (black), and NCEP analyses (cyan) are very similar, indicating that temperatures from NASA STROZ might be too low by 1 to 4 K in this altitude range. The pronounced increasing cold bias of the LTA data below 30 km arises from signal contamination by aerosols in the lower stratosphere. This bias is less evident in NASA STROZ and LiO_3S as these two lidars operate in the UV at 355 nm as opposed to LTA which operates in the visible at 532 nm and is more susceptible to contamination by aerosol scattering. Above 60 km, LTA (green), SABER (blue), and MLS (magenta) report lower temperatures than those provided by NASA STROZ. It appears that NASA STROZ might have a slight warm bias in the upper stratosphere and lower mesosphere, with respect to LTA, which gradually reaches 5 K near 80 km. Warm biases at the top of the lidar temperature profile are commonly associated with errors induced by the a priori used to initialise the lidar temperature calculation at the topmost levels or by underestimation of the background (Wing et al., 2018a; Sica and Haefele, 2015). A full study of the effects of the a priori selection, initialisation altitude, and tie-on uncertainty would be a good topic for another NDACC algorithm validation article where we are not constrained by the need to perform a "blind" comparison.

Several other interesting features appear in the temperatures difference profiles at middle altitudes:

- The higher temperatures reported by the LiO_3S below 22 km with respect to the other measurements.
- The higher temperatures between 30 and 55 km reported by the LTA. Compared to NASA STROZ, the LTA reports about 2 K higher temperatures near 40 km and 2 K lower temperatures near 70 km in Fig. 15. Interestingly, this is almost the exact opposite of the difference found between the same two systems in the July 1997 OTOIC intercomparison (Braathen et al., 2004). In OTOIC, NASA STROZ reported about 2 K higher temperatures than the LTA near 40 km, and 2 K lower temperatures near 70 km. On the other hand, the $\approx$ 1 K higher temperatures between 35 and 50 km from the LTA compared to the LiO_3S during LAVANDE in Fig. 15 are generally consistent with the similar, but slightly smaller, difference found between the same two systems over the 20-year period from 1993 to 2013 by Wing et al. (2018a).
- The already-mentioned lower temperatures reported by the LTA below 30 km. These are attributed to the much more significant contamination by aerosol scattering at the 532 nm wavelength used by this lidar (compared to 355 nm, used by the other lidars).
- The lower temperatures near 43 km and higher temperatures above 50 km provided by the NCEP analyses which may in part be due to the vertical averaging and data density differences between lidar measurements and Advanced Microwave Sounding Unit (AMSU) as demonstrated by Funatsu et al. (2008).

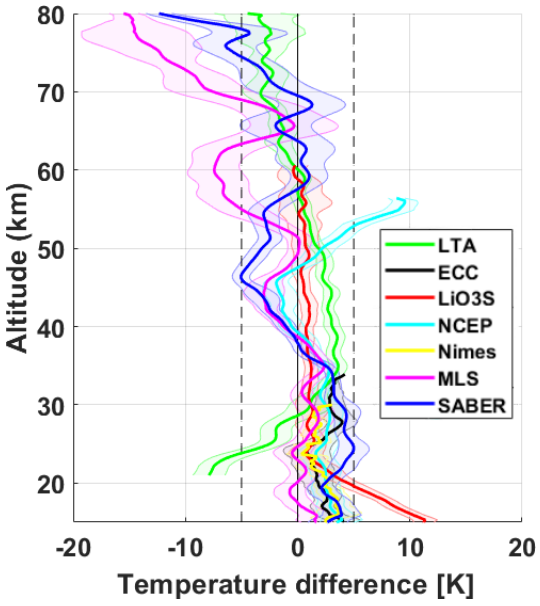

Figure 15. Average absolute difference profile between the temperature measured by the various LAVANDE instruments and temperature measured by NASA STROZ. The shaded range gives $\pm$2 standard deviations of the mean and indicates statistical uncertainty at the 95 % confidence level. Results for MLS and SABER are for the weighted average profiles, but very similar results are obtained using the closest match SABER or MLS profiles.

MLS and SABER temperatures stand out from the ground-based temperature observations as the temperatures exhibit oscillating biases between 35 and 80 km that can reach up to $-$5 K. A similar oscillating bias for MLS temperatures compared to the OHP lidars ($-$4 to $-$6 K near 42 km and near 60 km, no bias near 50 km) was also seen in the 2004–2018 long-term intercomparison by Wing et al. (2018b). The same study also found an "S-shaped" bias for SABER temperatures which also appears in Fig. 15. There, SABER temperatures have a warm bias compared to the three temperature lidars below 30 km and a cold bias between 40 and 50 km. Wing et al. (2018b) attributed a substantial part of these satellite temperature biases to altitude shifts introduced by the satellite retrieval algorithms.

Examining the scatter of the LAVANDE instrument temperatures in three different altitude regimes yields more detail about the relative biases of each instrument. Figure 16a compares the LAVANDE temperatures from 12 to 35 km to

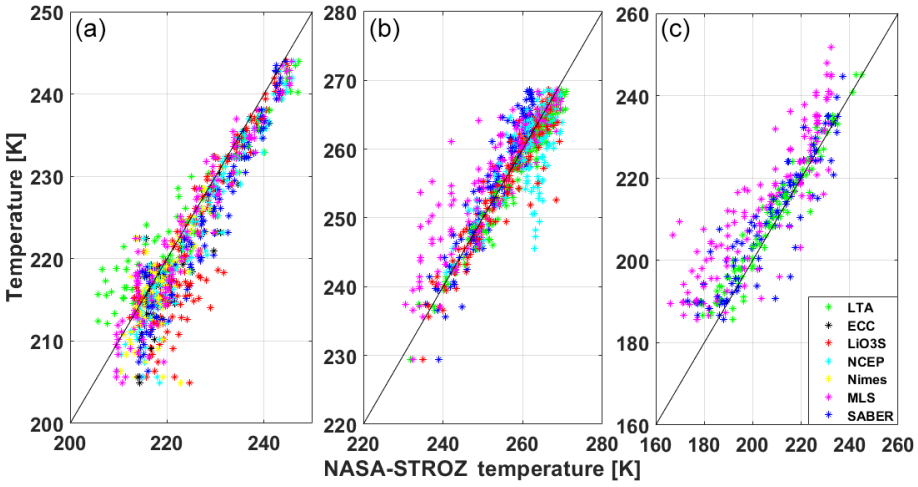

Figure 16. Scatter plots of temperature as measured by the various LAVANDE instruments (along the vertical axis) and temperature measured by NASA STROZ (along the horizontal axis). **(a)** Temperature from 12 to 35 km altitude. **(b)** Temperature from 35 to 60 km altitude. **(c)** Temperature from 60 to 80 km altitude.

NASA STROZ. We can see that LTA (green) has a clear aerosol-induced cold bias in the lower half of the panel as it is systematically colder than every other measurement. We can also see that most data points for the other instruments are below the black reference line, indicating that in this altitude range NASA STROZ reported reliably colder temperatures. Figure 16b represents measurements from 35 to 60 km and exhibits tight correlation between all measurements except MLS. As was noted in Fig. 15, MLS (magenta) has an oscillation in the sign of the temperature bias with respect to the other measurements which is seen here as increased scatter. We can also see the cold bias of NCEP (cyan) in the upper stratosphere. Figure 16c represents measurements from 60 to 80 km and includes only NASA STROZ (reference), LTA (green), MLS (magenta), and SABER (blue). There is generally good tacking between the two lidars with larger scatter for MLS and SABER. We can see some evidence that NASA STROZ is warmer than the other measurements but not on all nights.

The temperature correlation plot in Fig. 17 shows the extent to which temperatures reported by the various systems track the temperature variation measured by NASA STROZ. The highest correlations, ≥ 0.8, are seen below 35 km and above 55 km for LTA. Correlations drop significantly near 25 km and again around 50 km, which corresponds to regions just above the tropopause and around the stratopause. Similar to the case for ozone in Fig. 9, these drops are associated with small temperature variance at these altitudes, where temperature changes little with altitude, and night-to-night temperature variations are also small. Measurement noise/uncertainty then becomes prominent and decreases correlations.

Other points to note include the following. (1) The correlation between the NASA STROZ and OHP temperature profiles increases again above 50 km and exceeds 0.9 between 60 and 80 km. (2) Lower correlation is seen for temperature from the LiO_3S above 50 km. This is likely caused by increasing measurement uncertainty for temperature from the LiO_3S above 55 km which is associated with the lower laser output at 355 nm in this system. The 355 nm Nd:YAG energy output in LiO_3S is intentionally reduced by manually introducing delay in the laser oscillator. This is done to optimise the system for comparison with the 308 nm laser signal. (3) MLS and SABER temperatures show lower correlation with respect to all other instruments. (4) Excluding the region associated with the tropospheric temperature minimum,

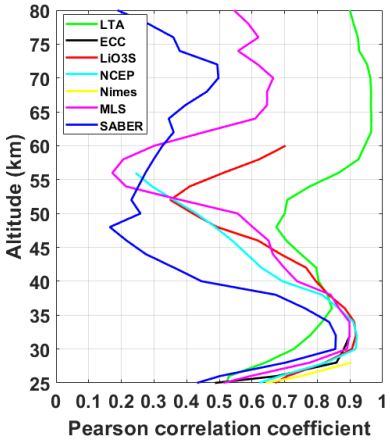

Figure 17. Vertical profiles of the correlation between temperatures reported by the various LAVANDE systems and temperature measured by NASA STROZ. Outliers were excluded. Correlation is taken over the 28 nights of LAVANDE, and over 2 km in altitude. For MLS and SABER, correlations are given for the weighted average profiles.

the correlation between NASA STROZ temperatures and the on-site ECC sondes, Nîmes radiosondes (up to 30 km), and NCEP analyses (up to about 40 km) is also good. Above 40 km, in the topmost NCEP analysis pressure levels at 1 and 0.4 hPa (≈48 and 54 km), correlation drops rapidly for the NCEP analyses. This has also been seen in previous intercomparisons (e.g. Steinbrecht et al., 2009b). At these top levels, the NCEP analyses are relaxed substantially towards a climatological state and are much less responsive to actual temperature variations.

5.1 Lidar temperature uncertainty analysis

A closer look at temperature measurement uncertainties is taken in Figs. 18–21. The approach in this section is the same as for ozone in the previous section. Figure 18 shows the estimated temperature measurement uncertainty for NASA STROZ (blue) and LiO_3S (red). The largest term contributing to the total uncertainty for lidar temperatures below 80 km comes from the Poisson statistics of the limited number of photons scattered back from high altitudes (see, e.g. Leblanc et al., 2016c; Sica and Haefele, 2015). The temperature measurement uncertainty for NASA STROZ is estimated to be less than 1 K between 15 and 50 km, increasing to 4 K near 80 km, very similar to the comprehensive uncertainty given for a typical stratospheric lidar in Fig. 10 of Leblanc et al. (2016c). For LiO_3S, temperature measurement uncertainty is also estimated to be less than 1 K below 30 km but increases to 10 K near 60 km. From these two measurement uncertainties, the combined uncertainty of the difference between coincident temperature profiles from LiO_3S and NASA STROZ can be estimated (similar to what was discussed in the previous section for ozone). This estimated combined uncertainty of the temperature difference is shown by the black line in Fig. 18.

If the estimated measurement uncertainties for the two lidars are correct, the black line in Fig. 18 should be very similar to the grey line, which shows the observed standard deviation of the temperature difference between LiO_3S and NASA STROZ over all the (nearly coincident) measurements during LAVANDE. Unfortunately, the agreement between the black and grey curves is not so good in Fig. 18. Above 30 km, the observed standard deviation is actually smaller than the estimated combined uncertainty by a factor of about 2. This indicates that the estimated temperature measurement uncertainty for the LiO_3S is too large above 30 km, by a factor of about 2. This may arise from incorrect accounting for the vertical integration and filtering of the temperature profile in the measurement uncertainty estimate for the LiO_3S. On the other hand, below 30 km, the observed standard deviation (grey line) is larger than the estimated combined uncertainty (black line) again by a factor of about 2. This would indicate that the estimated temperature measurement uncertainty for LiO_3S and/or NASA STROZ is too small, by a factor of 2 or more. It could mean that other sources of uncertainty, beyond statistical uncertainty, are important. Future work will be conducted using the results of this intercomparison campaign to refine the LiO_3S error budget for temperature.

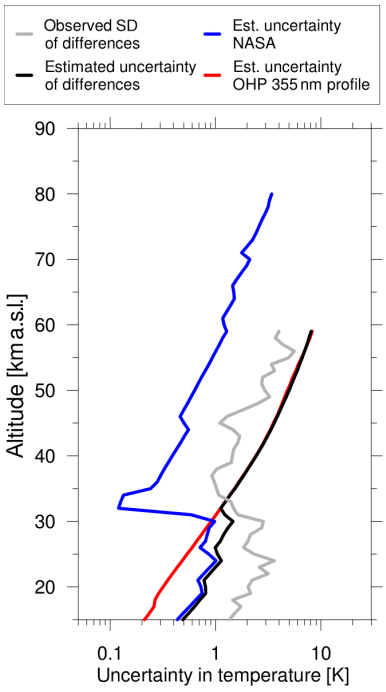

Figure 18. Similar to Figs. 10–13 but for temperature. The estimated measurement uncertainty for temperature measured by LiO_3S (red curve) and NASA STROZ (blue curve), estimated combined uncertainty for temperature differences between the two systems (black curve), and observed standard deviation of temperature differences between the two systems (grey curve) during LAVANDE are plotted.

The corresponding comparison of uncertainties for LTA and NASA STROZ are given in Fig. 19. Both systems have very similar estimates of temperature measurement uncertainty, which are also consistent with the recommendations of Leblanc et al. (2016c). Above 60 km, the estimated combined uncertainty of the temperature difference (black curve) is similar to the observed standard deviation during LAVANDE (grey curve), confirming the measurement uncertainty estimates for the two lidars above 60 km. However, at most altitudes below 60 km, the observed standard deviation (grey curve) remains at 2 to 3 K. This is substantially larger, by up to a factor of 10, than the estimate of the combined uncertainty (black curve).

This result indicates that the measurement uncertainty estimates for LTA and NASA STROZ are too optimistic during LAVANDE. Detector misalignment in one or both lidars is likely the main cause of the reported disagreement. At OHP, the alignment is made manually each night by operators and a slight misalignment may induce a detectable temperature bias. Given that even a small 1 % error in the slope of the density profile can induce a 2 to 2.5 K bias in the resulting

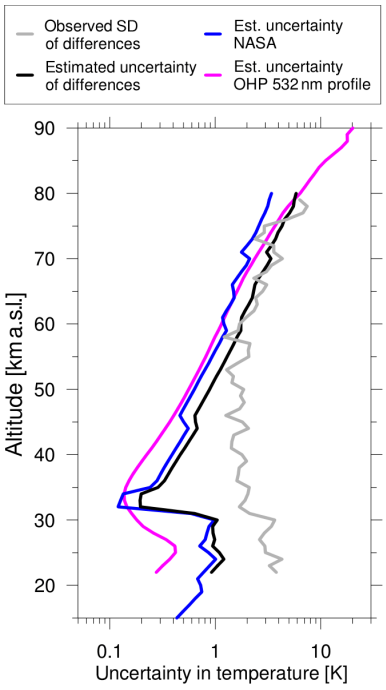

Figure 19. Same as Fig. 18 but for LTA (magenta curve) and NASA STROZ.

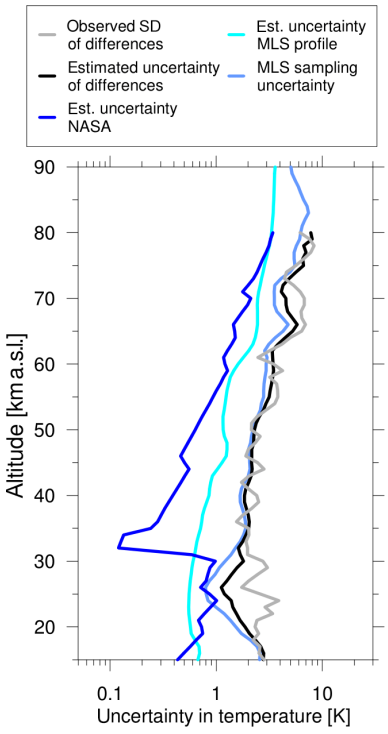

Figure 20. Same as Figs. 18 and 19 but comparing measurement and sampling uncertainties of MLS satellite temperature profiles (cyan and light blue curves) and NASA STROZ ground-based profiles (dark blue). Results are for the MLS weighted average profiles, but very similar results are obtained for closest match MLS profiles. MLS single-profile measurement uncertainty is included in the data distribution and is described in Schwartz et al. (2008).

temperature profile, the possibility of human errors exists. A key conclusion from this study is that automatic alignment systems for NDACC lidars are essential for measurement accuracy and long-term stability. Another source of error may come from the linearisation correction of the photon counting at high counting rate.

5.2 Satellite temperature uncertainty analysis

In the next section, we extend the comparison of uncertainty estimates and observed difference standard deviation to temperature profiles from the MLS and SABER satellite instruments. As with ozone, temporal and spatial mismatch between the lidar measurement at OHP and the number of satellite measurements within the chosen coincidence box (see Fig. 1) plays an important role. Figure 20 allows comparison of the measurement uncertainty given for the MLS data (cyan line) with the estimated sampling uncertainty for the weighted mean MLS profile (light blue curve). Sampling uncertainty is estimated by the weighted standard deviation of all MLS profiles in the coincidence box (which implicitly includes single-profile uncertainty). Clearly, for MLS, sampling uncertainty is larger than single-profile measurement uncertainty (e.g. Schwartz et al., 2008) by a factor of about 2. Sampling uncertainty is also the dominant source of uncertainty and accounts for most of the standard deviation of all MLS minus NASA STROZ temperature differences observed in LAVANDE (grey curve). When sampling uncertainty is included in the estimate for total temperature difference combined uncertainty (black curve in Fig. 20), good agreement is obtained with the observed standard deviation (grey curve). This good agreement would not be achieved if only the MLS single-profile measurement uncertainty would be considered (cyan line). Then the corresponding estimated temperature difference uncertainty would be too small. Overall, Fig. 20 confirms that (1) MLS single-profile temperature measurement uncertainty is of the order of 1 to 3 K; (2) NASA STROZ provides comparable single-profile measurement uncertainty (1 to 3 K); and (3) sampling uncertainty plays an important role in the total uncertainty budget for the satellite vs. ground-based intercomparison, contributing a combined uncertainty of 2 to 5 K during LAVANDE.

Similar results are obtained in Fig. 21 for SABER temperature profiles. Also for SABER, sampling uncertainty (purple curve) is larger than single-profile measurement uncertainty (pink curve, estimated following Rezac et al., 2015a, b). Sampling uncertainty must, again, be considered to explain the observed standard deviation of SABER – NASA STROZ temperature differences (black curve matching grey curve) particularly above 55 km. Below 40 km, however, the observed standard deviation (grey) is about 2 K larger than estimated from SABER sampling uncertainty and NASA

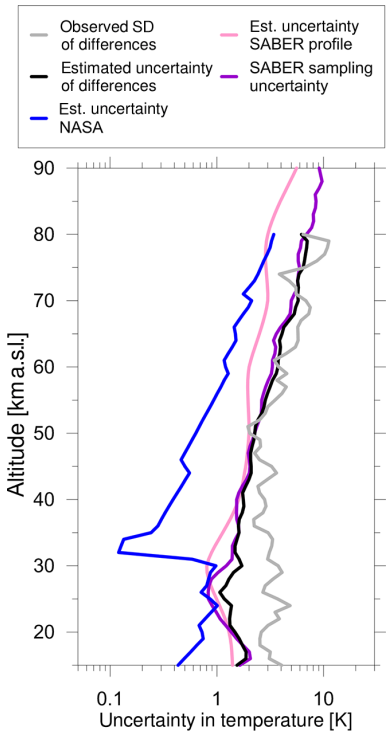

Figure 21. Same as Fig. 20 but for SABER satellite temperature profiles (pink for measurement uncertainty and purple for sampling uncertainty) and NASA STROZ (blue for measurement uncertainty). Results for SABER are shown for the weighted average profiles, but very similar results are obtained for closest match SABER profiles. SABER single-profile temperature measurement uncertainty was estimated following Rezac et al. (2015a, b).

STROZ temperature measurement uncertainty. A similar disagreement below ≈ 30 km was already mentioned for the LiO_3S vs. NASA STROZ comparison in Fig. 18 and for the LTA vs. NASA STROZ comparison below 50 km in Fig. 19. Using LiO_3S or LTA instead of NASA STROZ as the temperature reference produces very similar results, as are currently shown in Figs. 18–21.

Given that in Figs. 18 and 19 we see a larger standard deviation between pairs of coincident lidar measurements (grey) than the estimated combined uncertainty (black) gives us reason to expect, we suggest that additional uncertainty sources not considered in Leblanc et al. (2016c) may play a role (e.g. temporal changes in alignment, defocusing, multiple scattering). Additionally, the unexpectedly large standard deviation between the lidar and SABER results seen in Fig. 21 (grey), which may be due to unaccounted uncertainties in the SABER error budget, suggests a lower limit on the total temperature uncertainty budget of 1 to 3 K below 50 km. Taken together, these two suggestions imply that variations of approximately 3 K in the ensemble temperature differences seen in Fig. 15 are a reasonable threshold for validation of the participating lidar systems in the context of this LAVANDE campaign.

6 Conclusions

The LAVANDE intercomparison of the OHP lidars (tropospheric DIAL, stratospheric DIAL, and Rayleigh temperature), local radiosondes and ECCs, satellite instruments MLS and SABER, and the mobile NDACC reference lidar NASA STROZ has shown overall good tracking of both vertical profiles of temperature and ozone for all participating instruments. LAVANDE was a "blind" intercomparison; i.e. all ground-based measurements presented here were submitted "blind". There was no possibility to see results from the other instruments before submitting each group's data.

Agreement for ozone was within $\pm 10\,\%$ for all instruments between approximately 15 and 40 km. Agreement was closer, better than $\pm 5\,\%$, between 18 and 38 km for the two stratospheric DIAL systems. Some statistically significant differences are present in the two stratospheric systems when measuring low ozone densities below 14 km and above 40 km. The tropospheric DIAL, LiO_3T, also reported lower ozone concentrations than the local ECC and lower than LiO_3S above 10 km (bias $> 10\,\%$).

Although this may improve with further corrections of the ECC in the stratosphere, it is related to the increasing measurement uncertainty of the LiO_3T near its upper measurement range. Improvement of the lidar data processing and removal of this potential bias will be investigated in future work involving optimal estimation techniques (Farhani et al., 2019). Future tropospheric ozone lidar campaigns for NDACC lidars would be required to assess the new technique and fully characterise any residual biases. MLS and SABER ozone profiles agree with the profiles produced by lidars and ECCs from about 20 to above 40 km. Below 20 km, both sets of satellite profiles deviate significantly from the lidars and the ECCs. Above 40 km, ozone measurement uncertainties become large for the lidars, and differences increase while their significance goes down.

The assessment of the uncertainty budget for ozone concentration profiles for each instrument showed that the reported measurement uncertainties for both LiO_3S and NASA STROZ are well characterised and realistic. The reported measurement uncertainty estimates for ECCs from Tarasick et al. (2016) appear too optimistic for the sondes launched during LAVANDE. They seem to underestimate the combined uncertainty for the LAVANDE ECC sondes by a factor of 2. When comparing the ground-based profiles to the satellite measurements, it is necessary to account for sampling uncertainty, i.e. real ozone differences between the ground-based profile and the satellite profiles measured a couple hundred kilometres and a few hours away. This sampling uncertainty for MLS was greater than the reported single-profile measurement uncertainty below 30 km and dominates the er-

ror budget in this region. For SABER, sampling uncertainty is substantially larger than single-ozone-profile measurement uncertainty at all altitudes above 30 km. Above 35 km, MLS and SABER sampling uncertainty was less relevant, because lidar ozone measurement uncertainties become larger.

Agreement for temperature was within ±5 K for all instruments between approximately 25 and 80 km. Below 30 km, the LTA operating at 532 nm has a well-known aerosol-induced cold bias relative to the other instruments. This bias will be corrected in the future with the installation of a rotational Raman channel for lower atmospheric temperatures. The LiO_3S reports significantly higher temperatures below 23 km, which will be corrected in future data releases. NASA STROZ has an apparent warm bias above 70 km, likely due to a priori assumptions or background estimations made in the profile retrieval. Radiosondes and ECCs are in good agreement with the lidar profiles. MLS has a pronounced oscillating temperature bias throughout the middle atmosphere. SABER has a slight cold bias near the stratopause (45 km). Both of these biases are consistent with altitude distortions in the satellite retrieved altitude grid (see also Wing et al., 2018b).

The assessment of the uncertainty budget for temperature profiles showed that the reported measurement uncertainties for the LiO_3S may be underestimated below 30 km and overestimated at higher altitudes. Both the LTA and NASA STROZ appear to underestimate the combined uncertainty in the temperature profiles below 55 km. This may indicate that other sources of uncertainty, beyond those in Leblanc et al. (2016c), may need to be considered or that further work can be done in addressing potential sources of measurement bias (e.g. alignment, a priori temperature initialisation, deadtime corrections). When comparing ground-based the temperature profiles with satellite measured profiles from MLS and SABER, it is necessary to include sampling uncertainty. For MLS, sampling uncertainty during LAVANDE was between 2 and 8 K, about 2 times larger than single-profile measurement uncertainty at most altitudes from 20 to 80 km. Similar sampling uncertainty was found for temperature profiles measured by SABER during LAVANDE.

Overall, the LAVANDE campaign has successfully validated the NDACC lidar profiles for both temperature and ozone over a large vertical extent. We have identified a few minor biases existing at both the low and high limits of our profiles, which we shall address going forward. Additionally, we have shown that sampling uncertainty can be the largest contributing factor to the observed standard deviations in lidar–satellite comparisons and that NDACC temperature lidars have a larger standard deviation below 50 km than can be explained solely by the combined measurement and sampling uncertainties.

Data availability. The data that support the findings of this study are openly available. The data used in this publication were obtained from LATMOS as part of the Network for the Detection of Atmospheric Composition Change (NDACC) and are publicly available at ftp://ftp.cpc.ncep.noaa.gov/ndacc/station/ohp/ (LATMOS, 2018); local radiosoundings from Nîmes are available at http://weather.uwyo.edu/upperair/sounding.html (Météo-France, 2018); NCEP model profiles are available at ftp://ftp.cpc.ncep.noaa.gov/ndacc/ncep/temp/ (NOAA Physical Sciences Laboratory, 2018); MLS temperature and ozone profiles are available at https://disc.gsfc.nasa.gov/datasets?keywords=MLS (NASA, 2018); and SABER temperature and ozone profiles are available at http://saber.gats-inc.com/ (GATs, 2018).

Author contributions. RW, SGB, TJM, JTS, and GS conducted the measurement campaign at OHP. WS conducted the blind comparison of all LAVANDE data. RW and WS drafted the article. AH, SGB, SK, GA, PK, and TJM provided access to the data and instruments. All authors discussed the results and contributed to the final paper.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The authors would particularly like to thank the technicians at La Station Géophysique Gérard Mégie at OHP, who are so important for running the long-term measurement program.

Financial support. This research has been supported by the Institut National des Sciences de l'Univers/Centre National de la Recherche Scientifique (INSU/CNRS), Université de Versailles Saint-Quentin-en-Yvelines (UVSQ), Centre National d'Études Spatiales (CNES), the NASA Upper Atmospheric Research Program, and ARISE2.

Review statement. This paper was edited by Michel Van Roozendael and reviewed by two anonymous referees.

References

Ancellet, G. and Beekmann, M.: Evidence for changes in the ozone concentrations in the free troposphere over southern France from 1976 to 1995, Atmos. Environ., 31, 2835–2851, 1997.

Beekmann, M., Ancellet, G., Martin, D., Abonnel, C., Duverneuil, G., Eideliman, F., Bessemoulin, P., Fritz, N., and Gizard, E.: Intercomparison of tropospheric ozone profiles obtained by electrochemical sondes, a ground based lidar and an airborne UV-photometer, Atmos. Environ., 29, 1027–1042, 1995.

Braathen, G. O., Godin-Beekmann, S., Keckhut, P., McGee, T. J., Gross, M. R., Vialle, C., and Hauchecorne, A.: Intercomparison of stratospheric ozone and temperature measurements at the Observatoire de Haute Provence during the OTOIC NDSC validation campaign from 1–18 July 1997, Atmos. Chem. Phys. Dis-

cuss., 4, 5303–5344, https://doi.org/10.5194/acpd-4-5303-2004, 2004.

De Mazière, M., Thompson, A. M., Kurylo, M. J., Wild, J. D., Bernhard, G., Blumenstock, T., Braathen, G. O., Hannigan, J. W., Lambert, J.-C., Leblanc, T., McGee, T. J., Nedoluha, G., Petropavlovskikh, I., Seckmeyer, G., Simon, P. C., Steinbrecht, W., and Strahan, S. E.: The Network for the Detection of Atmospheric Composition Change (NDACC): history, status and perspectives, Atmos. Chem. Phys., 18, 4935–4964, https://doi.org/10.5194/acp-18-4935-2018, 2018.

Farhani, G., Sica, R. J., Godin-Beekmann, S., Ancellet, G., and Haefele, A.: Improved ozone DIAL retrievals in the upper troposphere and lower stratosphere using an optimal estimation method, Appl. Optics, 58, 1374–1385, 2019.

Froidevaux, L., Jiang, Y. B., Lambert, A., Livesey, N. J., Read, W. G., Waters, J. W., Browell, E. V., Hair, J. W., Avery, M. A., McGee, T. J., Twigg, L. W., Sumnicht, G. K., Jucks, K. W., Margitan, J. J., Sen, B., Stachnik, R. A., Toon, G. C., Bernath, P. F., Boone, C. D., Walker, K. A., Filipiak, M. J., Harwood, R. S., Fuller, R. A., Manney, G. L., Schwartz, M. J., Daffer, W. H., Drouin, B. J., Cofield, R. E., Cuddy, D. T., Jarnot, R. F., Knosp, B. W., Perun, V. S., Snyder, W. V., Stek, P. C., Thurstans, R. P., and Wagner, P. A.: Validation of Aura Microwave Limb Sounder stratospheric ozone measurements, J. Geophys. Res., 113, D15S20, https://doi.org/10.1029/2007JD008771, 2008.

Funatsu, B., Claud, C., Keckhut, P., and Hauchecorne, A.: Cross-validation of Advanced Microwave Sounding Unit and lidar for long-term upper-stratospheric temperature monitoring, J. Geophys. Res., 113, D23108, https://doi.org/10.1029/2008JD010743, 2008.

GATs, SABER data, available at: http://saber.gats-inc.com/, last access: 17 June 2018.

Gaudel, A., Ancellet, G., and Godin-Beekmann, S.: Analysis of 20 years of tropospheric ozone vertical profiles by lidar and ECC at Observatoire de Haute Provence (OHP) at 44° N, 6.7° E, Atmos. Environ., 113, 78–89, https://doi.org/10.1016/j.atmosenv.2015.04.028, 2015.

Godin, S., Mégie, G., and Pelon, J.: Systematic lidar measurements of the stratospheric ozone vertical distribution, Geophys. Res. Lett., 16, 547–550, https://doi.org/10.1029/GL016i006p00547, 1989.

Godin, S., Carswell, A., Donovan, L., Claude, H., Steinbrecht, W., Mc Dermid, I., Mc Gee, T., Gross, M., Nakane, H., Swart, D., Bergwerff, H., Uchino, O., Gathen, P., and Neuber, R.: Ozone differential absorption lidar algorithm intercomparison, Appl. Optics, 38, 6225–6236, https://doi.org/10.1364/AO.38.006225, 1999.

Godin-Beekmann, S., Porteneuve, J., and Garnier, A.: Systematic DIAL lidar monitoring of the stratospheric ozone vertical distribution at Observatoire de Haute-Provence (43.92° N, 5.71° E), J. Environ. Monitor., 5, 57–67, https://doi.org/10.1039/b205880d, 2003.

Google Earth Pro, M.: Observatoire de Haute Provence (CNRS) Kernel Description, available at: https://www.google.fr/maps/place/Observatoire+de+Haute+Provence+(CNRS)/@43.9236737,5.7183398 (last access: 12 June 2018), 2019.

Hauchecorne, A. and Chanin, M.-L.: Density and temperature profiles obtained by lidar between 35 and 70 km, Geophys. Res. Lett., 7, 565–568, https://doi.org/10.1029/GL007i008p00565, 1980.

Keckhut, P., Hauchecorne, A., and Chanin, M.: A critical review of the database acquired for the long-term surveillance of the middle atmosphere by the French Rayleigh lidars, J. Atmos. Ocean. Tech., 10, 850–867, https://doi.org/10.1175/1520-0426(1993)010<0850:ACROTD>2.0.CO;2, 1993.

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. Monitor., 6, 721–733, https://doi.org/10.1039/b404256e, 2004.

Kurylo, M. J., Thompson, A. M., and De Mazière, M.: The Network for the Detection of Atmospheric Composition Change: 25 Years Old and Going Strong, The Earth Observer, 28, 4–15, available at: https://eospso.gsfc.nasa.gov/earthobserver/sep-oct-2016 (last access: 15 March 2018), 2016.

LATMOS: NDACC lidar data, availablea at: ftp://ftp.cpc.ncep.noaa.gov/ndacc/station/ohp/, last access: 17 June 2018.

Leblanc, T., Walsh, T. D., McDermid, I. S., Toon, G. C., Blavier, J.-F., Haines, B., Read, W. G., Herman, B., Fetzer, E., Sander, S., Pongetti, T., Whiteman, D. N., McGee, T. G., Twigg, L., Sumnicht, G., Venable, D., Calhoun, M., Dirisu, A., Hurst, D., Jordan, A., Hall, E., Miloshevich, L., Vömel, H., Straub, C., Kampfer, N., Nedoluha, G. E., Gomez, R. M., Holub, K., Gutman, S., Braun, J., Vanhove, T., Stiller, G., and Hauchecorne, A.: Measurements of Humidity in the Atmosphere and Validation Experiments (MOHAVE)-2009: overview of campaign operations and results, Atmos. Meas. Tech., 4, 2579–2605, https://doi.org/10.5194/amt-4-2579-2011, 2011.

Leblanc, T., Sica, R. J., van Gijsel, J. A. E., Godin-Beekmann, S., Haefele, A., Trickl, T., Payen, G., and Gabarrot, F.: Proposed standardized definitions for vertical resolution and uncertainty in the NDACC lidar ozone and temperature algorithms – Part 1: Vertical resolution, Atmos. Meas. Tech., 9, 4029–4049, https://doi.org/10.5194/amt-9-4029-2016, 2016a.

Leblanc, T., Sica, R. J., van Gijsel, J. A. E., Godin-Beekmann, S., Haefele, A., Trickl, T., Payen, G., and Liberti, G.: Proposed standardized definitions for vertical resolution and uncertainty in the NDACC lidar ozone and temperature algorithms – Part 2: Ozone DIAL uncertainty budget, Atmos. Meas. Tech., 9, 4051–4078, https://doi.org/10.5194/amt-9-4051-2016, 2016b.

Leblanc, T., Sica, R. J., van Gijsel, J. A. E., Haefele, A., Payen, G., and Liberti, G.: Proposed standardized definitions for vertical resolution and uncertainty in the NDACC lidar ozone and temperature algorithms – Part 3: Temperature uncertainty budget, Atmos. Meas. Tech., 9, 4079–4101, https://doi.org/10.5194/amt-9-4079-2016, 2016c.

Leblanc, T., Brewer, M. A., Wang, P. S., Granados-Muñoz, M. J., Strawbridge, K. B., Travis, M., Firanski, B., Sullivan, J. T., McGee, T. J., Sumnicht, G. K., Twigg, L. W., Berkoff, T. A., Carrion, W., Gronoff, G., Aknan, A., Chen, G., Alvarez, R. J., Langford, A. O., Senff, C. J., Kirgis, G., Johnson, M. S., Kuang, S., and Newchurch, M. J.: Validation of the TOLNet lidars: the Southern California Ozone Observation Project (SCOOP), At-

mos. Meas. Tech., 11, 6137–6162, https://doi.org/10.5194/amt-11-6137-2018, 2018.

Margitan, J. J., Barnes, R. A., Brothers, G. B., Butler, J., Burris, J., Connor, B. J., Ferrare, R. A., Kerr, J. B., Komhyr, W. D., McCormick, M. P., McDermid, I. S., McElroy, C. T., McGee, T. J., Miller, A. J., Owens, M., Parrish, A. D., Parsons, C. L., Torres, A. L., Tsou, J. J., Walsh, T. D., and Whiteman, D. : Stratospheric ozone intercomparison campaign (STOIC) 1989: Overview, J. Geophys. Res.-Atmos., 100, 9193–9207, 1995.

Mauldin III, 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: Optical Remote Sensing of the Atmosphere and Clouds, vol. 3501, 355–365, International Society for Optics and Photonics, Bellingham, WA, USA, 1998.

McDermid, I., Bergwerff, J., Bodeker, G., Boyd, I., Brinksma, E., Connor, B., Farmer, R., Gross, M., Kimvilakani, P., Matthews, W., McGee, T., Ormel, F., Parrish, A., Singh, U., Swart, D., and Tsou, J.: OPAL: Network for the detection of stratospheric change ozone profiler assessment at Lauder, New Zealand 2. Intercomparison of revised results, J. Geophys. Res., 103, 28693–28699, https://doi.org/10.1029/98JD02707, 1998.

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, 31–40, 1991.

McGee, T. J., Ferrare, R., Whiteman, D., Butler, J., Burris, J., and Owens, M.: Lidar measurements of stratospheric ozone during the STOIC campaign, J. Geophys. Res., 100, 9255–9262, https://doi.org/10.1029/94JD02390, 1995.

Megie, G., Allain, J., Chanin, M., and Blamont, J.: Vertical profile of stratospheric ozone by lidar sounding from the ground, Nature, 270, 329–331, 1977.

Mertens, C. J., Mlynczak, M. G., López-Puertas, M., Wintersteiner, P. P., Picard, R. H., Winick, J. R., Gordley, L. L., and Russell, J. M.: Retrieval of mesospheric and lower thermospheric kinetic temperature from measurements of CO_2 15 µm Earth Limb Emission under non-LTE conditions, Geophys. Res. Lett., 28, 1391–1394, https://doi.org/10.1029/2000GL012189, 2001.

Météo-France: Nîmes radiosonde data, available at: http://weather.uwyo.edu/upperair/sounding.html, last access: 10 June 2018.

Milton, M. J., Ancellet, G., Apituley, A., Bösenberg, J., Carnuth, W., Castagnoli, F., Trickl, T., Edner, H., Stefanutti, L., Schaberl, T., and Sunesson, A.: Raman-shifted laser sources suitable for differential-absorption lidar measurements of ozone in the troposphere, Appl. Phys. B-Lasers O., 66, 105–113, https://doi.org/10.1007/s003400050363, 1998.

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, https://doi.org/10.5194/amt-5-1301-2012, 2012.

NASA: MLS data, available at: https://disc.gsfc.nasa.gov/datasets?keywords=MLS, last access: 17 June 2018.

NOAA Physical Sciences Laboratory: NCEP/NCAR Reanalysis, available at: ftp://ftp.cpc.ncep.noaa.gov/ndacc/ncep/temp/, last access: 1 June 2018.

Papayannis, A., Ancellet, G., Pelon, J., and Mégie, G.: Multiwavelength lidar for ozone measurements in the troposphere and the lower stratosphere, Appl. Optics, 29, 467–476, https://doi.org/10.1364/AO.29.000467, 1990.

Pelon, J., Godin, S., and Mégie, G.: Upper stratospheric (30–50 km) lidar observations of the ozone vertical distribution, J. Geophys. Res.-Atmos., 91, 8667–8671, 1986.

Rezac, L., Jian, Y., Yue, J., Russell, J., Kutepov, A., Garcia, R., Walker, K., and Bernath, P.: Validation of the global distribution of CO volume mixing ratio in the mesosphere and lower thermosphere from SABER, J. Geophys. Res., 120, 12,067–12,081, https://doi.org/10.1002/2015JD023955, 2015a.

Rezac, L., Kutepov, A., Russell, J., Feofilov, A., Yue, J., and Goldberg, R.: Simultaneous retrieval of T(p) and CO; VMR from two-channel non-LTE limb radiances and application to daytime SABER/TIMED measurements, J. Atmos. Sol.-Terr. Phy., 130-131, 23–42, https://doi.org/10.1016/j.jastp.2015.05.004, 2015b.

Rong, P. P., Russell III, J. M., Mlynczak, M. G., Remsberg, E. E., Marshall, B. T., Gordley, L. L., and López-Puertas, M.: Validation of Thermosphere Ionosphere Mesosphere Energetics and Dynamics/Sounding of the Atmosphere using Broadband Emission Radiometry (TIMED/SABER) v1.07 ozone at 9.6 µm in altitude range 15–70 km, J. Geophys. Res., 114, D04306, https://doi.org/10.1029/2008JD010073, 2009.

Russell III, J. M., Mlynczak, M. G., Gordley, L. L., Tansock Jr, J. J., and Esplin, R. W.: Overview of the SABER experiment and preliminary calibration results, Proc. SPIE Int. Soc. Opt. Eng., 3756, 277–288, 1999.

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 III, J. M., Santee, M. L., Snyder, W. V., Stek, P. C., Thurstans, R. P., Tompkins, A. M., Wagner, P. A., Walker, K. A., Waters, J. W., and Wu, D. L.: Validation of the Aura Microwave Limb Sounder temperature and geopotential height measurements, J. Geophys. Res.-Atmos., 113, D15S11, https://doi.org/10.1029/2007JD008783, 2008.

Sica, R. J. and Haefele, A.: Retrieval of temperature from a multiple-channel Rayleigh-scatter lidar using an optimal estimation method, Appl. Optics, 54, 1872–1889, https://doi.org/10.1364/AO.54.001872, 2015.

Smit, H. G.: Quality assurance and quality control for ozonesonde measurements in GAW, WMO, Geneva, Switzerland, 2013.

Smit, H. G. J., Straeter, W., Johnson, B. J., Oltmans, S. J., Davies, J., Tarasick, D. W., Hoegger, B., Stubi, R., Schmidlin, F. J., Northam, T., Thompson, A. M., Witte, J. C., Boyd, I., and Posny, F.: Assessment of the performance of ECC-ozonesondes under quasi-flight conditions in the environmental simulation chamber: Insights from the Juelich Ozone Sonde Intercomparison Experiment (JOSIE), J. Geophys. Res.-Atmos., 112, D19306, https://doi.org/10.1029/2006JD007308, 2007.

Steinbrecht, W., Neuber, R., von der Gathen, P., Wahl, P., McGee, T., Gross, M., Klein, U., and Langer, J.: Results of the 1998 Ny-

Ålesund Ozone Monitoring Intercomparison, J. Geophys.Res.-Atmos., 104, 30515–30523, 1999.

Steinbrecht, W., Claude, H., Schönenborn, F., McDermid, I. S., Leblanc, T., Godin-Beekmann, S., Keckhut, P., Hauchecorne, A., Van Gijsel, J. A. E., Swart, D. P. J., Bodeker, G. E., Parrish, A., Boyd, I. S., Kämpfer, N., Hocke, K., Stolarski, R. S., Frith, S. M., Thomason, L. W., Remsberg, E. E., Von Savigny, C., Rozanov, A., and Burrows, J. P.: Ozone and temperature trends in the upper stratosphere at five stations of the Network for the Detection of Atmospheric Composition Change, Int. J. Remote Sens., 30, 3875–3886, https://doi.org/10.1080/01431160902821841, 2009a.

Steinbrecht, W., McGee, T. J., Twigg, L. W., Claude, H., Schönenborn, F., Sumnicht, G. K., and Silbert, D.: Intercomparison of stratospheric ozone and temperature profiles during the October 2005 Hohenpeißenberg Ozone Profiling Experiment (HOPE), Atmos. Meas. Tech., 2, 125–145, https://doi.org/10.5194/amt-2-125-2009, 2009b.

Sullivan, J., McGee, T., DeYoung, R., Twigg, L., Sumnicht, G., Pliutau, D., Knepp, T., and Carrion, W.: Results from the NASA GSFC and LaRC Ozone Lidar intercomparison: New mobile tools for atmospheric research, J. Atmos. Ocean. Tech., 32, 1779–1795, https://doi.org/10.1175/JTECH-D-14-00193.1, 2015.

Tarasick, D. W., Davies, J., Smit, H. G. J., and Oltmans, S. J.: A re-evaluated Canadian ozonesonde record: measurements of the vertical distribution of ozone over Canada from 1966 to 2013, Atmos. Meas. Tech., 9, 195–214, https://doi.org/10.5194/amt-9-195-2016, 2016.

Wang, L., Newchurch, M. J., Alvarez II, R. J., Berkoff, T. A., Brown, S. S., Carrion, W., De Young, R. J., Johnson, B. J., Ganoe, R., Gronoff, G., Kirgis, G., Kuang, S., Langford, A. O., Leblanc, T., McDuffie, E. E., McGee, T. J., Pliutau, D., Senff, C. J., Sullivan, J. T., Sumnicht, G., Twigg, L. W., and Weinheimer, A. J.: Quantifying TOLNet ozone lidar accuracy during the 2014 DISCOVER-AQ and FRAPPÉ campaigns, Atmos. Meas. Tech., 10, 3865–3876, https://doi.org/10.5194/amt-10-3865-2017, 2017.

Waters, J. W., Froidevaux, L., Harwood, R. S., Jarnot, R. F., Pickett, H. M., Read, W. G., Siegel, P. H., Cofield, R. E., Filipiak, M. J., Flower, D. A., Holden, J. R., Lau, G. K., Livesey, N. J., Manney, G. L., Pumphrey, H. C., Santee, M. L., Wu, D. L., Cuddy, D. T., Lay, R. R., Loo, M. S., Perun, V. S., Schwartz, M. J., Stek, P. C., Thurstans, R. P., Boyles, M. A., Chandra, K. M., Chavez, M. C., Chen, G.-S., Chudasama, B. V., Dodge, R., Fuller, R. A., Girard, M. A., Jiang, J. H., Jiang, Y., Knosp, B. W., LaBelle, R. C., Lam, J. C., Lee, K. A., Miller, D., Oswald, J. E., Patel, N. C., Pukala, D. M., Quintero, O., Scaff, D. M., Snyder, W. V., Tope, M. C., Wagner, P. A., and Walch, M. J.: The Earth observing system microwave limb sounder (EOS MLS) on the aura Satellite, IEEE T. Geosci. Remote, 44, 1075–1092, https://doi.org/10.1109/TGRS.2006.873771, 2006.

Wing, R., Hauchecorne, A., Keckhut, P., Godin-Beekmann, S., Khaykin, S., McCullough, E. M., Mariscal, J.-F., and d'Almeida, É.: Lidar temperature series in the middle atmosphere as a reference data set – Part 1: Improved retrievals and a 20-year cross-validation of two co-located French lidars, Atmos. Meas. Tech., 11, 5531–5547, https://doi.org/10.5194/amt-11-5531-2018, 2018a.

Wing, R., Hauchecorne, A., Keckhut, P., Godin-Beekmann, S., Khaykin, S., and McCullough, E. M.: Lidar temperature series in the middle atmosphere as a reference data set – Part 2: Assessment of temperature observations from MLS/Aura and SABER/TIMED satellites, Atmos. Meas. Tech., 11, 6703–6717, https://doi.org/10.5194/amt-11-6703-2018, 2018b.