HAL open science

Trajectory matching of ozonesondes and MOZAIC measurements in the UTLS - Part 2: Application to the global ozonesonde network

J. Staufer, J. Staehelin, R. Stübi, T. Peter, F. Tummon, V. Thouret

To cite this version:

J. Staufer, J. Staehelin, R. Stübi, T. Peter, F. Tummon, et al.. Trajectory matching of ozonesondes and MOZAIC measurements in the UTLS - Part 2: Application to the global ozonesonde network. Atmospheric Measurement Techniques, 2014, 7, pp.241-266. 10.5194/amt-7-241-2014 . hal-04114654

HAL Id: hal-04114654

https://hal.science/hal-04114654v1

Submitted on 5 Jun 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

Atmos. Meas. Tech., 7, 241–266, 2014
www.atmos-meas-tech.net/7/241/2014/
doi:10.5194/amt-7-241-2014
© Author(s) 2014. CC Attribution 3.0 License.

Atmospheric Measurement Techniques
Open Access

Trajectory matching of ozonesondes and MOZAIC measurements in the UTLS – Part 2: Application to the global ozonesonde network

J. Staufer[1,*], J. Staehelin[1], R. Stübi[2], T. Peter[1], F. Tummon[1], and V. Thouret[3]

[1]Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland
[2]Federal Office of Meteorology and Climatology MeteoSwiss, Aerological Station, Payerne, Switzerland
[3]Laboratoire d'Aérologie, Université de Toulouse, CNRS, Toulouse, France
[*]now at: Laboratoire des Sciences du Climat et de l'Environnement, CEA-CNRS-UVSQ, IPSL, Gif-sur-Yvette, France

Correspondence to: J. Staufer (johannes.staufer@lsce.ipsl.fr)

Received: 4 June 2013 – Published in Atmos. Meas. Tech. Discuss.: 1 August 2013
Revised: 9 December 2013 – Accepted: 17 December 2013 – Published: 28 January 2014

Abstract. Both balloon-borne electrochemical ozonesondes and MOZAIC (measurements of ozone, water vapour, carbon monoxide and nitrogen oxides by in-service Airbus aircraft) provide very valuable data sets for ozone studies in the upper troposphere/lower stratosphere (UTLS). Although MOZAIC's highly accurate UV-photometers are regularly inspected and recalibrated annually, recent analyses cast some doubt on the long-term stability of their ozone analysers. To investigate this further, we perform a 16 yr comparison (1994–2009) of UTLS ozone measurements from balloon-borne ozonesondes and MOZAIC. The analysis uses fully three-dimensional trajectories computed from ERA-Interim (European Centre for Medium-Range Weather Forecasts Re-analysis) wind fields to find matches between the two measurement platforms. Although different sensor types (Brewer-Mast and Electrochemical Concentration Cell ozonesondes) were used, most of the 28 launch sites considered show considerable differences of up to 25 % compared to MOZAIC in the mid-1990s, followed by a systematic tendency to smaller differences of around 5–10 % in subsequent years. The reason for the difference before 1998 remains unclear, but observations from both sondes and MOZAIC require further examination to be reliable enough for use in robust long-term trend analyses starting before 1998. According to our analysis, ozonesonde measurements at tropopause altitudes appear to be rather insensitive to changing the type of the Electrochemical Concentration Cell ozonesonde, provided the cathode sensing solution strength remains unchanged. Scoresbysund (Greenland) showed systematically 5 % higher readings after changing from Science Pump Corporation sondes to ENSCI Corporation sondes, while a 1.0 % KI cathode electrolyte was retained.

1 Introduction

Over the last 40 yr electrochemical ozonesondes have been widely used for measuring ozone (O_3) up to the burst of the balloon at altitudes of 30–35 km. Electronically coupled with a standard meteorological radiosonde for data transmission to a ground receiver, they provide accurate measurements of O_3, with a typical vertical resolution of 100–200 m. Ozonesondes provide unique information that can be used to produce O_3 climatologies, validate satellite measurements, establish long-term atmospheric changes and trends, and for comparison with numerical model simulations.

Three main types of electrochemical ozonesondes have been developed since the 1960s: the Brewer-Mast (BM, Brewer and Milford, 1960), the Electrochemical Concentration Cell (ECC, Komhyr, 1969) and the Japanese ozonesonde (KC, Kobayashi and Toyama, 1966). At present, most sites use ECC sondes, and, since 2010, KC ozonesondes are no longer used operationally. The principle of operation is based on the titration of O_3, either in a potassium iodide (KI) sensing solution (ECC and BM sondes) or in a potassium bromide solution (KC sondes) (Smit et al., 2011). For each molecule of O_3 entering the solution, two iodide ions (I^-) are oxidised to form iodine (I_2), which is subsequently reduced back to I^- at the electrodes, generating an electric current of a few microamperes. This current is measured, and by assuming a 100 % reaction yield, can directly be related to the atmospheric O_3 partial pressure. Uncertainties may change during flight as the pump efficiency degrades with increasing altitude, or due to inaccurate pump temperature measurements or the presence of a background current that is subtracted from the measured current (Smit et al., 2007). The

background current has largest influence on the overall accuracy at low O_3 concentrations and therefore becomes particularly important in the tropical troposphere and below the mid-latitude tropopause (e.g. Smit et al., 2011). Conversely, the pump efficiency becomes the predominant uncertainty in the stratosphere (e.g. Stübi et al., 2008).

Although the primary principle of operation has not changed, ozonesondes have undergone several modifications, including changes to manufacturing, preparation, solution concentration and data processing, all of which may have affected the accuracy of the various sonde types and in turn the long-term trends estimated using these data (Smit et al., 2007). Over the past decades various research groups have put considerable effort into quantifying the precision and accuracy of ozonesondes, including balloon experiments using a multiple-instrument gondola (e.g. Hilsenrath et al., 1986; Deshler et al., 2008), dual flights (De Backer et al., 1998; Kivi et al., 2007; Stübi et al., 2008) and environmental chamber simulations (Smit et al., 2007; Thompson et al., 2007). A quantitative assessment of ozonesonde data quality is currently under way, following guidelines prepared by the ozonesonde data quality assessment panel (part of the SPARC/IO_3C/IGACO-O_3/NDACC initiative on "Past Changes in the Vertical Distribution of Ozone").

Comparison with continuous records from other instruments, for example, space-borne (e.g. Liu et al., 2006; Terao and Logan, 2007; Labow et al., 2013), ground-based (e.g. SPARC/IOC/GAW, 1998; Thompson et al., 2003a; Logan et al., 2012) or other aircraft-borne in situ measurements (Thouret et al., 1998; Schnadt Poberaj et al., 2009; Staufer et al., 2013), can also provide information about potential long-term changes in the performance of ozonesondes. Liu et al. (2006) demonstrated that, particularly for tropical stations, the variations in the bias to GOME (Global Ozone Monitoring Experiment) and SAGE-II (Stratospheric Aerosol and Gas Experiment II) found between different launch sites greatly depend on ozonesonde techniques, instrument type, sensor solution, and the total ozone normalisation. The quality of tropospheric data from earlier European BM sondes has been questioned by Schnadt Poberaj et al. (2009) and recently also by Logan et al. (2012), while stratospheric sonde data appear to have had few problems during the 1980s and 1990s (e.g. Terao and Logan, 2007).

Commercial airliners have also been used to provide high-quality tropospheric and lower stratospheric O_3 measurements, for example, as part of the MOZAIC aircraft program (Measurements of ozone, water vapour, carbon monoxide and nitrogen oxides by in-service Airbus aircraft, Marenco et al., 1998) or the Swiss NOXAR program (Nitrogen Oxides and Ozone along Air Routes, Brunner et al., 2001). For data in both programs, European long-range airliners are equipped with accurate UV photometers to measure O_3 and other trace gases. MOZAIC data are available from August 1994 onwards, while NOXAR data are available for 1995-97. In a companion paper (Staufer et al., 2013) we used both data sets to analyse ozonesonde data from Payerne (Switzerland) by using fully three-dimensional trajectories to find commonly sampled air masses. Comparison of the Payerne sonde data with MOZAIC showed mean differences of up to 20 % between 1994–1997, followed by differences of around 5–10 % in the subsequent years (1998–2009). The comparison of sonde data with the NOXAR data, however, showed a smaller offset of around 15 % from 1995–1997. The question arises as to whether these discrepancies indicate a small drift in the MOZAIC calibration or whether they are a particular feature of the Payerne data series. To answer this question, the analysis of Staufer et al. (2013) is extended to various other soundings sites in Europe, America, Japan, and Africa.

2 Data and methodology

2.1 Ozonesondes

The meteorological observatory at Hohenpeißenberg (MOHp), Germany, is the only ozonesonde station that continues to use BM sondes (manufactured by Mast Keystone Corporation, Reno, NV, USA), whereas Uccle (Belgium) and Payerne switched to ECC sondes in April 1997 and September 2002, respectively. KC sondes have only ever been flown at Japanese sites. ECC sondes are manufactured either by Science Pump Corporation (SP; model type 5A and 6A), or, since the early nineties, by the Environmental Science Corporation (ES; model type Z). In 2011 ES was taken over by Droplet Measurement Technologies. Originally, ES sondes were operated with a 1.0 % fully buffered KI cathode sensing solution, but after the environmental chamber tests of JOSIE (Juelich Ozone Sonde Intercomparison Experiment, Smit et al., 2007), the manufacturer recommended diluting the solution by half, to 0.5 % KI. This led some groups to change their technique (see Table 1).

ECC ozonesondes prepared according to the manufacturer's recommendations (SP, 1.0 % KI, ES 0.5 % KI) typically measured 5 % higher ozone mixing ratios in the (mid-latitude) upper troposphere/lower stratosphere (UTLS) compared to UV-photometers during the JOSIE and BESOS (Balloon Experiment on Standards for Ozonesondes) campaigns (Smit et al., 2007; Deshler et al., 2008). The ES sondes using a 1.0 % KI give 10–15 % higher ozone concentrations compared to a UV-photometer, while SP sondes prepared with 0.5 % KI agree within ±5 %, but underestimate the ozone column (Smit et al., 2007; Deshler et al., 2008). Thus, stations changing the ECC manufacturer without changing the cathode sensing solution strength accordingly can introduce changes of more than ±5 % in their records.

The US National Oceanic and Atmospheric Administration (NOAA) sites (in this study Boulder, Colorado and Huntsville, Alabama) experimented with different solution

Table 1. Overview of the sonde types used, data processing methods applied for the MOZAIC period (August 1994–March 2009), and soundings available for comparison (see text). Fields are left blank when information is missing or redundant. SST denotes the cathode sensing solution strength of the ECC sondes, whereby SST 1.0 denotes the fully buffered 1.0 % KI solution, SST 0.5 the half-buffered 0.5 % KI, SST 2.0 the unbuffered 2.0 % KI and SST 1.0b the 1.0 % KI, 1/10th buffered solution. MOHp denotes the Meteorological Observatory at Hohenpeißenberg, Germany, OHP the Observatory Haute Provence, France. W denotes the WOUDC archive, N the NDACC archive, and S the SHADOZ archive.

Station	Archive	lat	lon	Sonde type	SST	Background signal	Switch date	Scaled to column	Number of ascents total	Number of ascents used
Alert	W	82.5	−62.3	SP, ES	1.0	declining	2004	Yes	464	141 (30 %)
				ES					269	100 (37 %)
Boulder	N	39.9	−105.2	ES	1.0	constant	21 Aug 1997	No	160	101 (63 %)
					2.0				413	251 (61 %)
					1.0b		30 Nov 2005		210	107 (51 %)
Churchill	W	58.7	−94.1	SP, ES	1.0	declining	2004	Yes	419	209 (50 %)
				ES					209	132 (63 %)
De Bilt	N	52.1	5.2	SP	1.0	declining	1 Nov 1998	No	247	223 (90 %)
						constant			540	483 (89 %)
Edmonton	W	53.5	−114.1	SP, ES	1.0	declining	2004	Yes	461	318 (69 %)
				ES					256	171 (67 %)
Eureka	W	79.9	−85.9	SP, ES	1.0	declining	2004	Yes	670	151 (23 %)
				ES					356	130 (37 %)
Goose Bay	W	53.3	−60.3	SP, ES	1.0	declining	2004	Yes	424	313 (74 %)
				ES					247	176 (71 %)
Huntsville	W, NOAA	34.7	−86.7	ES	2.0	constant	1 Mar 2006	No	317	201 (63 %)
					1.0b				183	98 (54 %)
Irene	S	−25.9	28.2	SP	1.0				231	32 (14 %)
Izaña	N	28.5	−16.3	SP	1.0	constant		No	840	533 (63 %)
Legionowo	W	52.4	21.0	SP	1.0	declining		No	884	783 (89 %)
Lerwick	W	60.1	−1.2	SP, ES	1.0	declining			719	592 (82 %)
Lindenberg	W	52.2	14.1	SP	1.0	declining		Yes	748	500 (67 %)
Madrid	W	40.8	−3.6	SP	1.0	declining			538	413 (77 %)
MOHp	W	47.8	11.0	BM				Yes	1889	1335 (71 %)
Naha	W	26.2	127.7	KC		declining	13 Nov 2008	Yes	579	159(27 %)
				ES	0.5				12	5 (42 %)
Nairobi	W	−1.3	36.8	ES	1.0	declining			519	103 (20 %)
Natal	S, W	−5.4	−35.4	SP	0.5		1 Apr 1999		46	16 (35 %)
					1.0				389	44 (11 %)
OHP	N	43.9	5.7	SP	1.0	declining	1 Mar 1997	No	51	47 (92 %)
				ES					541	424 (78 %)
Paramaribo	S	5.8	−55.2	SP	1.0	declining			380	82 (22 %)
Payerne	W	46.7	6.6	BM			1 Sep 2002	Yes	1288	1065 (83 %)
				ES	0.5	constant		Yes	1009	834 (83 %)
Resolute	W	74.7	−94.9	SP, ES	1.0	declining	2004	Yes	304	83 (27 %)
				ES					205	82 (40 %)
Sapporo	W	43.1	141.3	KC		declining		Yes	647	415 (64 %)
Scoresbysund	N	70.5	−22.0	SP	1.0	constant	13 Jul 2001	No	397	284 (72 %)
				ES					341	243 (71 %)
Sodankylä	N, NILU	67.4	26.6	SP, ES	1.0	declining	25 Jan 2006	No	796	514 (65 %)
				ES	0.5	constant			266	115 (43 %)
Tsukuba	W	36.1	140.1	KC		declining		Yes	793	374 (47 %)
Uccle	W	50.8	4.4	BM			1 Apr 1997	Yes	385	346 (90 %)
				ES	0.5	constant		Yes	1775	1536 (87 %)
Wallops Island	N, W	37.9	−75.5	SP	1.0	constant			874	594 (68 %)

strengths and buffers (2.0 % KI unbuffered solution, Johnson et al., 2002) and recently with a 1.0 % KI 1/10th buffered solution. However, the solutions and techniques used by NOAA are unique and changes are carefully monitored to assure continuity of the record. Therefore, the concern is not with NOAA but with those stations that have used one of the standard buffered solutions, 1.0 % or 0.5 %, but not with the right ECC ozonesonde. NOAA also changed instruments and techniques at the Pacific tropical stations it operates (Pago Pago, American Samoa; Hilo, Hawaii; Suva, Fiji; San Cristobal, Ecuador). These stations, because they are not near MOZAIC flight routes, are not included in the present paper.

In this study we analyse 11 stations (Alert, Churchill, Edmonton, Eureka, Goose Bay, and Resolute in Canada; Lerwick, UK; Natal, Brazil; Observatory Haute Provence (OHP), France; Scoresbysund, Greenland; and Sodankylä, Finland) that switched from ES to SP (or vice versa) and/or have operated both sonde types with or without changing the solution strength. At Canadian stations both SP and ES-Z sondes were flown before 2004. After 2004 mainly ES sondes were launched but the 1.0 % solution strength was retained (see metadata at WOUDC, the World Ozone and Ultraviolet Radiation Data Centre).

Recently, the ozonesonde data user community has been addressing how to account for changes in radiosonde instrumentation that have accompanied ozonesonde changes at a number of stations in the past 5 yr or so (Stauffer et al., 2014). The radiosonde changes propagate to each ozone measurement, but mostly at pressures < 100 hPa; newer radiosondes mostly affect ozone data after 2009. Thus, these influences are neglected here.

The standard operating procedures (SOP) for the BM sondes are defined by Claude et al. (1987) and have been followed by MOHp and Payerne. Payerne has higher correction or scaling factors (CF) than MOHp, probably because the pump temperature was assumed constant at 280 K instead of 300 K (see also Jeannet et al., 2007). The CF is determined as the ratio between the total O_3 column measured by the ozonesonde and a nearby independent column measurement, such as from Dobson or Brewer spectrometers. Uccle data, as used here, are normalised following De Backer (1999) rather than utilising the SOPs.

The background current is a major error source for ozonesonde measurements in the upper troposphere, where ozone concentrations are small. The SOP for the BM sondes does not call for correction of the background current. For ECC sondes the conventional correction is to assume that the background current is proportional to the oxygen partial pressure and thus declines with altitude (Komhyr, 1986). This, however, is neither supported by lab studies (e.g. Thornton and Niazy, 1982), nor by the the study of Reid et al. (1996), who found a 7–8 % better agreement between ECC ozonesondes and an UV photometer for tropospheric O_3 concentrations when a constant background current was assumed. The recent assessment of ECC SOPs calls for a constant background current (Smit et al., 2011). The background current (ib) is measured three times during the prelaunch procedure: once the sondes are exposed to purified (ozone-free) air (ib_1), once after exposure to O_3 (ib_2), and just prior to flight (ib_3). We contacted the principle investigators (PI) in case the treatment of the background current could not be extracted from the different archives. Only De Bilt (the Netherlands), Huntsville, Legionowo (Poland), Madrid (Spain), OHP, Payerne, Sodankylä, and Uccle report their background current values without large data gaps for 1994–2009 (i.e. a few months per year or maximally one whole year is missing). Scoresbysund and Canadian stations have reported background currents to the archives since 2000. Due to the limitations of the Lagrangian match technique presented here (cf. Sects. 2.3 and 3), the influence of the background current on the sonde performance can be thoroughly investigated only at De Bilt, Legionowo, Payerne and Uccle. The latter three, however, report no change in treatment nor large variations of the background current in the 1994–2009 period.

Ozonesonde data can be downloaded from several archives: ftp servers at the WOUDC (World Ozone and Ultraviolet Radiation Data Centre), NDACC (Network for the Detection of Atmospheric Composition Change), SHADOZ (Southern Hemisphere ADditional OZonesondes), NILU (Norwegian Institute for Air Research), and NOAA (National Oceanic and Atmospheric Administration). Records from most stations can be found either on the WOUDC or NDACC homepages, or on both. Most tropical stations are now part of the SHADOZ network (Thompson et al., 2003a, b). NILU offers campaign data, for example, measurements from the VINTERSOL campaign (European field campaign studying stratospheric ozone), and high latitude station data. The Intensive Ozonesonde Network Study (IONS) experiments (Tarasick et al., 2010; Thompson et al., 2011) over North America have operated in 2004, 2006, 2008, and 2013. This has augmented regular US launches at Boulder (Colorado), Huntsville (Alabama), and Wallops Island (Virginia) as well as most of the Canadian stations listed in Table 1. The data are archived at NASA/Langley and WOUDC.

For some stations we needed to switch between the archives to obtain the highest number of soundings. We found that the archives do not necessarily contain the same number of soundings for the same period. Some years are missing at one archive but available at another. Some years are also missing in all archives. Data from Boulder, De Bilt, Izaña (Canary Islands), MOHp, OHP, and Sodankylä were obtained from the NDACC archive in March and April 2010. Sodankylä data from 2004 onwards were obtained from NILU's NADIR database in May 2010. Data from the Canadian sites (Alert, Churchill, Edmonton, Eureka, Goose Bay, Resolute), as well as for Huntsville, Legionowo, Lerwick, Lindenberg, Madrid, Naha, Nairobi, Payerne, Uccle, and Sapporo and Tsukuba (both in Japan) were also obtained from the WOUDC in March and April 2010. Uccle data from

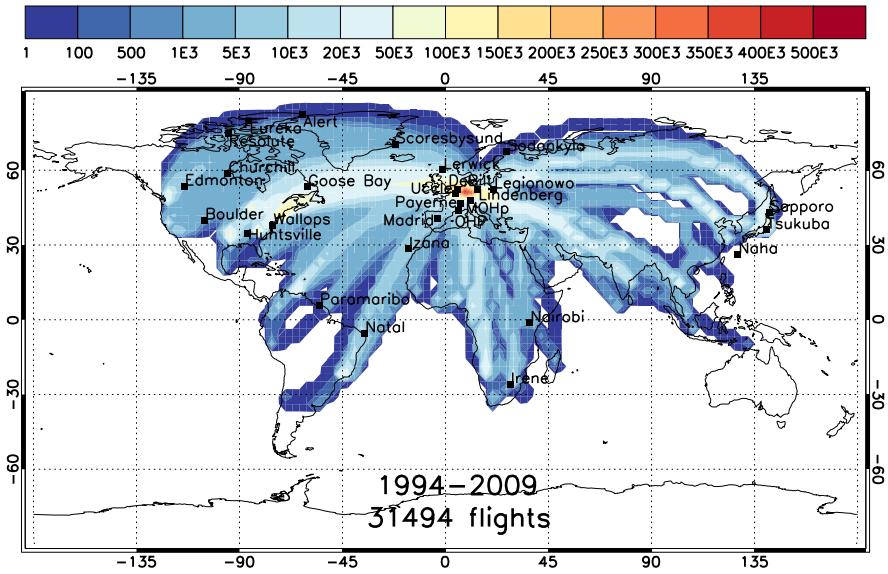

Fig. 1. Distribution of the ozonesonde stations considered in this work (black squares) and total number of measurements from the MOZAIC aircraft program (1994–2009), summed up over a $3° \times 3°$ latitude × longitude grid. The colour bar indicates the total number of 1 min averaged MOZAIC measurements. MOHp denotes the Meteorological Observatory at Hohenpeißenberg, Germany; OHP the Observatory Haute Provence, France.

2007 onwards were obtained in February 2011. Huntsville data for 2008 and 2009 were obtained from a NOAA ftp-server in February 2011. Irene (So. Africa), Natal (Brazil) and Paramaribo (Suriname) data were obtained from the SHADOZ database in April 2010, with the exception of data from the Natal site for 1997, which was obtained from the WOUDC. Wallops Island data for 1994, 1995, 2008, and 2009 were obtained from the WOUDC in March and April 2010, while all remaining data were obtained from the NDACC website in April 2010.

MOHp, Payerne, and Uccle typically launch two to three ozonesondes per week, whereas most other sites typically launch one sonde per week. It is important to note that not all sites flew ozonesondes for the entire MOZAIC period, with some stations starting later, particularly tropical ones. Consequently, the total number of launches for the entire MOZAIC period (August 1994–March 2009) is quite different from station to station, ranging from 300–400 (e.g. Paramaribo, Irene) to more than 2000 launches (e.g. MOHp, Payerne, Uccle, cf. Table 1).

2.2 MOZAIC ozone observations

The MOZAIC program and data from it are described and analysed in detail by Thouret et al. (1998). Here, only the main characteristics are summarised. Dual-beam UV absorption models from Thermo Environment were installed on several commercial aircraft participating in the MOZAIC project. These UV photometers have a response time of 4 s, a detection limit of 2 ppbv, and an uncertainty of ±[2 ppbv + 2 %]. For example, for $O_3 = 100$ ppbv this results in an uncertainty of ±4 ppbv. The quality assurance and control procedures have not changed since the project started in 1994. MOZAIC analysers are inspected annually and periodically calibrated (about every 12 months) with a reference analyser at the French National Institute of Standards and Technology. Additionally, the analysers are checked in-flight with a built-in ozone generator to detect any drift in instrument efficiency. MOZAIC is considered as a standard reference (for example in Thouret et al., 1998, 2009; Schnadt Poberaj et al., 2009) due to its regular inspections and checks as well as the unchanged quality assurance procedures. However, the recent analyses of both Logan et al. (2012) and Staufer et al. (2013) cast some doubts on the long-term stability of the MOZAIC ozone data.

MOZAIC's main flight route is the North Atlantic flight corridor, but aircraft also fly to airports in South America, East Asia, and Southern Africa. The flight distribution of the aircraft is shown in Fig. 1. The sounding sites investigated in this work are chosen according to these flight routes. In total, 31 494 flights were available when we downloaded the data (March 2010), covering the period from August 1994 to March 2009. We use 1 min averaged MOZAIC data, which correspond to a horizontal resolution of 10 to 15 km at cruise altitude.

2.3 Comparison methodology

For the comparison between routinely flown ozonesondes and ozone measurements from any MOZAIC aircraft we use

trajectories to ensure both instrument platforms observe the same air mass. In a companion paper (Staufer et al., 2013), we test and apply this method of comparisons between aircraft measurements from both MOZAIC and NOXAR aircraft and ozonesonde data from Payerne, Switzerland. Here, we summarise just the main points of this method, which is similar to the trajectory match technique used by Rex et al. (1998) or the trajectory hunting technique described by Danilin et al. (2002). After reconstructing the sonde's flight path using wind data (speed and direction) from the radiosonde, the trajectory tool LAGRANTO (Wernli and Davies, 1997) is used to calculate 6 day forward and backward trajectories for each sounding. Fully three-dimensional trajectories are used because it has been shown, for example by Stohl and Seibert (1998), that they are more accurate than kinematic isentropic or isobaric trajectories in the troposphere. In the stratosphere, isentropic trajectories are of similar accuracy to fully three-dimensional trajectories.

LAGRANTO is forced with six-hourly wind fields from ECMWF's (European Centre for Medium-Range Weather Forecasts) ERA-Interim reanalysis (1° horizontal resolution, 61 vertical levels). Trajectories ascending or descending by more than 450 hPa during the six days simulated are excluded to avoid air masses that transport polluted boundary layer air or air from deep stratospheric intrusions. For each trajectory all MOZAIC measurements matching the trajectory within $r \leq 75$ km and $\Delta\Theta \leq 0.6$ K are collected, then a weighted mean of the aircraft observations is calculated and compared to the ozonesonde measurements at initialisation of the trajectories. For the weighting a time lag compared to the soundings is used to account for the reduced accuracy of trajectories further away in time.

To assess the uncertainty of this technique, Staufer et al. (2013) checked the method for comparison of one instrument type with itself, i.e. MOZAIC–MOZAIC self-matches. Assuming that MOZAIC is noise-free, they found mean differences of ±2 %. However, it is important to note that this uncertainty was derived under most favourable conditions, namely a large number of matches and most matches found within the first 50 h of the trajectories. The uncertainty is expected to increase when most matches are found after 50 h since trajectory errors typically accumulate with time. Thus, the temporal distribution of the matches is a limiting factor for this comparison. This issue is discussed in more detail in the beginning of Sect. 3.

As shown by Staufer et al. (2013), the combination of forward and backward trajectories can be used to account for the potential effects of chemistry and mixing along the trajectory paths. These effects are typically more pronounced in the upper troposphere (UT) than in the lower stratosphere (LS). At Payerne, sonde biases of up to 10 % between forward-only and backward-only trajectories were found. Staufer et al. (2013) further showed that tropospheric photochemistry cannot alone account for the observed differences between forward and backward trajectories. Other factors such as the different temporal match distribution between the unidirectional trajectories or an inaccurate quantification of the meteorological conditions additionally contribute to the observed differences. For the data considered, very few trajectories were matched in both directions and Staufer et al. (2013) needed to analyse forward and backward trajectories separately. They showed that by surrounding each trajectory with four additional trajectories, each displaced by 0.5° latitude and longitude from the central trajectory, the biases could be reduced by half. Furthermore, the sonde bias at Payerne was found to be largely insensitive to the trajectory duration (one or six days). Due to this robustness, and because some sites do not allow reconstructing the balloon flight path since no wind direction or speed data are available, we use all trajectories (the central plus the displaced trajectories), referred to as the combined trajectory set, unless otherwise mentioned.

For tropical stations ($\phi < 30°$ N), trajectories were determined every 1 K in potential temperature at altitudes between 5–15 km. This is in contrast to the mid- and high latitude stations, where similar to Staufer et al. (2013), the trajectories originate every 5 hPa within the UTLS, which is defined as ±125 hPa around the local (lapse-rate defined) tropopause. Without this adaptation, no matches with MOZAIC would have been obtained because the tropical tropopause is much higher than the typical MOZAIC cruise altitude (8–12 km).

3 Ozonesonde comparisons with MOZAIC

Results are first presented as averages over the entire MOZAIC period; thereafter the differences between sonde and MOZAIC, $\Delta O_3 = 2(\text{sonde} - \text{MOZAIC})/(\text{sonde} + \text{MOZAIC})$, are analysed in more detail by separately discussing the behaviour and changes of ΔO_3 in both the UT and LS. We first focus on the mid-latitudes (30° N $\leq \phi \leq$ 60° N) where most stations are located, then show results for stations at high northern latitudes ($\phi > 60°$ N) and for tropical and southern latitudes ($\phi < 30°$ N).

Results of the comparisons are limited by the number of matches, which in turn depends on the number of ascents, the location of the station, and the MOZAIC flight paths (see Fig. 1, and Tables 1 and 2). The number of matches per launch site varies considerably. Table 1 shows, for example, that for the Swiss station Payerne, 1899 ascents can be compared to MOZAIC, while for Natal, which is located in Brazil, only 60 ozonesondes are available for comparison. Thus, stations like Payerne can be analysed in much more detail than stations like Natal.

Another important factor limiting the comparison is the temporal distribution of the MOZAIC matches. As already mentioned in Sect. 2.3, the ±2 % uncertainty of this Lagrangian matching approach derived by Staufer et al. (2013) using MOZAIC–MOZAIC self-matches was obtained with most matches found within 50 h of the trajectories. Figure 2 shows that many stations, particularly those in the tropics,

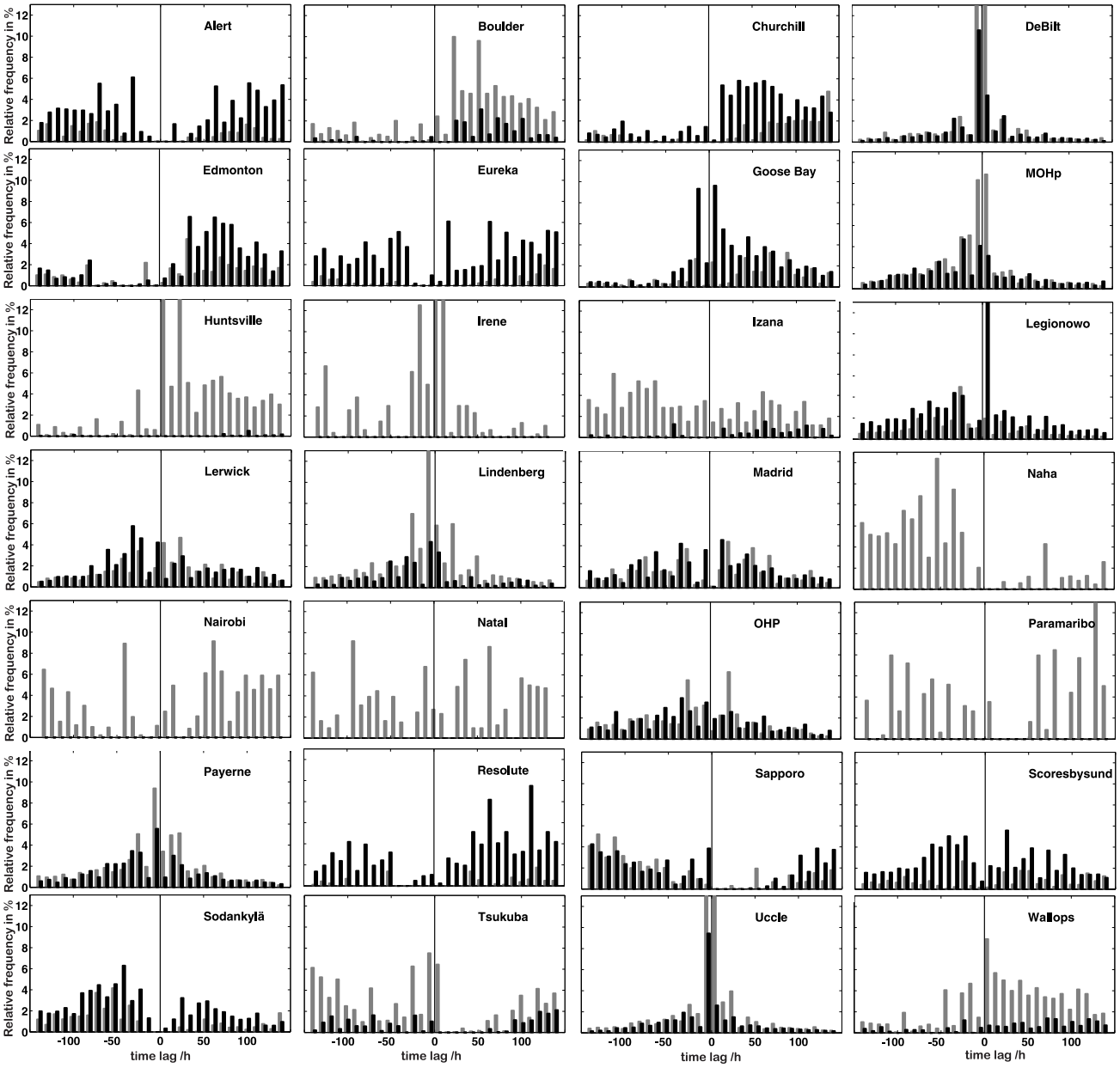

Fig. 2. Temporal distribution of the number of matches at different stations. The time lag is positive for forward-only trajectories and negative for backward-only trajectories. Matches in the stratosphere are shown in black, while matches in troposphere are shown in gray. The bin size is 10 h.

southern latitudes, or in Japan, have most matches after 50 h. Staufer et al. (2013) tried to systematically assess the uncertainty for different time lags but the uneven distribution of the MOZAIC self-matches, i.e. hardly any matches after 50 h, prevented them from doing so. However, by excluding matches from the first 24 h, uncertainty increased by 1–2 %, particularly in the UT, indicating an increasing uncertainty with increasing trajectory duration (Staufer et al., 2013). Thus, stations that are either closer to highly frequented MOZAIC airports or whose balloon trajectories cross the main MOZAIC flight path over the North Atlantic are favoured by this Lagrangian approach. These favoured stations are mainly located in Europe. For stations outside of Europe, larger errors bars in ΔO_3 certainly reflect the increasing uncertainty of the match technique and the lower number of matches.

3.1 Results for mid-latitude stations

3.1.1 MOHp/Payerne/Uccle

During the MOZAIC period (1994–2009) only three stations flew BM ozonesondes. Uccle and Payerne switched to ECC sondes (ES-Z type operated with 0.5 % KI half buffered sensing solution) in April 1997 and September 2002, respectively; Hohenpeißenberg (MOHp) still launches BM sondes, and thus is the only remaining BM station worldwide. Sample sizes from these three stations are largest because 2–3 sondes are flown per week at each of these sites, and the stations are located close to the main MOZAIC airports and flight routes. The percentage of matched ozonesondes is largest at Uccle (≈ 90 %) and smallest at MOHp (71 %). This disparity is related partly to the location of Uccle, which lies close to Brussels airport and the main flight route of aircraft over northern and western Europe (see Fig. 1). In addition, MOHp data from the WOUDC are reported on fewer pressure levels than at Payerne and Uccle.

The 16 yr mean O_3 concentrations from sondes and MOZAIC are shown in Fig. 3a. They show a striking agreement. The height of the lapse-rate defined tropopause is derived only from the sondes used in this comparison, resulting in a mean tropopause pressure of 250 hPa. Sonde–MOZAIC differences obtained from the unidirectional trajectories are negligibly small at all altitudes for both MOHP and Uccle, but range up to 5 % at Payerne (see also Staufer et al., 2013), although the differences in absolute concentrations are on the order of a few ppb in the troposphere. For all stations, the O_3 concentrations obtained from forward-only trajectories are systematically higher than from backward-only trajectories.

Figure 3b shows the mean relative differences between sonde and MOZAIC split into three different periods. The first period, 1994–1997, is characterised by large differences between BM sondes and early MOZAIC observations. At MOHp the sondes exceed MOZAIC by 10–15 %, while at Payerne and Uccle even higher offsets are found (up to 25 % in the vicinity of the tropopause). Mean 1994–1997 differences are lowest (5 % at MOHp and 10 % at Payerne, Uccle) at the 175 hPa level, where mean O_3 concentrations are on the order of 200 ppb. After 1997/1998 the mean differences drop to less than 10 % at all three stations, at all altitudes.

Figure 3c–e show the time series of 13 month moving average monthly mean differences. The lower stratosphere (Fig. 3c) includes only trajectories where the difference in pressure between the trajectory at initialisation ($p(t=0)$) and the tropopause pressure (p_{TP}) is smaller than 15 hPa. Figure 3d contains a narrow tropopause band with $|p(t=0)-p_{\mathrm{TP}}|<15$ hPa. The values chosen are similar to those of Thouret et al. (2006), who also considered a 30 hPa thick tropopause zone. The time series of the tropopause differences is more uncertain and more variable because the strongest O_3 gradients are typically found in the vicinity of the tropopause. Figure 3e comprises the upper troposphere and includes all trajectories satisfying $p(t=0)-p_{\mathrm{TP}} \geq 15$ hPa. All calculations follow the methodology laid out in Staufer et al. (2013); however, 50 hPa pressure intervals are used here instead of 1 km altitude bins.

At MOHp, the CF corrects for the low BM sonde bias in the LS, except for the 1994–1997 period when the application of the CF results in a high bias compared to MOZAIC. For the UT, application of the CF is counterproductive for almost all periods. In contrast, at Payerne the agreement with MOZAIC in both the UT and LS is better when no scaling is applied. Whereas Stübi et al. (2008) recommended scaling both sonde types to column O_3, Staufer et al. (2013) suggest that the transition from BM to ECC sondes is smoother when the BM sondes remain unscaled, at least for the LS as defined here. The homogenised Uccle and MOZAIC data show differences of less than 5 % in the LS (Fig. 3c), but the homogenisation does not remove the high offset in the mid-1990s in the troposphere (Fig. 3e). However, their CF, whose calculation differs from the usual approach, reduces the bias compared with MOZAIC to 5 % in the UT after 1996. Our analysis qualitatively confirms results from Schnadt Poberaj et al. (2009) for the European ozonesonde stations. Furthermore, our analysis shows that the mean discrepancies at Uccle from 1994–2001 can be traced back to the use of BM sondes. Our results for the free troposphere ($p > 430$ hPa) also qualitatively agree with the recent study of Logan et al. (2012), who found that the tropospheric portion of BM sonde data before 1998 should be discarded for trend analysis due to the mismatch with MOZAIC and long-term ozone measurements from alpine sites. The anomalous peak in the Uccle tropospheric data in 2007 is present in our analysis, although in 2002 a peak of similar magnitude is also found. However, note that the peak in 2002 is not present in the UT, when trajectories satisfying $p(t=0)-p_{\mathrm{TP}} \geq 30$ hPa are used instead (not shown). The high bias in UT O_3 compared to MOZAIC observed at all three BM sonde sites for the 1994–1997 period remains unexplained.

3.1.2 De Bilt

In contrast to Payerne and Uccle, at De Bilt SP sondes operated with a 1.0 % KI cathode electrolyte have been used during the entire MOZAIC period. A smaller number of ozone soundings is available from this launch site compared to MOHp, Payerne, and Uccle because typically just one sounding is launched per week. Despite this, 90 % of the soundings could be matched with MOZAIC (Table 1).

Similar to MOHP, Payerne, and Uccle, the 16 yr mean O_3 concentrations from both De Bilt and MOZAIC agree to within 10 % at all altitudes (Fig. 4a). However, the difference between ozonesondes and MOZAIC shows a distinct time dependence, ΔO_3 amounts to 15 % in the LS and 20 % in the UT from 1995–1996, then slowly decreases to below 0 % by the end of the 1999, and then increases again to up to 10 % after 2003 (Fig. 4c and d). This is likely related to the

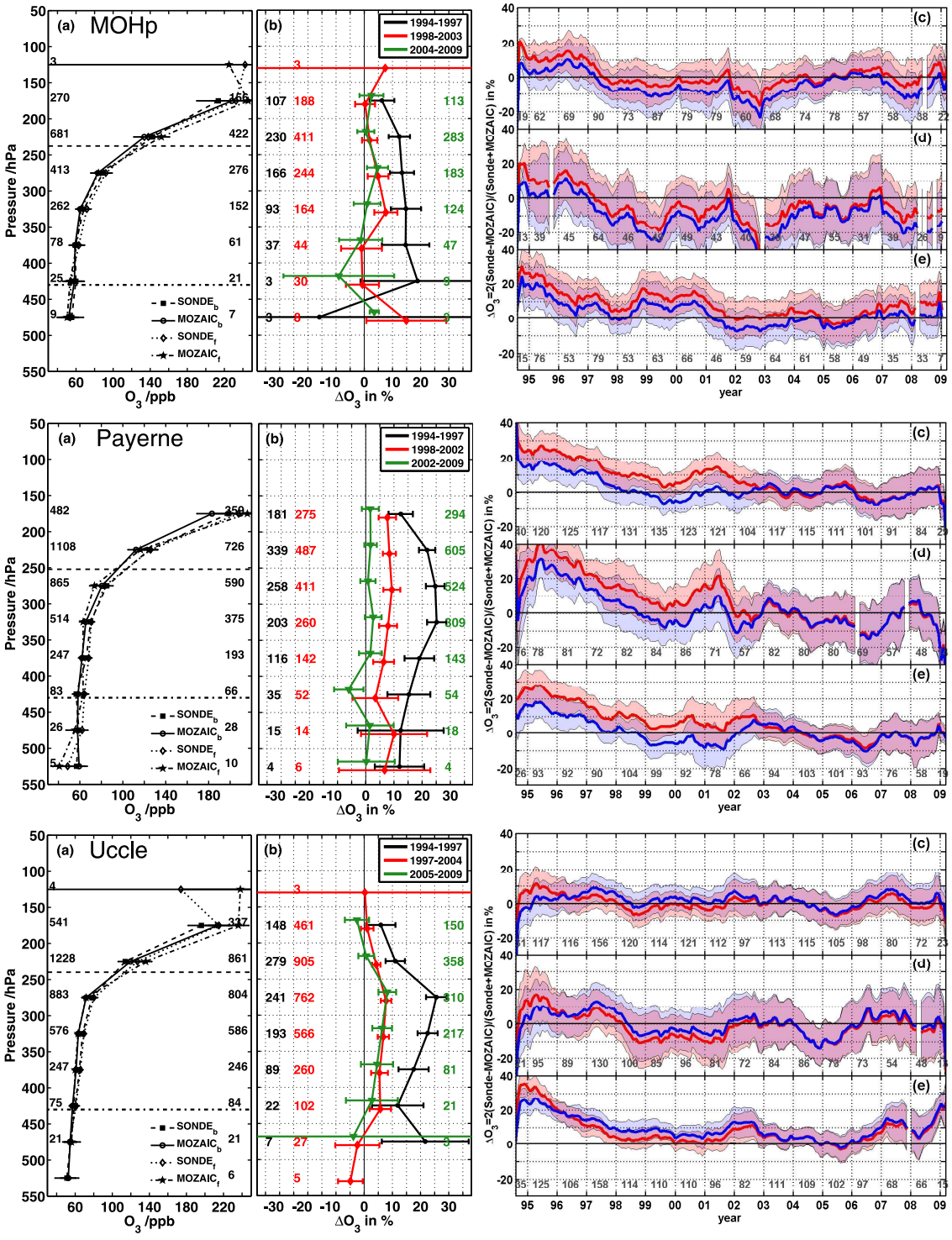

Fig. 3. Comparison between MOZAIC O_3 measurements and ozonesondes from MOHp, Payerne, and Uccle. **(a)** 16 yr average O_3 profiles from sonde and MOZAIC binned into 50 hPa layers. Numbers on the left and right denote the number of soundings using 6 day backward-only and forward-only trajectories, respectively. The dashed horizontal line denotes the tropopause, the dash-dotted horizontal line the level up to which Logan et al. (2012) compared ozonesondes with MOZAIC. **(b)** Relative differences $\Delta O_3 = 2(\text{sonde} - \text{MOZAIC})/(\text{sonde} + \text{MOZAIC})$ split into three periods. The number of sondes available for comparison is displayed for each period on the sides. Time series of ΔO_3 with CF (red) and without CF (blue) are displayed for the LS **(c)**, a narrow tropopause band **(d)**, and the UT **(e)**. Numbers at the bottom indicate the number of soundings used for calculating monthly mean differences. The error bars in panels **(a)** and **(b)** denote the 90 % confidence of the median, while in **(c)**–**(e)** the shaded areas denote the standard error (68 % confidence). Overlapping areas are displayed in light purple.

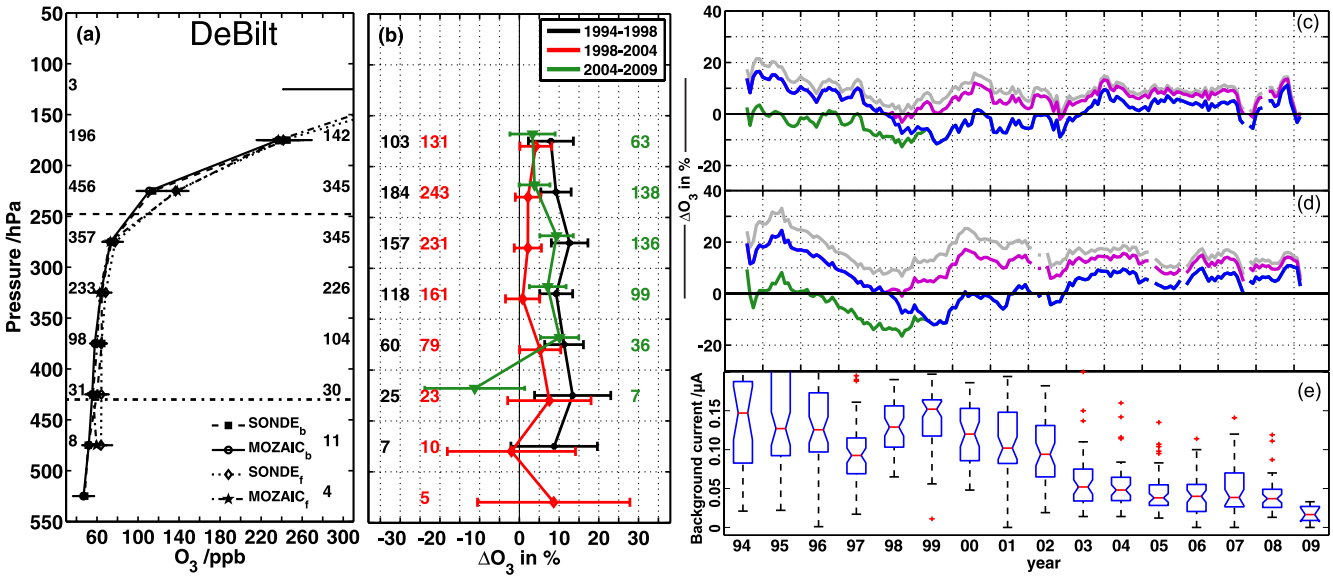

Fig. 4. Comparisons of De Bilt ozonesondes with MOZAIC measurements. **(a)** and **(b)**: As for Fig. 3a–b. **(c)** and **(d)**: time series of ΔO_3 for the LS **(c)** and UT **(d)**, respectively. Time series are shown for the difference resulting from using the operating procedures described in Table 1 (blue line), with no correction for background signal (gray line), with a constant background current correction (green line), and with the application of an altitude-dependent background correction (purple line). **(e)**: annual statistical distribution of the background current used to process the ozonesonde data. The median is depicted as a horizontal solid red line, the 90 % confidence interval as the notches. The box limits correspond to the 75 % and 25 % quartiles. The whiskers extend out to the maximum or minimum values, or to 1.5 times either the 75 % or 25 % quartile if there are data beyond this range. Outliers are identified with red crosses.

background signal, the main source of uncertainty in calculating tropospheric O_3 partial pressures, which shows a significant trend over this period (Fig. 4e). Prior to 2003 the background current was high (0.10–0.16 µA) and highly variable. The reduced background current after 2003 is a consequence of a change in prelaunch procedure. Chemicals are renewed more often and the signal is measured outdoors just prior to launch instead of being measured indoors (A. Piters, personal communication, 2012). With typical values of < 0.06 µA, the background current at De Bilt now agrees well with the signals measured at Payerne and Uccle (0.03–0.04 µA). The drop of ΔO_3 to below 0 % from 1998–2002 can be explained by two factors, the large background current values and the change in background current treatment. Since November 1998 a constant background current has been used to process the data instead of a background current that declines with altitude. In the case of having large background current values, when a constant value is subtracted from the measured cell current much lower O_3 partial pressures are obtained than when an altitude-dependent background current is subtracted. This feature is more pronounced in the upper troposphere than in the lower stratosphere since smaller O_3 partial pressures are measured. The hypothesis that the background current values and the changing background current treatment are responsible for the low O_3 partial pressures measured by sondes is further supported by the fact that the differences between sonde and MOZAIC are more stable (i.e. ΔO_3 do not drop to below 0 %) before 2003 if the data are processed with an altitude-dependent background current, especially in the LS (Fig. 4c). After 2003, when the mean background values drop below 0.06 µA, the spread in the results resulting from different correction schemes (altitude dependent or constant) is significantly smaller. During this period, the agreement with MOZAIC in the LS is better than 5 % when a constant background is subtracted, and better than 10 % when an altitude-declining background current is subtracted. In the UT, the agreement is better than 10 % with a constant background current and better than 15 % with an altitude-declining background current.

3.1.3 Legionowo

At Legionowo, Poland, SP sondes (1.0 % KI) were also flown for the entire period considered here. In terms of data treatment, a background current declining with altitude is applied. The number of ozonesondes launched is similar to that at De Bilt, and 89 % of the launched sondes can be matched with MOZAIC. Mean differences between the Legionowo sondes and MOZAIC are substantial in the troposphere (10–15 %), but smaller in the stratosphere (< 10 %) (see Fig. 5a and b). The differences remain relatively constant in time, similar to the background current values (Fig. 6). The only exception is 1995, when the sondes exceed MOZAIC in the troposphere by up to 20 % (Fig. 5e), as observed at Payerne,

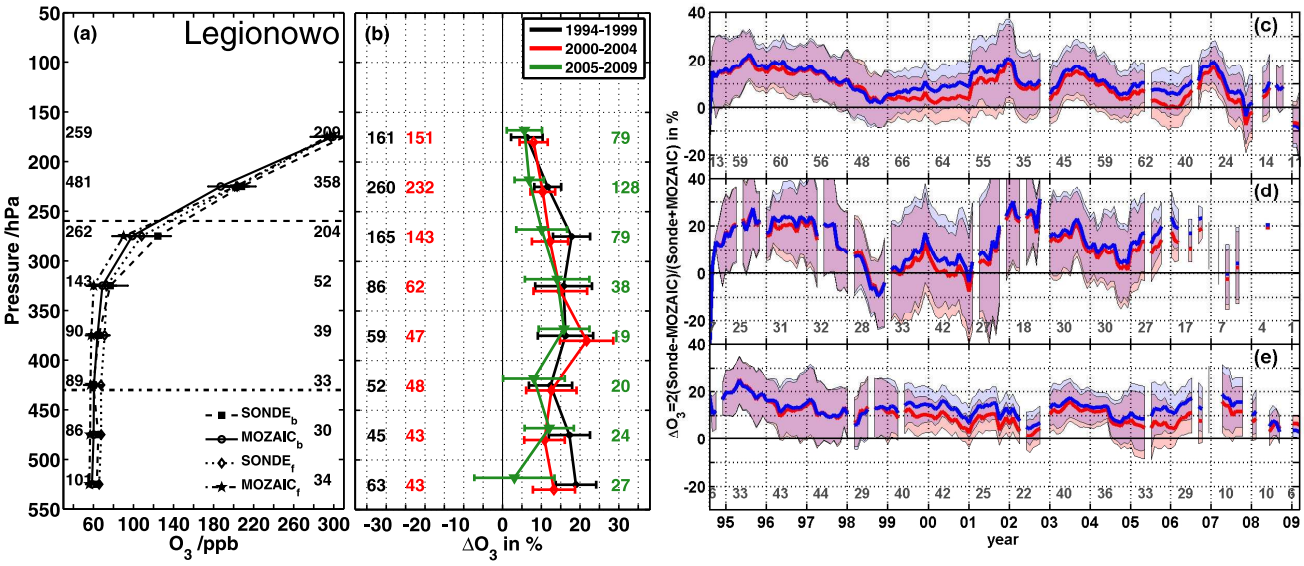

Fig. 5. Comparison of ozonesondes with MOZAIC measurements. As for Fig. 3, but for Legionowo.

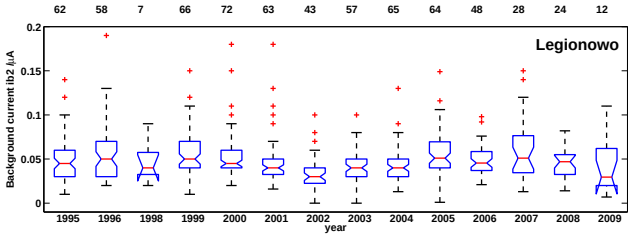

Fig. 6. Annual statistical distribution of the background current ib_2 used to process Legionowo ozonesonde data. Note that ib_2 is not reported to the archives for every launch and data prior to 26 January 1995 were not reported to the WOUDC archive.

Uccle, MOHp, and De Bilt. In the LS, differences between sonde and MOZAIC are large from 1994–1997 (15–20 %), then decrease to below 10 % from 1998–2001, and remain around 10 % thereafter. The LS time series of the ΔO_3 at both Legionowo and De Bilt show similar patterns if both data are processed assuming an altitude-declining background signal.

The five European stations analysed above reveal pronounced similarities in the sonde–MOZAIC differences. All show large discrepancies of 20–25 % in the mid-nineties (1994–1997), followed by smaller differences of 5–10 % in the subsequent years. MOZAIC's UV photometers are regularly checked for significant variations and recalibrated each year but it is remarkable that both sonde types (BM and ECC) reveal temporal similarities in ΔO_3. It is not straightforward to understand this behaviour, but it appears premature to attribute these discrepancies only to errors derived from the ozonesonde observations. Although it cannot be ruled out completely that the Lagrangian technique may have systematic problems, a slight error in the long-term stability of MOZAIC in the mid-nineties needs to considered. This has also been discussed by Logan et al. (2012), who found an increase in the MOZAIC bias from 1994–2009 over Frankfurt/Munich compared with the alpine surface site Zugspitze (Germany).

As already mentioned, the number of matches and the corresponding temporal distribution affect the application of the match approach. The ozonesonde stations presented above are clearly favoured by their location, the large number of ascents and matches (cf. Table 2), and the temporal distribution of their matches, which are mostly found within the first 50 h (cf. Fig. 2). These conditions, however, do not apply to the other stations presented below. Consequently, less precise deductions about ΔO_3 and the influence of instrument variations on the sonde performance can be made.

3.1.4 Lindenberg/Madrid/OHP

The comparison between ozonesonde data from Lindenberg, Madrid, and the Observatory Haute Provence (OHP) with MOZAIC observations is presented in Fig. 7. The Lindenberg and Madrid sites followed the ECC flight instructions of Komhyr (1986) and Komhyr et al. (1995) for SP sondes (1.0 % KI full buffered cathode solution, processed assuming an altitude-declining background current signal). At OHP they changed from using the SP sondes (flown with 1.0 % KI) to using ES sondes (1.0 % KI) in March 1997, and the data are post-processed assuming a background current that is constant with altitude. Between 400–500 ozonesondes are matched at these three sites, substantially less compared to De Bilt and Legionowo (700–800), or MOHp, Uccle, Payerne (Table 1). About 80 % of the sondes flown from Madrid and OHP can be matched with MOZAIC, while

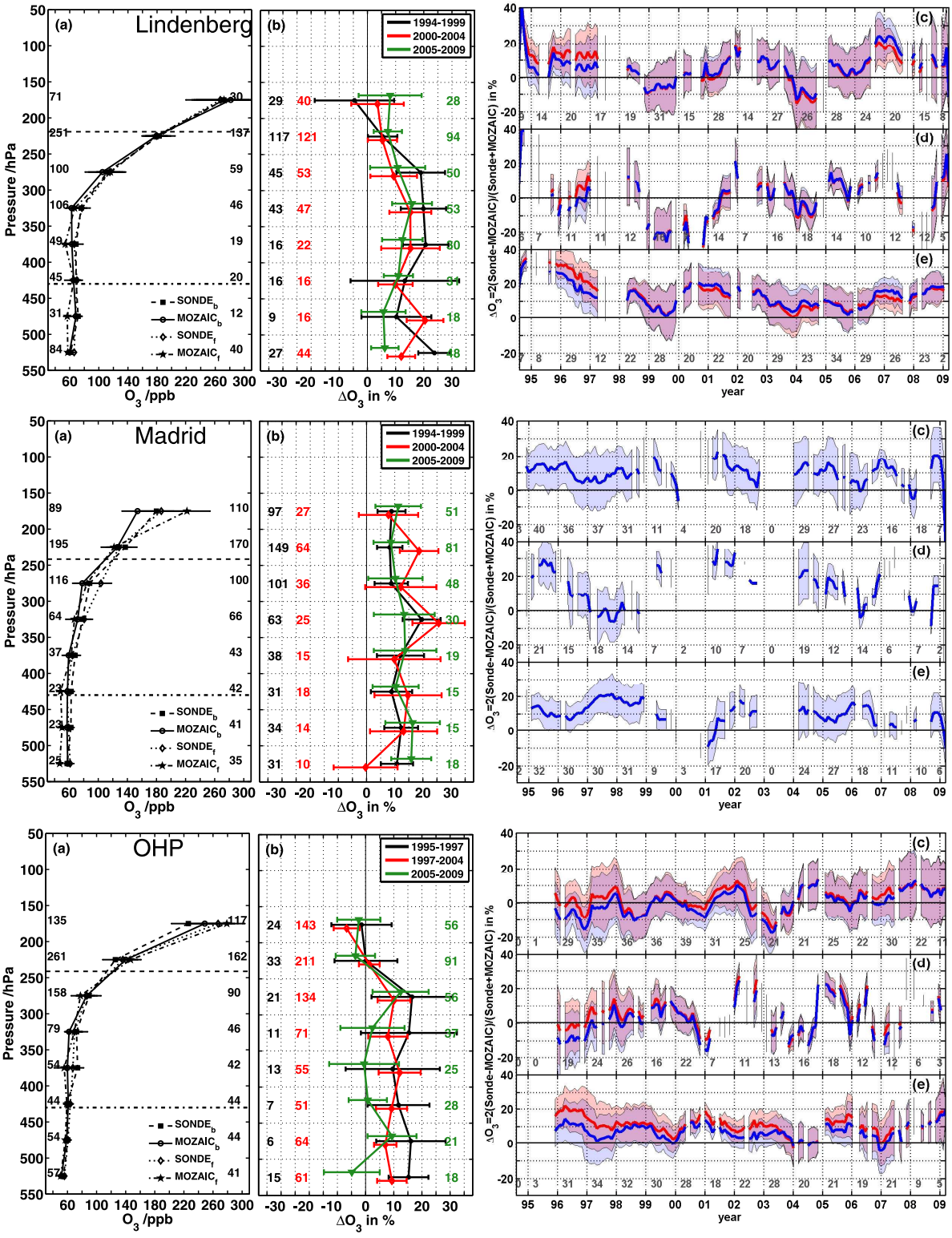

Fig. 7. Comparison of ozonesondes with MOZAIC measurements. As for Fig. 3, but for Lindenberg, Madrid, and the Observatory Haute Provence (OHP). ES sondes have been flown at OHP since March 1997, where SP sondes were used before. Both are operated with a 1.0 % electrolyte.

Table 2. Total number of matches, matched trajectories and ascents used for the different launch sites included in this study.

Station	Total number of		
	matches	matched trajectories	matched sonde ascents
Alert	2703	1051	241
Boulder	5986	1816	459
Churchill	5271	1994	341
De Bilt	41 520	13 277	706
Edmonton	9301	3269	489
Eureka	2935	1077	281
Goose Bay	10 942	3979	489
Huntsville	5250	1543	299
Irene	743	219	32
Izaña	7450	2261	533
Legionowo	22 384	8118	783
Lerwick	16 048	5958	592
Lindenberg	5718	2080	500
Madrid	10 204	3381	413
MOHp	17 576	5885	1335
Naha	2589	693	164
Nairobi	1783	397	103
Natal	739	142	60
OHP	11 060	3702	471
Paramaribo	792	204	82
Payerne	70 336	21 439	1899
Resolute	1599	627	165
Sapporo	8198	2890	415
Scoresbysund	13 004	4699	572
Sodankylä	9559	3530	629
Tsukuba	3909	1468	374
Uccle	96 746	30 007	1882
Wallops	10 854	3631	594
total	395 203	129 340	14 859

at Lindenberg less than 70 % are matched, partly because fewer trajectories are initialised. Similar to MOHp, the other DWD (German Weather Service) station, data from Lindenberg are reported on fewer pressure levels than at Madrid, OHP, and all other sites. The tropopause height at Lindenberg calculated using the soundings available for comparison is 20–30 hPa higher than several European stations, including Madrid, OHP, and Legionowo (Fig. 7a).

Between 300–400 hPa there are some sonde–MOZAIC differences (up to 10 %) between the backward- and forward-only trajectories, with the backward-only trajectories yielding larger sonde biases (Fig. 7a). As mentioned previously, and as described in detail by Staufer et al. (2013), this may result from chemical processing along the trajectories, different temporal match distributions between the unidirectional trajectories, and inaccurate quantification of the meteorological conditions in the UT.

The lower stratospheric sonde–MOZAIC differences at Lindenberg, Madrid, and OHP range from −5 % (OHP) to 5–10 % (Lindenberg, Madrid), while in the troposphere they are somewhat larger (Fig. 7b). Large discrepancies between sondes and MOZAIC are found in the stratosphere at both Lindenberg and Madrid from 1994–1996 (up to 15 %) (Fig. 7c). In the troposphere the discrepancies increase in the 1990s at Madrid, while at Lindenberg the sonde data are 15–30 % larger than MOZAIC from 1994–1998.

Our analysis indicates that there is no obvious break in the ΔO_3 time series over OHP resulting from the switch of ECC sonde manufacturer in March 1997. There is, however, a decrease in tropospheric bias after 1997, although this cannot be attributed to the change in ECC sonde type (and retaining a 1.0 % KI) since several other stations show similar deviations during this period as well. The times series in the LS is too noisy to draw any firm conclusions.

3.1.5 Churchill/Edmonton/Goose Bay

The sonde–MOZAIC comparison at the Canadian mid-latitude stations Churchill, Edmonton, and Goose Bay is presented in Fig. 8. Because of the MOZAIC flight distribution, most matches are obtained from forward trajectories (Fig. 8a), with most trajectories originating from the lower stratosphere. In total, 350–500 ozonesondes can be matched with MOZAIC at these stations, the equivalent of only one third the sample size of the European BM stations.

In the lower stratosphere, the ΔO_3 time series are qualitatively similar, especially for Edmonton and Goose Bay. The sondes overestimate MOZAIC by up to 5 % at Goose Bay, and by up to 15–35 % at Edmonton and Churchill from 1994–1996, but then underestimate O_3 compared to MOZAIC from 1997–1999 (Fig. 8c). Thereafter, the sonde–MOZAIC bias becomes positive again, ranging between 5–15 %, depending on the station. The results suggest no statistically significant differences in the mean lower stratospheric deviations (at the 90 % confidence level; see Fig. 8b). Although there are only a few tropospheric matches, the discrepancies at Goose Bay from 1995–1996 (ranging between 15–20 %) are similar to those observed at European stations (Fig. 8e).

3.1.6 Boulder/Huntsville/Wallops Island

Matches for the United States stations are obtained mostly from forward trajectories originating in the upper troposphere, particularly at Huntsville and Wallops Island (see Fig. 9a). Note that as for the Canadian stations, the number of matched ozone soundings is much lower than at MOHp, Payerne, or Uccle. This is likely in part due to the low measurement frequency and the position of these stations.

The electrolytic solutions and techniques used by NOAA are unique. The sonde solution chemistry is therefore quite different to all other stations that use one of the standard recipes. At Boulder, the sensing solution strength changed

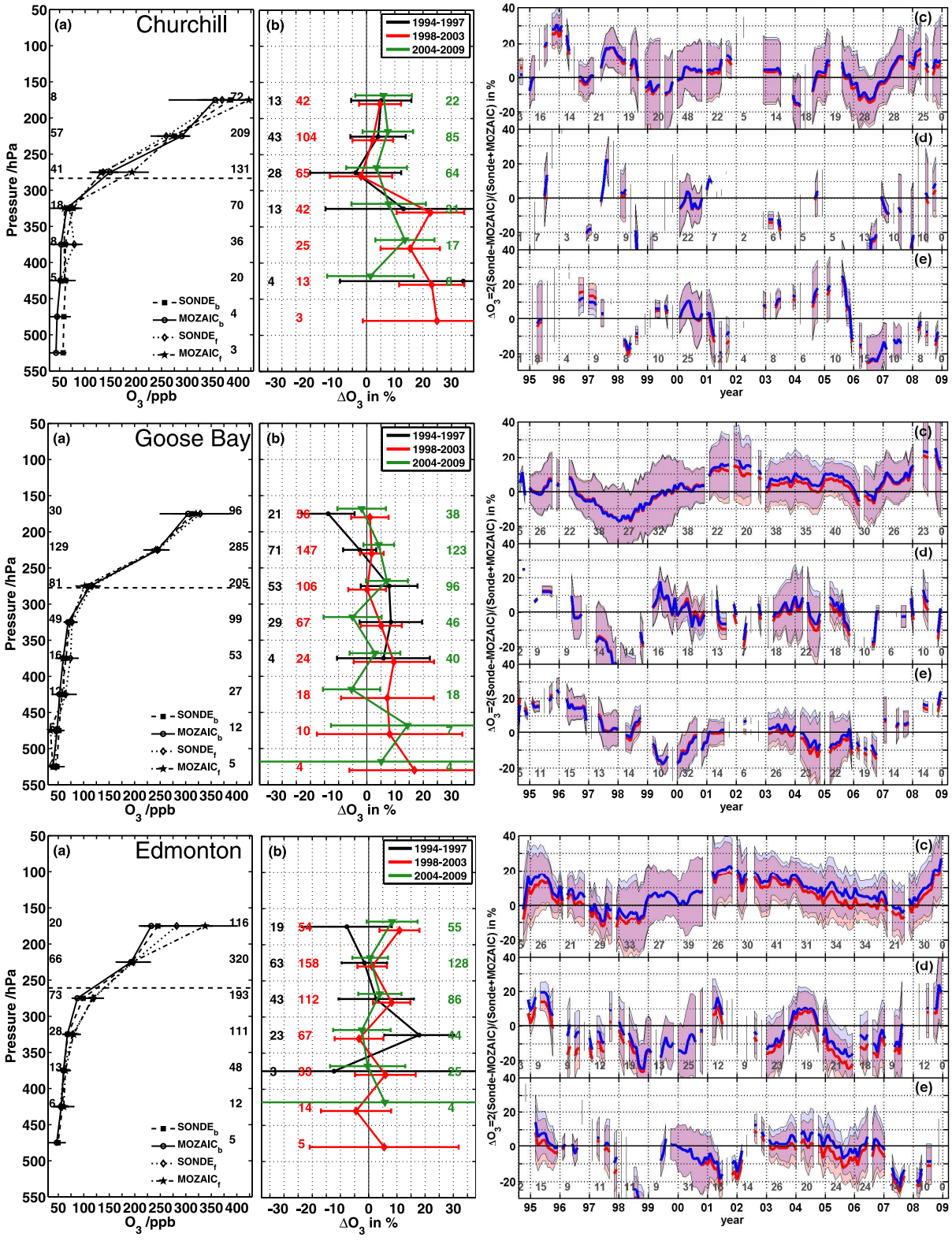

Fig. 8. Comparisons of ozonesondes with MOZAIC measurements. As for Fig. 3, but for Churchill, Edmonton, and Goose Bay. At all three launch sites both SP and ES-Z sondes were flown before 2004. After 2004 mainly ES sondes were launched but the 1.0 % solution strength was retained.

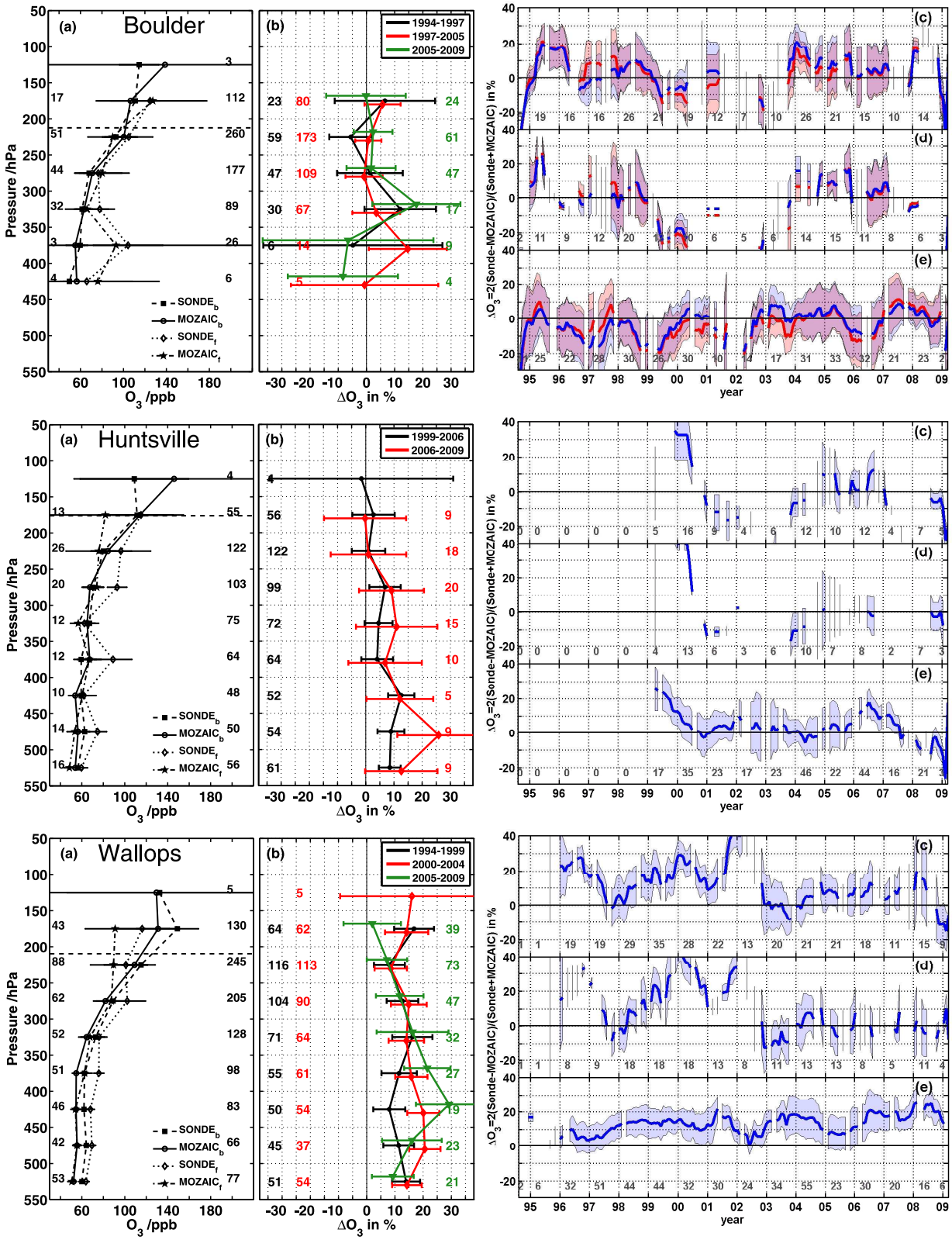

Fig. 9. Comparison of ozonesondes with MOZAIC measurements. As for Fig. 3, but for Boulder, Huntsville, and Wallops Island. For each period in **(b)** a different sensing solution was used at Boulder and Huntsville.

twice, in August 1997 and November 2005 (Table 1). The Boulder sondes exceed MOZAIC observations by < 5 % at pressure < 300 hPa for most of the time periods considered, but show a positive offset of up to 20 % in the lower stratosphere from 1995–1996 (Fig. 9c). This provides further evidence that MOZAIC perhaps underestimates O_3 compared to the sondes from 1994–1997.

Ozone soundings began at Huntsville in 1999. From all observations considered here, only very few backward trajectories are matched so the O_3 time series is almost entirely determined from forward trajectories. Sondes exceed MOZAIC by up to 10 %, depending on altitude (Fig. 9b), with slightly higher values being observed after the change from a 2.0 % KI unbuffered solution to a 1.0 % KI 1/10th buffered solution in March 2006.

Above 350 hPa, ozonesondes flown from Wallops Island show similar results from 1994–2009 with sonde measurements exceeding MOZAIC by 5–15 % (Fig. 9b). These values agree well with results from Schnadt Poberaj et al. (2009), who also show a positive sonde bias of 5–20 % compared to MOZAIC from 1994–2001. Below 350 hPa the sondes tend to measure more O_3 from 2005–2009 than in previous periods, in particular between 350–450 hPa. Such a trend is not visible at other sites.

3.1.7 Sapporo/Tsukuba

The 16 yr mean O_3 concentrations obtained from sondes flown at Tsukuba and Sapporo, and from MOZAIC agree within 5 % in the troposphere (Fig. 10a). However, there are differences in the stratospheric performance (typically at pressures < 250 hPa) from forward-only and backward-only trajectories at Tsukuba. The agreement of sondes with MOZAIC also tends to evolve differently at Sapporo and Tsukuba (Fig. 10b). Although both stations used the same sonde type (KC-79 until summer 1997, KC-96 from summer 1997 to December 2009), from 2005–2009 ΔO_3 is positive (10–20 %) at pressures < 350 hPa at Tsukuba but negative (−10 %) at Sapporo.

As already mentioned, the time lag between sonde–MOZAIC matches can be an important factor for the comparison. The temporal distribution of the individual matches is provided in Fig. 2; however, this distribution might be biased since it does not account for the averaging of the matches along each trajectory – some trajectories contain more individual matches than others – nor for the weighting of matches along the trajectories. It does, though, provide an idea of the mean time lag between MOZAIC measurements and the soundings. In contrast to most of the European stations, the majority of stratospheric matches at non-European stations are not found within the first 50 h of air parcel travel time. Rather, many matches occur at the end of the trajectories where they have already travelled more than 100 h. In our companion paper (Staufer et al., 2013), a 2 % uncertainty was found when testing this matching technique using MOZAIC–MOZAIC self-matches. This comparison found that most matches (50 %) occurred within the first two days (48 h) of trajectories. Thus, in the case of the Japanese stations with almost no matches in the first two days of the trajectories, this uncertainty is likely to be higher than 2 % due to accumulated trajectory errors and may explain the discrepancies observed. The results for Sapporo and Tsukuba are therefore less reliable.

3.2 High latitude stations

3.2.1 Lerwick/Scoresbysund/Sodankylä

Results for the high latitude stations included in this study are presented in Fig. 11. At these sites the tropopause is located at pressures < 300 hPa except for Lerwick, where it is located at pressures < 250 hPa. A larger sample is obtained for the LS compared to the UT because the height of the tropopause is lower at high latitudes than at mid-latitudes, while MOZAIC's cruise altitude remains constant (8–12 km) independent of latitude. For all three high latitude stations the sondes exceed MOZAIC by 5–10 % in the stratosphere and by 10–15 % in the troposphere. For Lerwick and Sodankylä the differences between sondes and MOZAIC obtained from backward-only trajectories are systematically larger than for forward-only trajectories (5 %) (Fig. 11a).

Scoresbysund is among those stations that have used ES ozonesondes without the right standard buffered solution. Extensive laboratory and field work showed that differences of 5–10 % can be observed if the wrong solution is used (e.g. Kivi et al., 2007; Smit et al., 2007; Deshler et al., 2008). At Scoresbysund, after the change from SP to ES sondes in 2001, measured O_3 concentrations are systematically 5 % higher at all altitudes (Fig. 11b). This is in accordance with the JOSIE (Smit et al., 2007) and BESOS experiments (Deshler et al., 2008), which showed that ES sondes had systematic high bias of around 5 % compared to SP sondes when both were operated with a 1.0 % KI cathode sensing solution.

At Sodankylä both SP and ES sondes were flown with a 1.0 % KI cathode electrolyte before February 2006. The majority of the sondes during this period were SP sondes, although for short periods, ranging from several weeks to 3 months, ES sondes were used. From February 2006 onwards only ES sondes with a 0.5 % KI solution have been flown. During both periods the sondes were operated following recommendations of the scientific community and manufacturers, and therefore only small differences of a few percent are to be expected (e.g. Smit et al., 2007; Deshler et al., 2008). As shown in Fig. 11b, no statistically significant change in the sonde performance is observed. Note that since February 2006 data are processed with a constant background current instead of with an altitude-dependent background current. In contrast to De Bilt, the trend in background current values measured as part of preflight

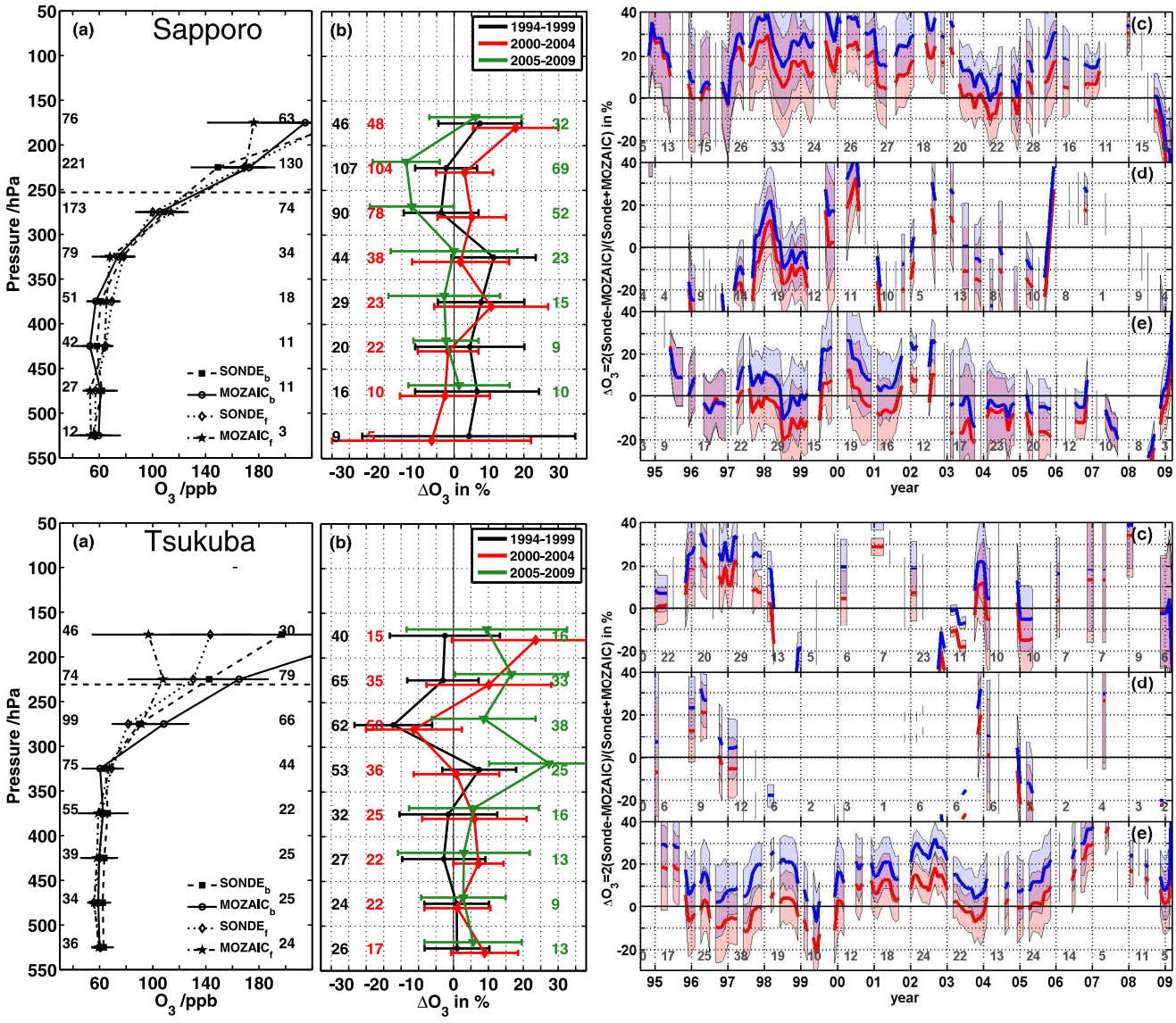

Fig. 10. Comparisons of ozonesondes and MOZAIC measurements. As for Fig. 3, but for Sapporo and Tsukuba.

preparations is small, and therefore this change is not likely to influence the sonde performance at Sodankylä.

At Lerwick sondes exceed MOZAIC, in particular in the upper troposphere in the early years (> 20 % from 1994–1996; see Fig. 11c). ΔO_3 in the LS is < 10 % for most of 1994–2009, similar to most ECC stations. According to the information given at the WOUDC, Lerwick frequently changed between SP and ES sondes, both flown with a 1.0 % KI electrolyte. We contacted the PIs for more information on the exact switch dates of the ECC sensors so as to analyse the data set more thoroughly. However, satisfying answers could not be provided by them. We therefore cannot assess whether and how the frequent changes from SP to ES sondes and vice versa influences the agreement with MOZAIC data.

3.2.2 Alert/Eureka/Resolute

Only one third of all ozone soundings from these sites are available for comparison with MOZAIC (see Table 1). Tropospheric data from these stations are particularly scarce. Because of the very small sample size, only the 16 yr 1994–2009 average is provided (see Fig. 12). Most sonde–MOZAIC matches are found at altitudes between 200–300 hPa and indicate that the sondes exceed MOZAIC measurements by 5–10 %.

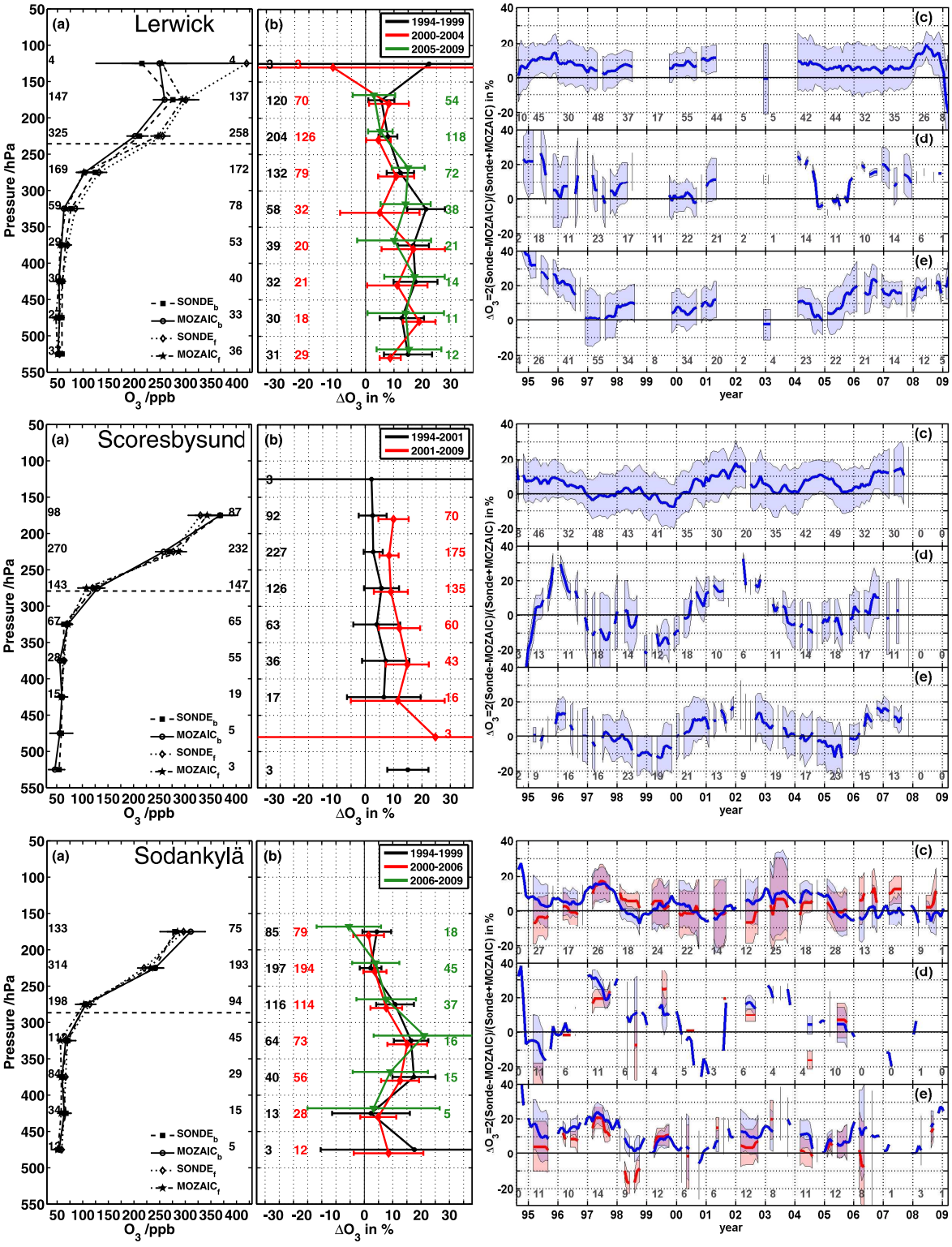

Fig. 11. Comparisons of ozonesondes and MOZAIC measurements. As for Fig. 3, but for Lerwick, Scoresbysund, and Sodankylä. Scoresbysund changed from SP to ES sondes in July 2001. The 1.0 % KI was retained. Sodankylä mainly used SP sondes (1.0 % KI) before switching to ES sondes (0.5 % KI) in January 2006.

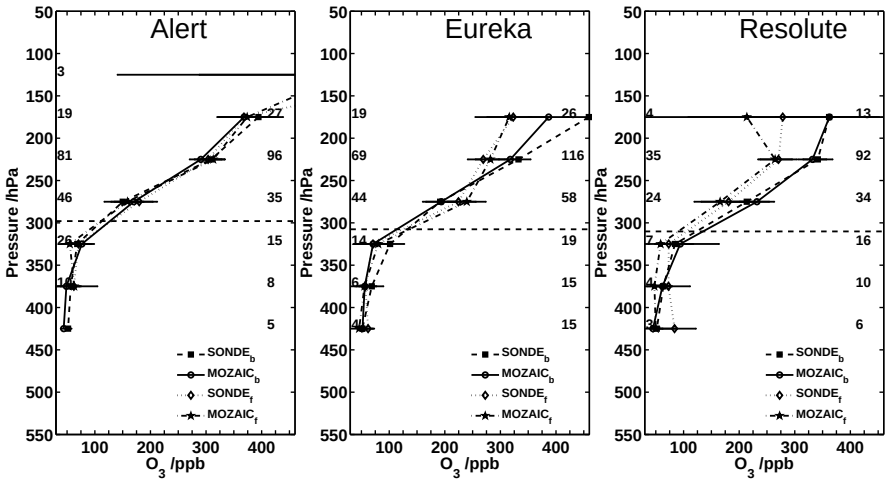

Fig. 12. 16 yr average (1994–2009) O_3 profiles from sondes flown at Alert, Eureka, and Resolute compared to O_3 profiles from MOZAIC. Data are grouped into 50 hPa layers. Numbers on the left and right indicate the number of soundings using 6 day backward-only (subscript b) and forward-only trajectories (subscript f), respectively.

3.3 Tropical stations

3.3.1 Izaña

Results for the troposphere using only forward trajectories reveal a sonde–MOZAIC bias similar to that observed at European sites such as Legionowo or Madrid. In contrast, results using only backward trajectories reveal a large offset between sondes and MOZAIC (> 20 %, see Fig. 13a). The geographical distribution of matches (Fig. 14) shows two major peaks in the backward direction, one over the Canary Islands and one over the east coast of the United States, while most matches in the forward direction are found over the Mediterranean Sea, with no pronounced peaks. The spatial distribution is also reflected in the temporal distribution of matches, since no pronounced peaks are observed (Fig. 2). Most matches with MOZAIC observations therefore occur after 3 days travel along the trajectory paths, where the trajectories are expected to be less accurate. It may also be possible that O_3 production takes place over the course of the 12 day trajectories given that the photochemical lifetime of tropospheric O_3 is expected to be shorter in the subtropics and tropics than in the mid- and high latitudes (e.g. Logan et al., 1981).

Differences between the two data sets are less than 5 % in the stratosphere, and no trend in performance is found (Fig. 13b). It appears that Fig. 13c does not support this statement, but it can be explained by the different calculation of ΔO_3. The calculation of ΔO_3 in Fig. 13b is based on monthly mean differences. Some months contain more ascents than others. Ascents in months with few matches are therefore weighted stronger than ascents in months with many matches, which is not the case for the calculation used in Fig. 13b, where all matched sonde ascents contribute with the same weight to ΔO_3. No statistically significant changes in bias are found in the troposphere either, but again note that the the sonde data have a greater positive bias from 1994–1995 compared to the subsequent four to five years (Fig. 13e).

3.3.2 Nairobi/Irene/Naha/Paramaribo/Natal

As a result of the distribution of MOZAIC flights, only very few matches with ozonesondes in the tropics and Southern Hemisphere were found (less than one third of all sondes flown). Furthermore, comparison was only possible in the troposphere because the aircraft cruise altitude is usually below the height of the tropopause. Because of the very small sample size, only the 16 yr (1994–2009) average results are provided (see Fig. 15). In addition, the rather poor temporal distribution of matches adds to the uncertainty (Fig. 2). Most stations agree with MOZAIC to within 10 % at pressures > 400 hPa and within 20 % (10–20 ppbv) at pressures < 400 hPa. The bias found at Paramaribo is considerably larger than for the other tropical stations, ranging between 30–40 % higher than MOZAIC at pressures > 300 hPa. Comparison with TOMS (Total Ozone Mapping Spectrometer) overpass observations and co-located Brewer measurements further indicate different error characteristics for Paramaribo (Thompson et al., 2012). While Paramaribo data showed a positive bias of around 10 %, most other tropical stations showed a smaller bias. However, Paramaribo data were reprocessed in 2012 with a constant background current and a pump flow of correction of Komhyr (1986) was used instead of Komhyr et al. (1995) (M. Allaart, personal communication, 2013). Thompson et al. (2012) showed that the reprocessing could nearly eliminate total ozone offsets with satellites and spectrometers. The data used here (downloaded

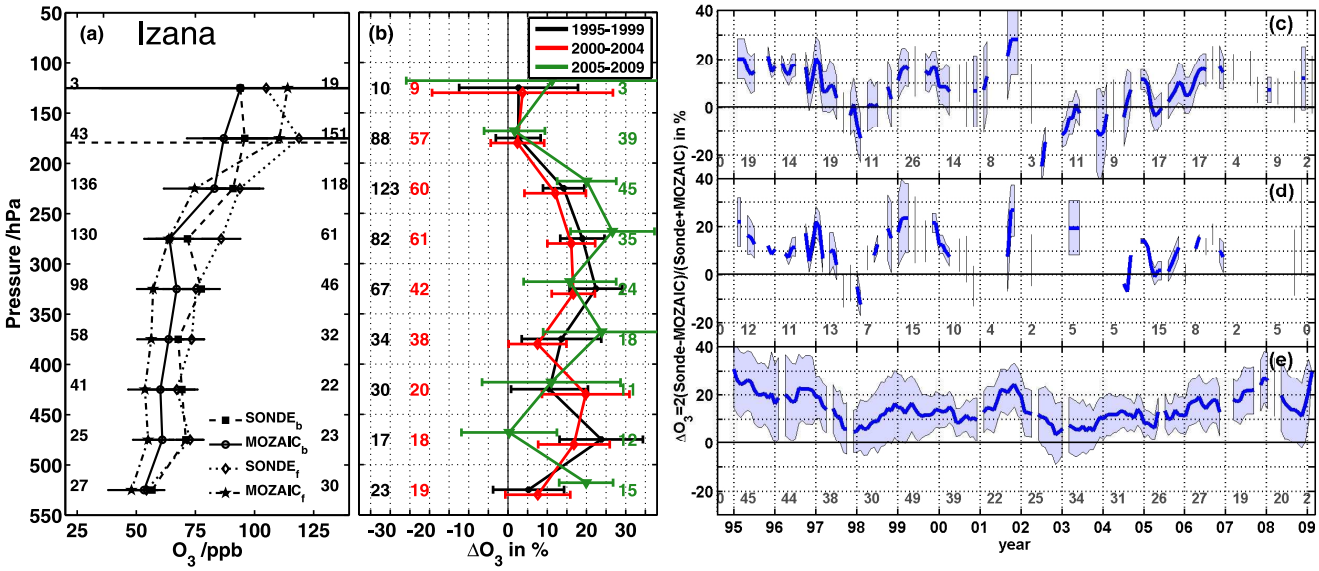

Fig. 13. Comparison of ozonesonde and MOZAIC ozone measurements. As for Fig. 3, but for Izaña. SP sondes operated with 1.0 % electrolyte are used for the entire period.

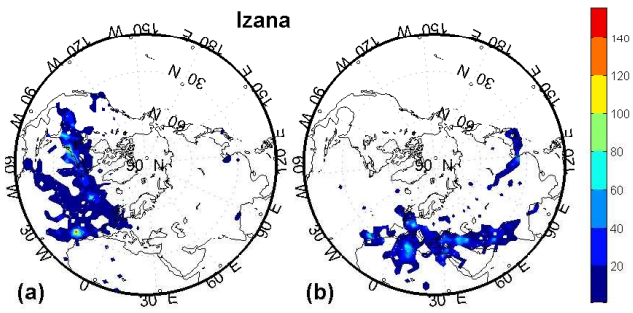

Fig. 14. Spatial distribution of matches between MOZAIC aircraft observations and **(a)** backward trajectories or **(b)** forward trajectories initialised at Izaña at altitudes between 5–15 km. The colour bar shows the total number of matches, summed up over a 3° × 3° grid.

in April 2010) evidently reflect high Paramaribo ozone archives.

4 Summary and conclusions

Because of the annual inspection and calibration of the highly accurate (±2 % uncertainty) UV-photometers, the MOZAIC data are often considered as the standard data set to validate other ozone data sets (e.g. Thouret et al., 1998; Schnadt Poberaj et al., 2009) or chemistry-transport models (e.g. Law et al., 2000; Brunner et al., 2003; Teyessèdre et al., 2007). However, recent analyses cast some doubt on the long-term stability of the MOZAIC data; particularly the differences between MOZAIC and NOXAR in 1995–1997 (Staufer et al., 2013) and the increase in bias between Frankfurt/Munich MOZAIC and the alpine surface site Zugspitze from 1994–2009 (Logan et al., 2012) raised some concern. One of the main purposes of this paper was to investigate the long-term stability and consistency of both MOZAIC and ozonesonde data. To do so, 16 yr (1994–2009) of O_3 observations from MOZAIC aircraft were compared with measurements from balloon-borne ozonesondes using 6 day, three-dimensional trajectories. Match criteria of 75 km maximum horizontal distance and 0.6 K maximum potential temperature difference (≈ ±20 m) were chosen to ensure that measurements from both platforms sampled the same air mass. This method relies on 14 859 balloon ascents that are matched with observations from MOZAIC flights, yielding a total of 129 340 independent match trajectories. O_3 measurements from soundings and airliners are averaged in 50 hPa segments to examine tropical to high latitude data.

The present analysis confirms that, at least during the MOZAIC period, ozonesondes provide a reliable tool for investigating atmospheric O_3 climatologies, even in proximity of the tropopause, where O_3 partial pressures are low and measurement uncertainty high. The differences between ozonesondes and MOZAIC are typically smaller in the lower stratosphere (5–10 %) than in the upper troposphere (10–15 %), where the uncertainty of ozonesondes is higher. Stratospheric O_3 climatologies from ozonesondes also agree within 5–10 % with satellites such as MLS (Microwave Limb Sounder, e.g. Jiang et al., 2007) and OMI (Ozone Monitoring Instrument, e.g. Thompson et al., 2012).

At mid- and high latitude stations, ozonesondes typically differ from MOZAIC by 5–10 % after 1998, in very good agreement with previous lab and field studies (for example, the JOSIE and BESOS experiments, respectively, Smit

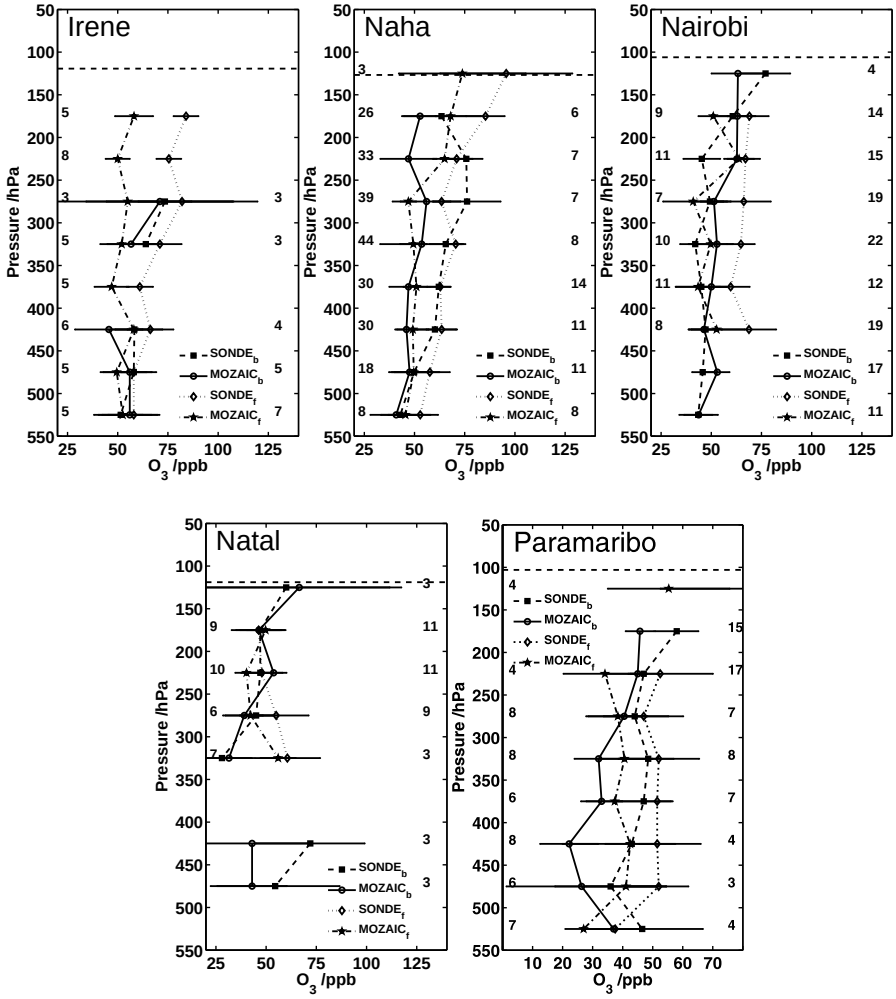

Fig. 15. Comparison of MOZAIC ozone measurements at tropical and southern latitude ozonesonde stations. As for Fig. 12, but for Irene, Naha, Nairobi, Natal, and Paramaribo.

et al., 2007; Deshler et al., 2008, where ozonesondes were also compared to the UV-photometer technique). However, before 1998 discrepancies of up 25 % are observed and one therefore could argue that the ozonesondes should only be trusted after 1998 since the UV photometry technology used by the MOZAIC program is expected to be more precise, particularly at the low ozone concentrations typical of the UTLS. The MOZAIC instruments are also inspected annually and regularly calibrated, which is not the case for single-use balloon-borne ozonesondes. The fact, however, that 10 of the 20 time series shown in Fig. 16 indicate large positive differences compared to MOZAIC in the mid-1990s, followed by a systematic tendency to smaller differences in subsequent years, casts some doubt on the explanation that the differences are due solely to errors stemming from the ozonesondes. It is remarkable that various sonde types reveal similar behaviour, namely BM (Brewer-Mast) and ECC (electrochemical cells manufactured by either SP or ES). In view of the fact that three different manufacturers were involved in building these instruments, it is not straightforward to understand this behaviour. Likewise, however, MOZAIC operated five aircraft instruments simultaneously, and it is also not clear how these instruments could explain the observed differences, even though they are identically constructed, maintained, and calibrated. A comparison between Payerne BM sondes and O_3 measurements made during the NOXAR B747 project from 1995–1996 showed a smaller offset of around 15 % (scaled) compared to MOZAIC, which may indicate a small drift in the MOZAIC calibration (see Staufer et al., 2013).

The method developed in the companion paper (Staufer et al., 2013) and applied here provides the most reliable results for launch sites with a large number of matches within the first 2 or 3 days. Mainland Europe-based launch sites such as De Bilt, MOHP, Payerne, Uccle, and Legionowo are clearly favoured over most other stations, particularly over

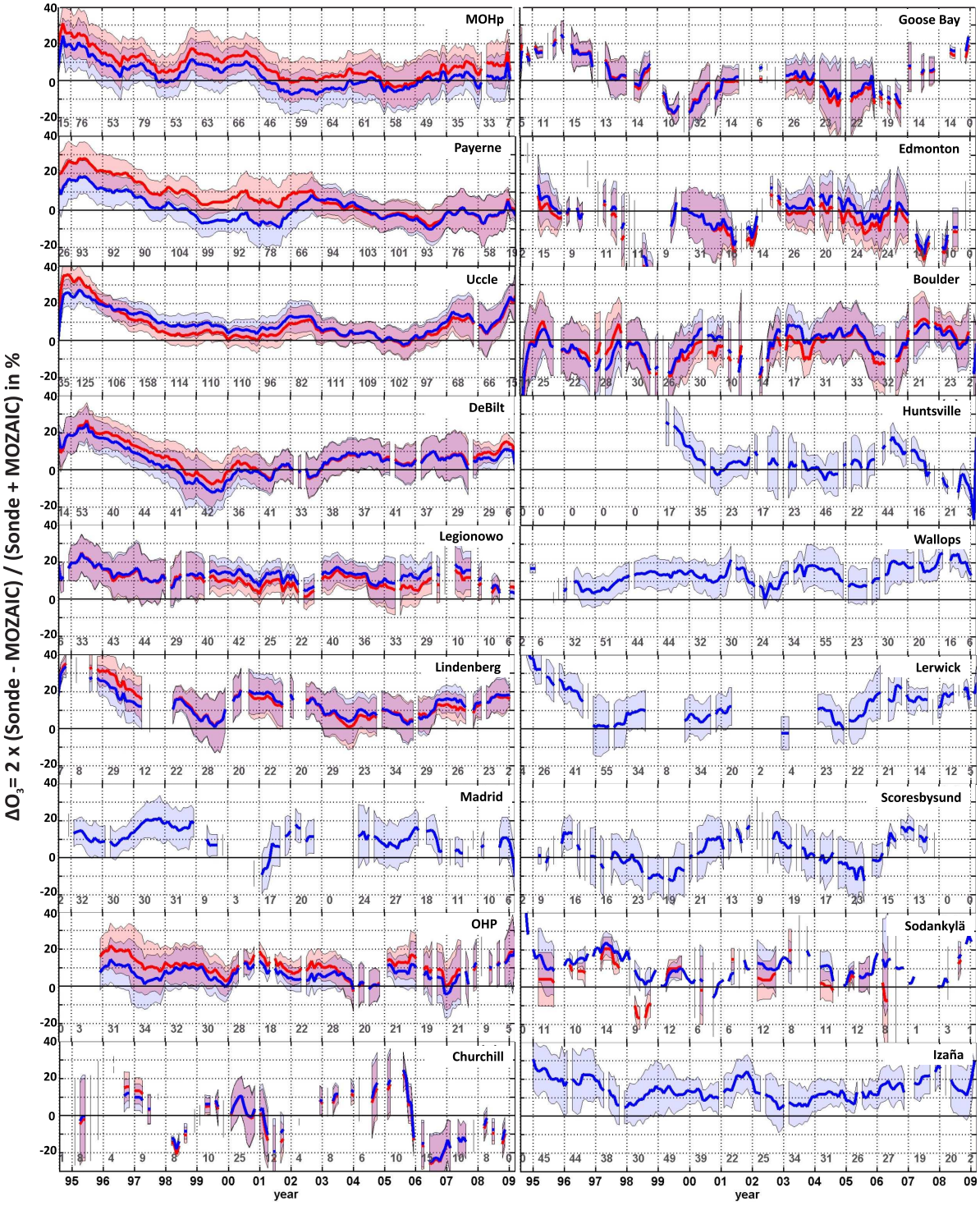

Fig. 16. Time series of the relative differences between ozonesondes and MOZAIC measurements (ΔO_3 in %) in the upper troposphere. These time series comprise 20 of the 28 launch sites considered in this work (the 8 remaining have too few matches with MOZAIC to be included here). Note the large differences at nearly all sites from 1994 to 1998, and the systematic tendency for smaller differences at 10 of these stations thereafter, three using BM sondes (MOHp, Payerne, Uccle) and seven using SP or ES ECC sondes (De Bilt (SP), Legionowo (SP), Lindenberg (SP), Goose Bay (SP, ES), Edmonton (SP, ES), Lerwick (SP, ES), and Izaña(SP)). Bold lines: ΔO_3 time series with CF (red) and without CF (blue). Numbers at the bottom indicate the number of sondes used for calculating the monthly mean differences. The shaded areas denote the standard error (68 % confidence). Overlapping areas are displayed in light purple.

those in the tropics, southern latitudes, and Japan, which all suffer from long trajectories with larger uncertainty and fewer matches. These European stations therefore could be analysed more thoroughly, while deductions drawn from most other stations certainly are less clear and conclusive.

The BM sondes flown operationally at MOHp, Payerne, and Uccle from 1994–1997 overestimate O_3 by up to 25 % in the upper troposphere compared to MOZAIC. These results agree well with previous studies (Logan et al., 2012; Schnadt Poberaj et al., 2009). Due to the more favourable conditions, measurements in the lower stratosphere show a smaller offset during this period, especially at Uccle. After 1998, the sonde–MOZAIC deviations decrease to values below 10 % in both the UT and LS. By 1998 most stations had switched from using BM to ECC sondes, with, for example, Uccle having switched in 1997 and Payerne in 2002. In comparison, MOHp continues to operate BM sondes. From 2000 onwards, the sondes flown at these three stations agree with MOZAIC to better than 5 %. Thus, the ES sondes flown at both Uccle and Payerne resemble the expected bias from previous lab and field studies (e.g. Smit et al., 2007; Deshler et al., 2008).

De Bilt and Legionowo flew all SP sondes with the recommended 1.0 % KI fully buffered cathode sensing solution. It is interesting to note that Legionowo data are not as close (ΔO_3 is typically around 10 %) to MOZAIC as Uccle and Payerne ES sondes. As shown by Smit et al. (2007) (for example, their Fig. 12), a very similar bias to the UV-photometer technique is expected. It is difficult to explain and understand this offset, and it cannot be completely ruled out that it is the result of the Lagrangian match technique. However, the results for De Bilt after 2002, which show a better agreement of typically 5 % between sonde and MOZAIC when a constant background current is subtracted from the cell current, may suggest that the constant background current correction is more appropriate. Indeed, the recent assessment of ECC SOPs calls for a constant background current (Smit et al., 2011). However, the physico–chemical description of the background current is not well understood and further research is required to better understand its origin and its appropriate measurement and treatment (see discussion in Vömel and Diaz, 2010; Smit et al., 2011). In contrast, the unusual negative bias of De Bilt sondes compared to MOZAIC's UV-photometers from 1998–2002 can be explained by large background current values and the change in treatment.

In other published studies, for example Kivi et al. (2007); Smit et al. (2007); Deshler et al. (2008), ECC sondes operated not following the manufacturer's recommendation for appropriate cathode sensing solution strength have been found to have systematic offsets compared to those operated accordingly. The stations affected by this, however, are not located in mainland Europe (except OHP) and thus not favoured by the matching technique. The influences of the different solution strengths were not generally apparent in the comparisons with MOZAIC UTLS O_3 measurements. There are too few matches to analyse this impact at the Natal site. OHP changed in 1997 from SP to ES sondes and kept a 1.0 % KI solution, but discerning this impact is made difficult by the large discrepancies between sondes and MOZAIC in the mid-nineties. The analysis of the Canadian sites does not reveal a change in agreement with MOZAIC introduced by the change of sensor type without changing the cathode electrolyte. The only launch station at which we find a systematic increase in the sonde O_3 measurements resulting from a change from SP to ES sondes (and the solution strength being retained) is Scoresbysund. The results for Scoresbysund are in agreement with conclusions from the JOSIE 2000 experiments (Smit et al., 2007), who reported a higher bias of 5 % for ES sondes compared to SP sondes when both are operated with a 1.0 % KI sensing solution. Boulder and Huntsville also changed solutions. The cathode solutions, however, used by NOAA are unique and their own data processing techniques may account for these changes. The agreement with MOZAIC appears to hardly be affected by changes in solution strength.

There is an ongoing debate on the application of a correction factor (CF) to normalise sonde profiles to a nearby column O_3 measurement (Dobson or Brewer), in particular concerning the application of a CF to the tropospheric fraction of the measurements (e.g. SPARC/IOC/GAW, 1998; Thouret et al., 1998; Stübi et al., 2008; Schnadt Poberaj et al., 2009). Since the CF largely depends on stratospheric O_3 levels, doubts have been raised with respect to its application to the tropospheric part of profiles. The application of a CF implies making assumptions about the O_3 content above burst altitude, which can introduce biases originating from the independent column measurements used. We find no systematic behaviour of sondes to the application of a CF, but rather to a dependence on time; for example, at MOHp, soundings in the lower stratosphere show better agreement with MOZAIC before 1998 if not normalised, but after 1998 the normalisation decreases the sonde–MOZAIC differences. In the troposphere, however, better agreement is obtained without normalisation over the entire time period. The only two exceptions are at Uccle and Legionowo where upper tropospheric soundings agree better with MOZAIC when the data are normalized using a CF.

Acknowledgements. J. Staufer's research was funded by a MeteoSwiss GAW-CH grant. The authors would like to thank the European Commission, Airbus, CNRS-France, and FZJ-Germany for their strong support of the MOZAIC project, as well as the airlines (Lufthansa, AirFrance, Austrian, Air Namibia, and former Sabena) involved in the project who have carried the MOZAIC instrumentation free of charge since 1994. MOZAIC is presently funded by INSU-CNRS (France), Météo-France, CNES, Université Paul Sabatier (Toulouse, France) and Research Center Jülich (FZJ, Jülich, Germany). IAGOS is additionally funded by the EU projects IAGOS-DS, IAGOS-ERI, and IGAS. MOZAIC-IAGOS

data are available via the CNES/CNRS-INSU Ether web site http://www.pole-ether.fr. We gratefully acknowledge all staff at the soundings sites for carefully preparing and performing all launches. The ozone sonde data used in this publication were obtained from the World Ozone and Ultraviolet Radiation Data Centre (WOUDC; data publicly available at http://www.woudc.org, the Network for the Detection of Atmospheric Composition Change (NDACC; data publicly available at http://www.ndsc.ncep.noaa.gov), SHADOZ (http://croc.gsfc.nasa.gov/shadoz/), a NOAA ftp server (ftp://ftp.cmdl.noaa.gov/ozwv/ozone/), and from NILU's NADIR database. We would also like to thank Bryan J. Johnson and Alberto Redondas for providing information regarding the sonde configuration used at the NOAA sites and Izaña, respectively. Thanks are also due to Ankie Piters and Marc Allaart from KNMI (Royal Netherlands Meteorological Institute) for providing the information on data processing and sonde preparation changes at De Bilt and Paramaribo. Finally, we thank the anonymous reviewers for their constructive and helpful suggestions and comments that helped to improve the manuscript.

Edited by: R. Eckman

References

Brewer, A. and Milford, J.: The Oxford Kew ozonesonde, P. Roy. Soc. Lond., 256, 470–495, doi:10.1098/rspa.1960.0120, 1960.

Brunner, D., Staehlin, J., Jeker, D., Wernli, H., and Schumann, U.: Nitrogen oxides and ozone in the tropopause region of the Northern Hemisphere: measurements from commercial aircraft in 1995/1996 and 1997, J. Geophys. Res., 106, 27673–27699, 2001.

Brunner, D., Staehelin, J., Rogers, H. L., Köhler, M. O., Pyle, J. A., Hauglustaine, D., Jourdain, L., Berntsen, T. K., Gauss, M., Isaksen, I. S. A., Meijer, E., van Velthoven, P., Pitari, G., Mancini, E., Grewe, G., and Sausen, R.: An evaluation of the performance of chemistry transport models by comparison with research aircraft observations – Part 1: Concepts and overall model performance, Atmos. Chem. Phys., 3, 1609–1631, doi:10.5194/acp-3-1609-2003, 2003.

Claude, H., Hartmannsgruber, R., and Köhler, U.: Measurement of Atmospheric Ozone Profiles Using the Brewer/Mast Sonde – Preparation, Procedure, Evaluation, WMO Global Ozone Research and Monitoring Project, World Meteorol. Organ., Geneva, Switzerland, 17, 1987.

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

De Backer, H.: Homogenisation of ozone vertical profile measurements at Uccle, Wetenschappelijke en technische publicaties van het K. M. I., no. 7, 26, available at: ftp://ftp.kmi-irm.be/dist/meteo/hugo/publ/1999/o3prof.pdf (last access: 1 March 2010), 1999.

De Backer, H., De Muer, D., and De Sadelaer, G.: Comparison of ozone profiles obtained with Brewer–Mast and Z-ECC sensors during simultaneous ascents, J. Geophys. Res., 103, 19641–19648, 1998.

Deshler, T., Mercer, J. L., Smit, H. G. J., Stübi, R., Levrat, G., Johnson, B. J., Oltmans, S. J., Rigel, K., Thompson, A. M., Witte, J., Davies, J., Schmidlin, F. J, Brothers, G., and Sasaki, T.: Atmospheric comparison of electrochemical cell ozonesondes from different manufacturers and with different cathode solution strenghts: the Ballon Experiment on standards for ozonesondes, J. Geophys. Res., 113, D04307, doi:10.1029/2007JD008975, 2008.

Hilsenrath, E. W., Attmannspacher, W., Bass, A., Evans, W., Hagemeyer, R., Barnes, R. A., Komhyr, W., Mauersberger, K., Mentall, J., Proffitt, M., Robbins, D., Taylor, S., Torres, A., and Weinstock, E.: Results from the balloon ozone intercomparison campaign (BOIC), J. Geophys. Res., 91, 13137–13152, 1986.

Jeannet, P., Stübi, R., Levrat, G., and Staehelin, J.: Ozone balloon soundings at Payerne (Switzerland): reevaluation of the time series 1967–2002 and trend analysis, J. Geophys. Res., 112, D11302, doi:10.1029/2005JD006862, 2007.

Jiang, Y. B., Froidevaux, L., Lambert, A., Livesey, N. J., Read, W. G., Waters, J. W., Bojkov, B., Leblanc, T., McDermid, I. S., Godin-Beekmann, S., Filipiak, M. J., Harwood,, R. S., Fuller, R. A., Daffer, W. H., Drouin, B. J., Cofield, R. E., Cuddy, D. T., Jarnot, R. F., Knosp, B. W., Perun, V. S., Schwartz, M. J., Snyder, W. V., Stek, P. C., Thurstans, R. P., Wagner, P. A., Allaart, M., Andersen, S. B., Bodeker, G., Calpini, B., Claude, H., Coetzee, G., Davies, J., De Backer, H., Dier, H., Fujiwara, M., Johnson, B., Kelder, H., Leme, N. P., König-Langlo, G., Kyro, E., Laneve, G., Fook, L. S., Merrill, J., Morris, G., Newchurch, M., Oltmans, S., Parrondos, M. C., Posny, F., Schmidlin, F., Skrivankova, P., Stubi, R., Tarasick, D., Thompson, A., Thouret, V., Viatte, P., Vömel, H., von Der Gathen, P., Yela, M., and Zablocki, G.: Validation of Aura Microwave Limb Sounder Ozone by ozonesonde and lidar measurements, J. Geophys. Res., 112, D24S34, doi:10.1029/2007JD008776, 2007.

Johnson, B. J., Oltmans, S. J., Vömel, H., Smit, H. G. J., Deshler, T., and Kröger, C.: Electrochemical concentration cell (ECC) ozonesonde pump efficiency measurements and tests on the sensitivity to ozone of buffered and unbuffered ECC sensor cathode solutions, J. Geophys. Res., 107, 4393, doi:10.1029/2001JD000557, 2002.

Kivi, R., Kyrö, E., Turunen, T., Harris, N. R. P., von der Gathen, P., Rex, M., Andersen, S. B., and Wohltmann, I.: Ozonsonde observations in the Arctic during 1989–2003: Ozone variability and trends in the lower stratosphere and free troposphere, J. Geophys. Res., 112, doi:10.1029/2006JD007271, 2007.

Kobayashi, J. and Toyama, Y.: On various methods of measuring the vertical distribution of atmospheric ozone (III) – carbon iodine type chemical ozonesonde, Pap. Meteorol. Geophys., 17, 113–126, 1966.

Komhyr, W. D.: Electrochemical concentration cells for gas analysis, Ann. Geophys.-Italy, 25, 203–210, 1969.

Komhyr, W. D.: Operations Handbook – Ozone Measurements to 40 km Altitude with 4A-ECC-Ozone Sondes, NOAA Techn. Memorandum ERL-ARL-149, Air Resources Laboratory, Boulder, CO, 1986.

Komhyr, W. D., Barnes, R. A., Brothers, G. B., Lathrop, J. A., and Opperman, D. P.: Electrochemical concentration cell ozonesonde

performance evaluation during STOIC 1989, J. Geophys. Res., 100, 9231–9244, 1995.

Labow, G. J., McPeters, R. D., Bhartia, P. K., and Kramarova, N.: A comparison of 40 years of SBUV measurements of column ozone with data from the Dobson/Brewer network, J. Geophys. Res. Atmos., 118, 7370–7378, doi:10.1002/jgrd.50503, 2013.

Law, K. S., Plantevin, P.-H., Thouret, V., Marenco, A., Asman, W. A. H., Lawrence, M., Crutzen, P. J., Muller, J.-F., Hauglustaine, D. A., and Kanakidou, M.: Comparison between global chemistry transport model results and Measurements of Ozone and Water Vapor by Airbus In-Service Aircraft (MOZAIC) data, J. Geophys. Res., 105, 1503–1525, 2000.

Liu, X., Chance, K., Sioris, C. E., Kurosu, T. P., and Newchurch, M. J.: Intercomparison of GOME, ozonesondes, and SAGE II measurements of ozone: demonstration of the need to homogenize available ozonesonde data sets, J. Geophys. Res., 111, D14305, doi:10.1029/2005JD006718, 2006.

Logan, J. A., Prather, M. J., Wofsy, S. C., and McElroy, M. B.: Tropospheric chemistry: a global perspective, J. Geophys. Res., 86, 7210–7254, 1981.

Logan, J. A., Staehelin, J., Megretskaia, I. A., Cammas, J.-P., Thouret, V., Claude, H., De Backer, H., Steinbacher, M., Scheel, H.-E., Stübi, R., Fröhlich, M., and Derwent, R.: Changes in ozone over Europe: analysis of ozone measurements from sondes, regular aircraft (MOZAIC) and alpine surface sites, J. Geophys. Res., 117, D09301, doi:10.1029/2011JD016952, 2012.

Marenco, A., Thouret, V., Nédélec, P., Helten, M., Kley, D., Karcher, F., Simon, P., Law, K., Pyle, J., Poschmann, G., Von Wrede, R., Hume, C., and Cook, T.: Measurement of ozone and water vapour by Airbus in-service aircraft: the MOZAIC airborne program, an overview, J. Geophys. Res., 103, 25631–25642, 1998.

Reid, S. J., Vaughan, G., Marsh, A. R. W., and Smit, H. G. J.: Accuracy of ozonesondes measurements in the troposphere, J. Atmos. Chem., 25, 215–226, 1996.

Rex, M., von der Gathen, P., Harris, N. R. P., Lucic, D., Knudsen, B. M., Braathen, G. O., Reid, S. J., De Backer, H., Claude, H., Fabian, R., Fast, H., Gil, M., Kyrö, E., Mikkelsen, I. S., Rummukainen, M., Smit, H. G., Stähelin, J., Varotsos, C., and Zaitcev, I.: In situ measurements of stratospheric ozone depletion rates in the Arctic winter 1991/1992: a Lagrangian approach, J. Geophys. Res., 103, 5843–5853, 1998.

Schnadt Poberaj, C., Staehelin, J., Brunner, D., Thouret, V., De Backer, H., and Stübi, R.: Long-term changes in UT/LS ozone between the late 1970s and the 1990s deduced from the GASP and MOZAIC aircraft programs and from ozonesondes, Atmos. Chem. Phys., 9, 5343–5369, doi:10.5194/acp-9-5343-2009, 2009.

Smit, H. G. J. and Kley, D.: JOSIE: the 1996 WMO international intercomparison of ozonesondes under quasi flight conditions in the environmental simulation chamber at Jülich, WMO Global Atmos. Watch Rep., 130, World Meteorol. Organ., Geneva, Switzerland, 1998.

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., 112, D19306, doi:10.1029/2006JD007308, 2007.

Smit, H. G. J., De Backer, H., Davies, J., Deshler, T., Fujimoto, T., Hoegger, B., Johnson, B.J., Kyrö, E., Levrat, G., Northam, T., Oltmans, S. J., Posny, F., Schmidlin, F. J., Stübi, R., Thompson, A. M., Viatte, P., Witte, J. C., and Proffitt, M.: Quality Assurance and Quality Control for Ozonesonde Measurements in GAW, WMO Global Atmos. Watch Rep., No. 201, World Meteorol. Organ., Geneva, Switzerland, 2011.

SPARC/IOC/GAW: Assessment of Trends in the Vertical Distribution of Ozone, edited by: Harris, N., Hudson, R., and Phillips, C., SPARC Report No. 1, WMO Ozone Research and Monitoring Project, Report No. 43, World Meteorol. Organ., Geneva, Switzerland, 1998.

Staufer, J., Staehelin, J., Stübi, R., Peter, T., Tummon, F., and Thouret, V.: Trajectory matching of ozonesondes and MOZAIC measurements in the UTLS – Part 1: Method description and application at Payerne, Switzerland, Atmos. Meas. Tech., 6, 3393–3406, doi:10.5194/amt-6-3393-2013, 2013.

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

Stohl, A., and Seibert, P.: Accuracy of trajectories as determined from the conservation of meteorological tracers, Q. J. R. Meteorol. Soc., 124, 1465–1484, 1998.

Stübi, R., Levrat, G., Hoegger, B., Viatte, P., Staehelin, J., and Schmidlin, F. J.: In-flight comparison of Brewer–Mast and electrochemical concentration cell ozonesondes, J. Geophys. Res., 113, D13302, doi:10.1029/2007JD009091, 2008.

Tarasick, D. W., Jin, J. J., Fioletov, V. E., Liu, G., Thompson, A. M., Oltmans, S. J., Liu, J., Sioris, C. E., Liu, X., Cooper, O. R., Dann, T., and Thouret, V.: An ozone climatology for INTEX and ARCTAS from IONS ozonesondes, J. Geophys. Res.,115, D20301, doi:10.1029/2009JD012918, 2010.

Terao, Y., and Logan, J. A.: Consistency of time series and trends of stratospheric ozone as seen by ozonesonde, SAGE II, HALOE, and SBUV(/2), J. Geophys. Res., 112, D06310, doi:10.1029/2006JD007667, 2007.

Teyssèdre, H., Michou, M., Clark, H. L., Josse, B., Karcher, F., Oliviér, D., Peuch, V.-H., Saint-Martin, D., Cariolle, D., Attié, J.-L., Nédélec, P., Ricaud, P., Thouret, V., van der A, R. J., Volz-Thomas, A., and Chéroux, F.: A new tropospheric and stratospheric Chemistry and Transport Model MOCAGE-Climate for multi-year studies: evaluation of the present-day climatology and sensitivity to surface processes, Atmos. Chem. Phys., 7, 5815–5860, doi:10.5194/acp-7-5815-2007, 2007.

Thompson, A. M., Witte, J. C., McPeters, R. D., Oltmans, S. J., Schmidlin, F. J., Logan, J. A., Fujiwara, M., Kirchhoff, V. W. J. H., Posny, F., Coetzee, G. J. R., Hoegger, B., Kawakami, S., Ogawa, T., Johnson, B. J., Vömel, H., and Labow, G.: Southern Hemisphere Additional Ozonesondes (SHADOZ) 1998–2000 tropical ozone climatology 1. Comparison with Total Ozone Mapping Spectrometer (TOMS) and ground-based measurements, J. Geophys. Res., 108, 8238, doi:10.1029/2001JD000967, 2003a.

Thompson, A. M., Witte, J. C., Oltmans, S. J., Schmidlin, F. J., Logan, J. A., Fujiwara, M., Kirchhoff, V. W. J. H., Posny, F., Coetzee, G. J. R., Hoegger, B., Kawakami, S., Ogawa, T., For-

tuin, J. P. F., and Kelder, H. M.: Southern Hemisphere Additional Ozonesondes (SHADOZ) 1998–2000 tropical ozone climatology 2. Tropospheric variability and the zonal wave-one, J. Geophys. Res., 108, 8241, doi:10.1029/2002JD002241, 2003b.

Thompson, A. M., Witte, J. C., Smit, H. G. J., Oltmans, S. J., Johnson, B. J., Kirchhoff, V. W. J. H., and Schmidlin, F. J.: Southern Hemisphere Additional Ozonesondes (SHADOZ) 1998–2004 tropical ozone climatology. 3. Instrumentation, Station Variability, Evaluation with Simulated Flight Profiles, J. Geophys. Res., 112, D03304, doi:10.1029/2005JD007042, 2007.

Thompson, A. M., Oltmans, S. J.,, Tarasick, D. W., Von der Gathen, P., Smit, H. G. J., and Witte, J. C.,: Strategic ozone sounding networks: Review of design and accomplishments, Atmos. Environ., 45, 2145–2163, doi:10.1016/j.atmosenv.2010.05.002, 2011.

Thompson, A. M., Miller, S. K., Tilmes, S., Kollonige, D. W., Witte, J. C., Oltmans, S. J., Johnson, B. J., Fujiwara, M., Schmidlin, F. J., Coetzee, G. J. R., Komala, N., Maata, M., bt Mohamad, M., Nguyo, J., Mutai, C., Ogino, S Y., Raimundo Da Silva, F., Paes Leme, N. M., Posny, F., Scheele, R., Selkirk, H. B., Shiotani, M., Stübi, R., Levrat, G., Calpini, B., Thouret, V., Tsuruta, H., Valverde Canossa, J., Vömel, H., Yonemura, S., Andrés Diaz, J., Tan Thanh, N. T., and Thuy Ha, H. T.: Southern Hemisphere Additional Ozonesondes (SHADOZ) ozone climatology (2005-2009): Tropospheric and tropical tropopause layer (TTL) profiles with comparisons to OMI-based ozone products., J. Geophys. Res., 117, D23301, doi:10.1029/2011JD016911, 2012.

Thornton, D. C. and Niazy, N.: Sources of background current in the ECC ozonesonde: implications for total ozone measurements, J. Geophys. Res., 87, 8943–8950, 1982.

Thouret, V., Marenco, A., Logan, J. A., Nédélec, P., and Grouhel, C.: Comparisons of ozone measurements from the MOZAIC airborne program and the ozone sounding network at eight locations, J. Geophys. Res., 103, 25695–25720, 1998.

Thouret, V., Cammas, J.-P., Sauvage, B., Athier, G., Zbinden, R., Nédélec, P., Simon, P., and Karcher, F.: Tropopause referenced ozone climatology and inter-annual variability (1994–2003) from the MOZAIC programme, Atmos. Chem. Phys., 6, 1033–1051, doi:10.5194/acp-6-1033-2006, 2006.

Thouret, V., Saunois, M., Minga, A., Mariscal, A., Sauvage, B., Solete, A., Agbangla, D., Nédélec, P., Mari, C., Reeves, C. E., and Schlager, H.: An overview of two years of ozone radio soundings over Cotonou as part of AMMA, Atmos. Chem. Phys., 9, 6157–6174, doi:10.5194/acp-9-6157-2009, 2009.

Vömel, H. and Diaz, K.: Ozone sonde cell current measurements and implications for observations of near-zero ozone concentrations in the tropical upper troposphere, Atmos. Meas. Tech., 3, 495–505, doi:10.5194/amt-3-495-2010, 2010.

Wernli, H. and Davies, H. C.: A Lagrangian-based analysis of extratropical cyclones, I: The method and some applications, Q. J. Roy. Meteorol. Soc., 123, 467–489, 1997.