Validation of global gas-phase chemistry models with particular emphasis on Chemistry Climate Models (CCM) and hindcast modelling

Prepared by Johannes Staehelin for the IGAC/SPARC Chemistry-Climate Model Initiative (CCMI)

With contributions from Greg Bodeker, Dominik Brunner, Neil Harris, David D. Parrish, Martin Schultz, and Simone Tilmes

December, 2012

1. Introduction

This document aims to contribute to the validation of Chemistry Climate Models (CCM) using ground-based measurements, as planned within the IGAC/SPARC Chemistry-Climate Model Initiative (CCMI). It is important to keep in mind which ground-based measurements are available and, more importantly, suitable for comparison with global numerical simulations. Ground-based measurements play different roles in the validation of stratospheric and tropospheric numerical simulations: satellite measurements have been extensively used for validation in the stratosphere (for the period after 1979), but less so in the troposphere. It is therefore useful to distinguish between validation of stratospheric (Section 2) and tropospheric numerical models (Section 3). Section 4 attempts to summarise the main results of the report. The document was sent for discussion to the members of the working group on "Evaluation of numerical simulations by ground based measurements" (Isabelle Bey, Greg Bodeker, Dominik Brunner, Veronika Eyring, Arlene Fiori, Neil Harris, Jean-Francois Lamarque, Jennifer Logan, David D. Parrish, Martin Schultz, and Simone Tilmes) and was used as a basis for discussions at the IGAC/SPARC Global Chemistry-Climate Modelling and Evaluation workshop, which took place 21-24 May, 2012 in Davos (Switzerland). The document was finalized in December 2012.

2. Validation of stratospheric model results with ground-based measurements

2.1 Introduction and General Remarks

Since the second part of the 1970s a global network of ground-based sun spectrophotometers has been operated under the auspices of the WMO, providing high quality measurements of column ozone at selected sites (as part of Global Atmosphere Watch (GAW)). The network has long-term calibration stability with a drift of less than 1% per decade (Komhyr et al., 1989). Other long-term, ground-based ozone profile measurements with known data quality are made within the networks of GAW, NDACC (Network for the Detection of Atmospheric Composition Changes) and SHADOZ (Southern Hemispheric ADitional OZonsondes) (see section 2.2).

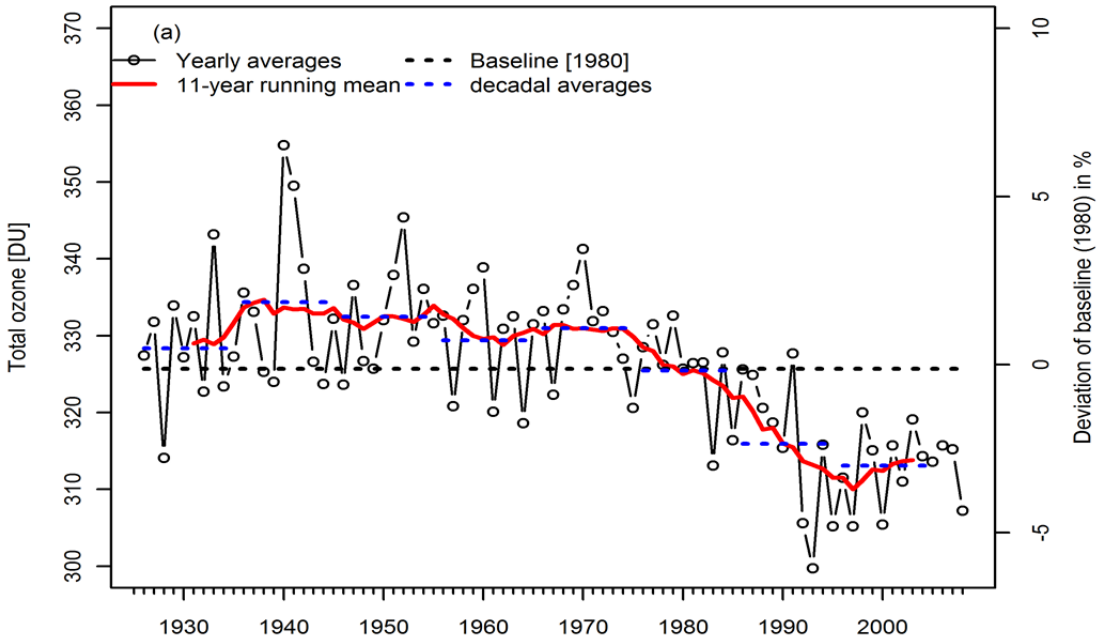

Figure 1: Total ozone series of Arosa (Switzerland) (Staehelin et al., 1998, updated).

Since 1979, quasi-continuous and quasi-global satellite measurements are available, allowing the documentation of the spatial distribution and temporal changes of ozone and several other important stratospheric trace gases (see for example the SPARC "Data Initiative" activity). High quality, long-term, ground-based measurements are vital not only for satellite validation but also to ensure the temporal stability of satellite products and for critical assessment of merged satellite data (note that all satellite instruments have limited lifetimes, which requires combination (merging) of satellite series to produce complete long-term datasets. Ground-based instruments on the other hand can be examined more easily). After 1979 the direct use of ground-based measurements for validation of numerical simulations (such as used in hindcast modelling) is less attractive than the use of satellite products, if the quality of the satellite data has been properly assessed (again, see for example the SPARC "Data Initiative").

Section 2.2 of the document contains an inventory of ground-based measurements suitable for comparison with numerical simulations of the stratosphere and Section 2.3 includes some additional remarks important in the context of the validation of these simulations using ground-based measurements.

2.2 Inventory of ground-based stratospheric measurements suitable for model validation

2.2.1 Column ozone

Precise total (column) ozone measurements became possible with the construction of the Dobson spectrophotometer in the 1920s. Today, Dobson spectrophotometers with the same basic instrumental design as that of the 1950s remain one of the backbones of the GAW total ozone network (Komhyr, 1980). In the 1980s the Brewer instruments (Kerr et al., 1988), essentially based on the same general design (also making use of atmospheric ozone absorption in the Huggins band), extended the Dobson network, and from the 1990s onwards, SAOZ instruments (measuring ozone absorption in the visible, see 2.2.3; Pommerau and Goutail, 1988) have also been operated.

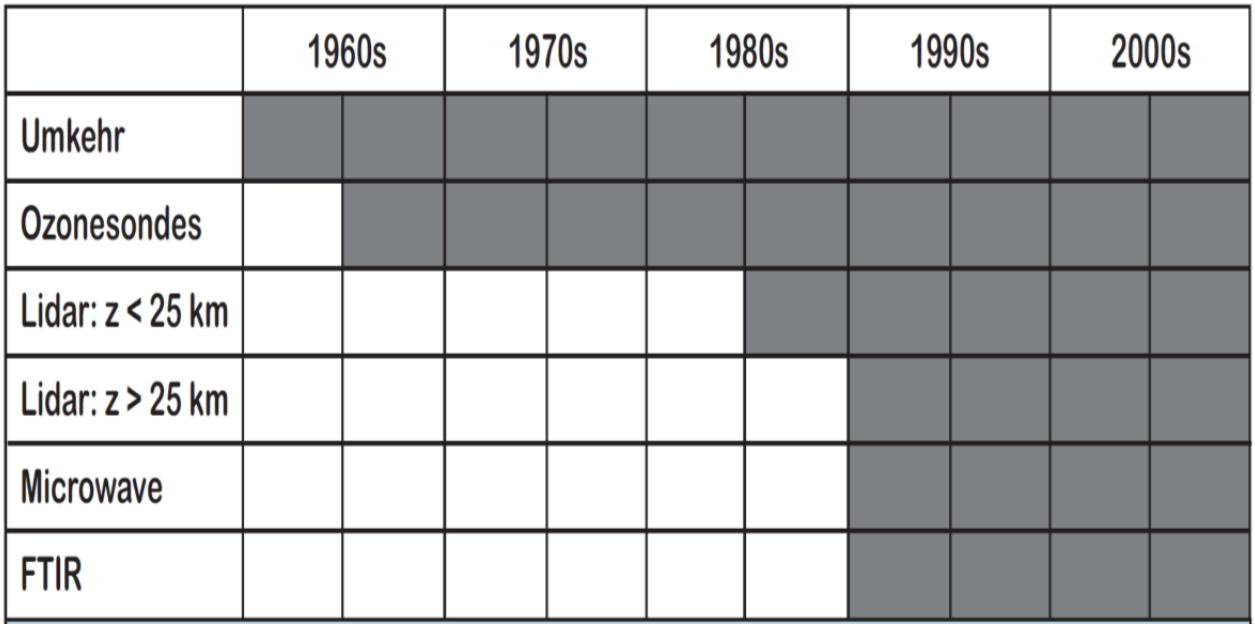

Figure 2: Availability of ground based measurements providing information regarding the vertical distribution of stratospheric ozone and its long-term changes (from Harris et al., 2011).

Ground-based total ozone measurements are most directly relevant for comparison with numerical simulations for the period prior to 1979, i.e. before satellite measurements became available. They are the only data source providing reliable information of (stratospheric) ozone variability and its changes on inter-decadal temporal scales, covering the decades prior to the period when Ozone Depleting Substances (ODSs) started to affect the ozone layer (see Fig. 1). However, since such measurements require sunlight, they are more restricted and less precise in the polar areas and strongly limited outside the northern mid-latitudes. Note that total ozone measurements from before the middle of the 1970s need careful evaluation, and in most cases homogenisation is required as well (Brönnimann et al., 2003; Vogler et al., 2007). These historical measurements can be used for comparison with numerical simulations of the stratosphere at single sites after 1980, but they may be more useful to evaluate the long-term reliability of satellite measurements than for direct use in validation of numeric models. Total ozone observations from Dobson and Brewer spectrophotometers are available at the Word Ozone and Ultraviolet Data Centre (WOUDC), operated by Environment Canada (www.woudc.org/).

2.2.2 Vertical distribution of Ozone

Fig. 2 shows an overview of ground-based measurements that provide information regarding the vertical distribution of ozone. The first ozone profile observations were obtained using the Umkehr method (Götz et al., 1934) and continuous Umkehr series are available from a few sites from about 1960. Umkehr profile measurements can be performed using both Dobson and Brewer spectrophotometers (measurements stored at the WOUDC). Ozonesondes also provide in-situ measurements, with continuous data available since the second part of the 1960s. For stratospheric measurements (where ozone concentrations are much larger than in the troposphere) Brewer Mast (BM) sensors are believed to provide reliable data (even if long-term tropospheric ozone changes deduced from BM sensors are judged to be unreliable, see section 3.3). Note that some of the ECC sonde measurements require homogenisation (see section 3.3 for information regarding the archives). The measurements of NDACC, which started around 1990, are shortly described in Section 2.3.

Every instrument has individual characteristics and strengths (e.g. regarding optimal height and vertical resolution), which need to be taken into account when comparing with numerical simulations. All remote sensing observations included in Fig. 2 require the application of averaging kernels to ensure accurate comparison with numerical simulations. Long-term ozone profile evolution is presently studied within the "Past Changes in the Vertical Distribution of Ozone" initiative (Harris et al., 2011). This project aims to use all available long-term ozone profile observations to critically assess ozone changes, particularly for the period after 2005 when the SAGE-II instrument stopped operating.

2.2.3 Ground-based remote sensing measurements performed by NDACC

Several ground-based instruments are operated within the NDACC network, which is also responsible for data quality assurance of the measurements. Most NDACC measurements started in the 1990s (see Fig. 2). The NDACC network includes the following instruments measuring the listed species (see www.ndsc.ncep.noaa.gov/ for more information regarding instrument locations and data availability):

FTIR: Partial and total column of selected species, including CH_4, C_2H_6, $ClONO_2$, CO, HCl, HCN, HF, HNO_3, N_2O, O_3. Total column of a broad range of species, including CCl_2F_2, CHF_2Cl, CO_2, COF_2, H_2O, HDO, NO, NO_2 and OCS. FTIR measurements of ODSs show that the emissions of the substances regulated by the Montreal Protocol have been reduced as expected.

Microwave radiometer: Profiles of O_3, H_2O, ClO, CO, HNO_3 and N_2O.

LIDAR: Aerosol profiles: Backscatter LIDAR; Ozone profiles: Differential Absorption Lidar (DIAL); Tropospheric ozone profiles: DIAL; Temperature profiles: Raman and Rayleigh Lidar; Water Vapour profiles: Raman Lidar.

UV/Visible Spectrometer: Total O_3 column (e.g. using the SAOZ instrument) and NO_2.

2.3 Remarks concerning use of ground based measurements for validation of stratospheric numerical simulations

For accurate comparison of numerical simulations with field observations, the measurement uncertainty needs to be determined (although estimating uncertainty is often not a simple task for historical data). As discussed above, direct comparison of ground-based measurements is most important for the period prior to the start of continuous satellite measurements, but is expected to be less important than satellite measurements after 1980. In the latter period, reliable ground-based measurements play a crucial role for satellite validation.

In order to document the ability of CCMs (and CTMs – Chemistry Transport Models) to describe the long-term effects of the anthropogenic release of ODSs on the ozone layer (and its recovery), the comparison of reliable stratospheric ozone observations with CCMs (either as a time series when run with nudged meteorology or in the form of comparison of variability) is important (e.g. Eyring et al., 2010). In addition, it is recommended to move beyond direct comparison of CCM output with observed fields. It can be very valuable to derive 'sensitivities' in the CCMs and to compare these with their counterparts in measurements, e.g. the sensitivity of ozone to Cl_y, the sensitivity of chlorofluorocarbon lifetimes to the strength of the Brewer Dobson Circulation, or the sensitivity of the rate of changes of HNO_3 to PSC volume. Deriving semi-empirical equations that describe key atmospheric processes and then fitting those equations to observations and model fields, and then comparing those fit parameters, can be a very powerful use of observations to validate models. Some examples are shown in the PhD thesis of Stefanie Kremser (to be published).

3. Validation of tropospheric model results with ground-based and regular aircraft measurements

3.1 Introduction

In this section ground-based measurements suitable for comparison with CCMs and CTMs describing global tropospheric gas-phase chemistry are discussed. This report also addresses data quality problems. Recently, Logan et al. (2012) studied the coherence of long-term European ozone measurements. Another recent publication by Parrish et al. (2012) describes long-term changes in baseline ozone in northern mid-latitudes. The "base state of ozone" between 1995 and 2009 is discussed in Tilmes et al. (2012), including information regarding longitudinal differences in several latitude bands and considering ozone probability distribution functions.

Urban and suburban air quality measurements are not covered here and this report only includes regular measurements (including those from aircraft). Measurements made during research campaigns are not part of this report (although they can provide valuable information, for example, for comparison with CTM models (e.g. Emmons et al., 2000; Brunner et al., 2003, 2005). Remote sensing measurements, e.g. from LIDAR and FTIR (such as provided by NDACC, see Section 2.2.3) and tropospheric satellite trace gas measurements are also not included in this document. Section 3.2 describes available surface measurements, Section 3.3 ozonesondes, Section 3.4 measurements from regular aircraft, and Section 3.5 highlights some important points for comparisons with numerical simulations, with some particular emphasis on the validation of hindcast simulations.

3.2 Surface measurements of trace gases

For the availability of surface measurements also see Oltmans et al. (2006) and Gilge et al. (2010).

3.2.1 Ozone

Europe

Surface ozone measurements at European sites document an increase of approximately a factor of two between the late 1950s and the early 1990s, at rural and remote sites extending from the boundary layer to the free troposphere (as captured by the high alpine sites such as Jungfraujoch) (Staehelin et al., 1994; Parrish et al, 2012). The results mainly originate from campaigns, but they are confirmed by continuous measurements from *Arosa* (Switzerland, 46°46'N, 9°12'E, 1860 meters above sea level (m asl.); see Fig. 3). The 1950s likely reflect atmospheric conditions at the beginning of the period of large economic growth after World War II, in which anthropogenic emissions from industrialised countries increased strongly. Note that the instruments used during this period are sensitive to SO_2 interference and only historical measurements from rural and remote sites should be used for comparison with numerical simulations, since a large part of the urban European planetary boundary layer was strongly polluted by SO_2 during the time of the historical measurements.

The only quasi-continuous time series (measured using chemical methods at the beginning) going back to the late 1950s originates from *Arkona* (1956-83). Measurements were combined with surface ozone observations of *Zingst*, a site close to *Arkona* (see Fig. 4). The sites are located close to the North Sea. Continuous ozone measurements at the rural village *Hohenpeissenberg* (southern Germany, 47°63'N, 11.°01'E, 985 m asl.) started in 1971 (see Gilge et al., 2010, and Fig. 4).

High mountain (alpine) sites largely remain in the free troposphere in winter, whereas spring and summer measurements can be affected by air from the polluted planetary boundary

layer, which is transported by thermally induced wind systems. Occasionally convectively mixed air can reach altitudes of the high mountain sites, however, effects of such transport are rather limited for ozone, particularly for long-term ozone changes (e.g. Cui et al., 2011).

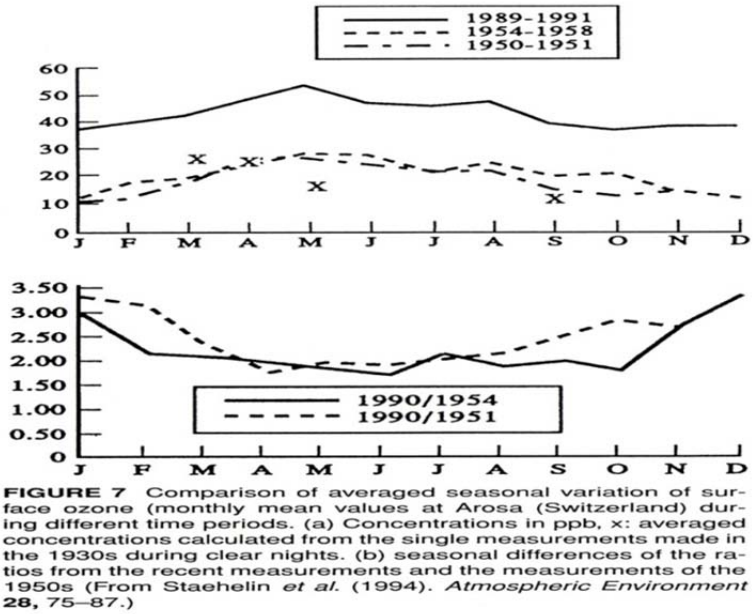

Figure 3: From Staehelin et al., 1994.

FIGURE 7 Comparison of averaged seasonal variation of surface ozone (monthly mean values at Arosa (Switzerland) during different time periods. (a) Concentrations in ppb, x: averaged concentrations calculated from the single measurements made in the 1930s during clear nights. (b) seasonal differences of the ratios from the recent measurements and the measurements of the 1950s (From Staehelin *et al.* (1994). *Atmospheric Environment* **28,** 75–87.)

Measurements at *Zugspitze*, Germany (47°42'N, 10°98'E, 2962/2670 m asl.) started in 1975, at *Jungfraujoch*, Switzerland (46.55 N, 7.98 E, 3580 m asl.) in 1986 (however, measurements up to the early 1990s have been identified as suspicious, see Zanis et al. (1999)) and at *Sonnblick* ("Hoher Sonnblick"), Austria (47°03'N, 12°57'E, 3105 m asl.) in 1989.

The three mountain sites show coherent variability only since about 1998 (see Fig. 5; for more details of the consistency of the three records see Logan et al., 2012), nevertheless, they show a qualitatively similar picture with increasing concentrations until 1998, followed by stabilisation or a decreasing tendency (see Fig. 4, comp. 3.5.3).

Arosa. Continuous ozone measurements are available from Arosa from 1989 onwards. Arosa is the site where continuous ozone measurements were performed during the 1950s. Ozone concentration changes observed at Arosa are qualitatively similar to the alpine site Jungfraujoch (see Fig. 4 and Parrish et al., 2012).

Surface ozone has been measured since the late 1980s close to *Kislovodsk* (43.70°N, 42.70°E, 2070 m asl.), a resort village in the Caucasus, Russia. The measurements show a distinct downward trend particularly during the 1990s, different from the measurements at Jungfraujoch (Tarasova et al., 2009).

Measurements at the marine site *Mace Head* (coastal site in Ireland, 53.2N 9.54W 25 m asl.) started in 1987 (see Derwent et al., 2007), and at *Izana*, Tenerife, Spain (28.3N 16.5W 2800 m asl.) in 1988. Mace Head shows a similar tendency in the long-term ozone evolution as the high alpine sites, with an increase until the end of the 20th century followed by a stabilisation (see Fig. 4).

Extended and reliable surface ozone data from many European rural sites are available for analysis of long-term changes since 1995 (see e.g. Wilson et al., 2012).

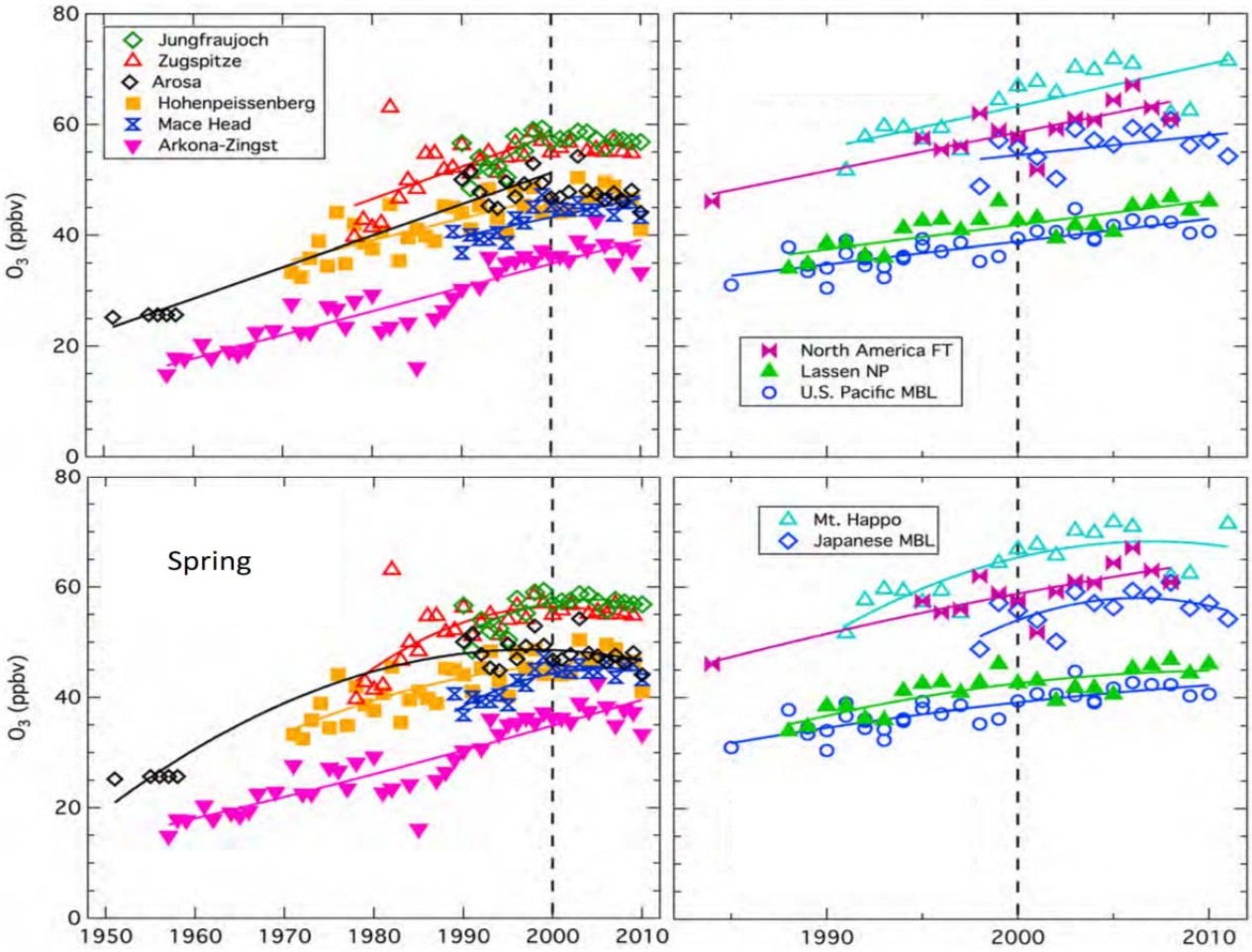

Figure 4: Long-term changes in ozone during spring from reliable baseline sites (from Parrish et al. (2012); HTAP, 2010). The left panel includes European sites, the right panel sites from the American West coast and Japan. The top panel shows fitting of the data with a linear fit, the bottom panel fitting with a quadratic fit (for more detail see Parrish et al., 2012).

North America

Surface ozone measurements are available at:
Whiteface Mtn., *NY*, USA (44.4°N, 73.9°W, 1484 m asl.) from 1973,
Lassen National Park (NP), Calif., USA (40.5°N, 121.6°W, 1756 m asl.) from 1987,
Glacier NP, Montana, USA (48.5°N 114.0°W, 976 m asl.) from 1989.
and from polar sites at:
Barrow, Alaska, USA (71.1°N, 156.6°W, 11 m asl.) from1973,
Alert, Nunavut, Canada (82.5°N, 62.3°W, 62 m asl.) from 1992, and
Denali NP, Alaska, USA (63.7°N, 149.0°W, 661 m asl.) from 1987.

A large number of surface ozone series at non-urban (i.e. including rural and remote) sites is available since 1997 from the Canadian Air and Precipitation Monitoring Network (CAPMoN), (http://www.msc-smc.ec.gc.ca/natchem/), the Canadian National Air Pollution Surveillance Network (NAPS) (http://www.etc-cte.ec.gc.ca/) and the United States Clean Air Status and Trends Network (CASTNET) of the US Environmental Protection Agency (US EPA) and the National Park Service (NPS). Data from CASTNET and NPS are available from http://epa.gov/castnet/. Chan and Vet (2010) used ninety-seven non-urban measurement sites (starting in 1997), spanning latitudes from approximately 29° N to 55° N, longitudes from 65° W to 123° W and altitudes from 2-3178 meters above sea level (asl.).

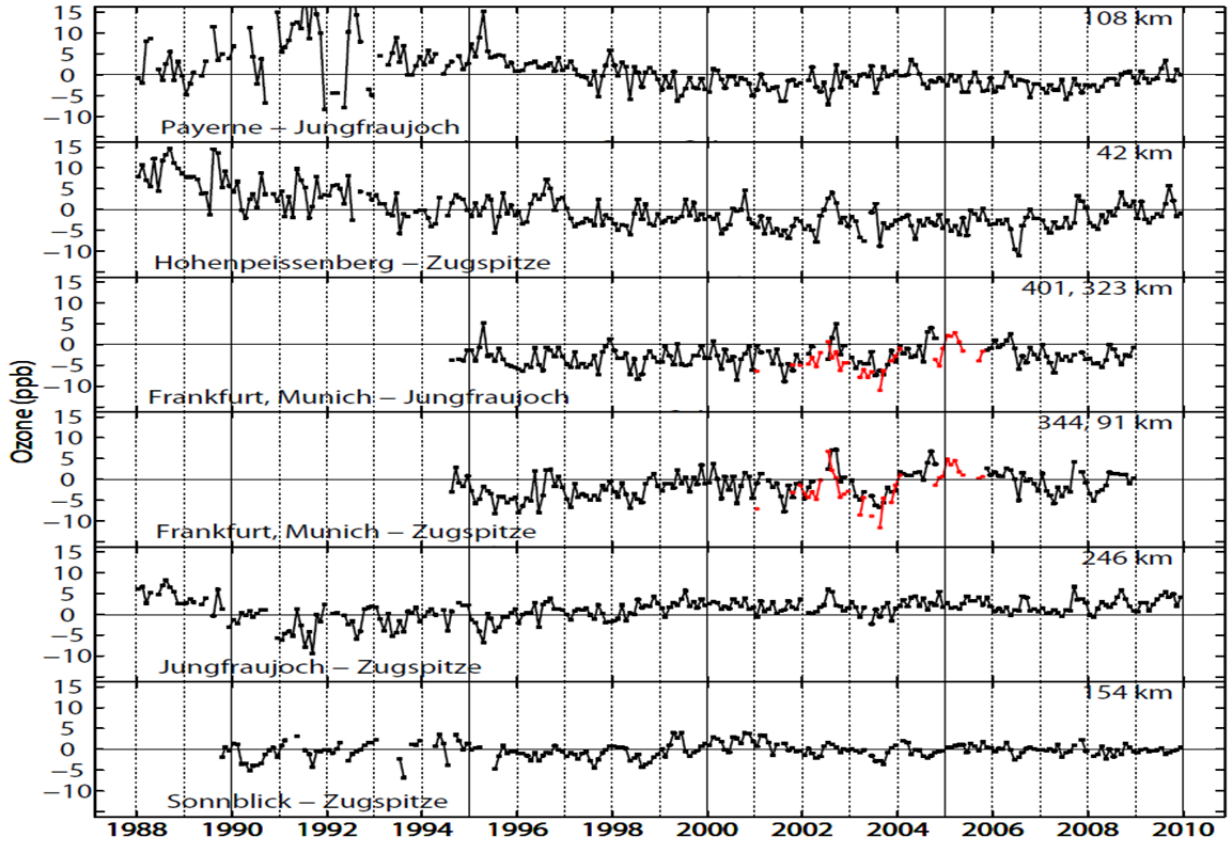

Figure 5: Difference of monthly mean ozone concentrations between pairs of locations. The level at 681 hPa was used for the profile data sets. Distances between the locations are given in the upper right corner (Logan et al., 2012).

Surface baseline ozone evolution in the Western part of North America was studied in several papers (see Fig. 4, Parrish et al., 2009). Cooper et al. (2010) presented the most comprehensive analysis of these data, showing a long-term increase in ozone mixing ratios in the free tropospherie during spring.

Downwind of North America:

Observations are available from

Bermuda (32° 16' N, 64° 53' W, 30 m asl.) from 1989, and

Sable Island, Nova Scotia, Canada (43° 56' N, 60° 01' W , 20 m. asl.) from1991.

Asia

Long-term surface ozone measurements are available from the Acid Deposition Monitoring Network in East Asia (EANET) (http://www.eanet.cc/product/index.html) (comp. e.g. Fiori et al., 2009; HTAP, 2010). High quality ozone observations are made at 11 EANET sites in Japan. The longest continuous surface sites are available from the Japanese site at *Mt. Happo* (36.69°N, 137.80°E, 1850 m asl., started in 1991, see Fig. 4 and Tanimoto (2009). One concern, however, is that the most recent data from Mt. Happo are 10-20 ppbv below the mixing ratios measured in earlier years. After 2007, the traceability system and operational protocol for monitoring ambient O_3 in Japan was modified, and the EANET instruments were replaced in 2009, so a systematic error is suspected. Efforts are underway to attempt to resolve this discrepancy. Regarding regional representativeness for model comparison, the following six additional sites are expected to be most suitable: *Rishiri* Island (45.12°N, 141.24°E, 40 m asl., see Fig. 4), Cape *Tappi* (41.25°N,

140.35°E, 106 m asl.), *Sado-seki* (Island) (38.25°N, 138.40°E, 110 m. asl.), *Oki* Island (36.28°N, 133.18°E, 90 m. asl.), *Hedo* and *Ogasawara* Island (27.09°N, 142.21°E, 230 m. asl.). The observation periods of these sites are around 10 years (H. Akimoto, pers. communication and Tanimoto et al., 2009). Measurements from other EANET sites need to be checked for suitability and data quality.

Tropics (subtropics) of the Pacific (see Oltmans et al., 2006).
Surface ozone measurements are available from:
Mauna Loa, Hawaii, USA (19.5°N, 155.6°W, 3397 m asl.) from 1973, and
Matatula Pt., *Am. Samoa* (14.3°S, 170.6°W, 82 m asl.) from 1975.

Southern hemisphere, extratropics
The longest continuous surface ozone data originate from *Cape Grim*, Australia (40.7°S, 144.°7E 104 m asl.), with observations starting in 1982. Other data are available from:
Cape Point, South Africa (34.4°S, 18.5°E, 230 m asl.) from 1983,
Baring Head, New Zealand (41.4°S, 174.9°E, 85 m asl.) from 1991, and
South Pole, Antarctica (90.0°S, 2840 m asl.) from 1975.

Table 1: Availability of carbon monoxide measurements from flask sampling of NOAA/CMDL (Novelli et al., 1998)

Name of site	longitude	latitude	elevation (1)	duration	MBL *
Alert, N.W.T.	82°27'N	62°31'W	210	1985-present	y
Ascension Island	7°55'S	14°25'W	54	1989-present	y
Baltic Sea (Balanga Sister)	55°30'N	16°40'E	7	1992-present	n
St. David's Head, Bermuda	32°22'N	64°39'W	30	1989-present	y
Southhampton Bermuda	32°16'N	64°53'W	30	1989-present	y
Barrow, Alaska	71°19'N	156°36'W	11	1988-present	y
Cold Bay, Alaska	55°12'N	162°43'W	25	1992-present	y
Cape Grim, Tasmania	40°4 l'S	144°4l' E 94		1991-present y	
Christmas Island	1°42'N	157°10'W	3	1989-present	y
Cape Meares, Oregon	45°29'N	123°58'W	30	1992-present	n
Guam, Mariana Islands	13°26'N	144°47'E	2	1989-present	y
Dwejera Point, Gozo, Malta	36°03'N	14°11'E	30	1993-present	n
Hegyhatsal, Hungary	46°58'N	16°23'E	240	1993-present	n
Vestmannaeyjarl, Iceland	63°15'N	20°09'W	100	1992-present	y
Tenerife, Canary Islands	28°18'N	16°29'W	2300	1991-present	n
Key Biscayne, Florida	25°40'N	80°12'W	3	1991-present	y
Cape Kumukahi, Hawaii	19°3l' N	154°49'W	3	1989-present	y
Mould Bay, N.W.T.	76ø15'N	119°21'W	58	1992-19975	y
Mace Head, Ireland	53°20'N	9°54'W	25	1991-present	y
Sand Island, Midway	28°13'N	177°22'W	4	1992-present	y
Mauna Loa, Hawaii	19°32'N	155°35'W	3397	1989-present	n
Niwot Ridge, Colorado	40°03'N	105°35'W	3475	1989-present	n
Pacific Ocean (California Star)			8	1993-present	y
Pacific Ocean (Wellington tar)			7	1990-19935	y
Palmer Station, Antarctica	64°55'S	64°00'W	10	1994-present	n
Qinghai Province, China	36°16'N	100°55'E	3810	1990-present	n
Ragged Point, Barbados	13°10'N	59°26'W	3	1992-present	y
South China Sea (Carla A Hills)			15	1991-1993	n
South China Sea (Great Promise)			15	1993-present	n
Seychelles, Mahe Island	4°40'S	55°10'E	3	1990- present	y
SHM Shemya Island, Alaska	52°43'N	174°06'E	40	1992-present	y
Tutuila, American Samoa	14°15'S	170°34'W	42	1988-present	y
South Pole, Antarctica	89°59'S	24°48'W	2810	1993-present	y
Syowa, Antarctica	69°00'S	39°35'E	11	1983-present	y
Tae-ahn Peninsula, Korea	36°44'N	126°08'E	20	1990-present	n
Ulaan Uul, Mongolia	44°27'N	111°06'E	914	1992-present	n
Wendover, Utah	39°54'N	113°43'W	1320	1993-present	n

*Elevation in metres above mean sea level

Measurements from Cape Grim show a continuous increase in ozone mixing ratios until the beginning of the 21st century (I. Galbally, pers. commun.). This increase is much smaller in magnitude (0.07 ppbv/yr) than the trend in the free troposphere over Europe (about 0.4 to 0.5 ppb/yr).

Further surface ozone observations from rural and remote sites (including shorter time series') are archived at the World Data Centre for Greenhouse gases (WDCGG, http://gaw.kishou.go.jp/wdcgg/).

3.2.2 Carbon monoxide

Global network

Extended carbon monoxide measurements are available from the NOAA/CMDL flask sampling program (see Novelli et al., 1998 and Table 1). The longest data series started in 1985 at Alert, Canada (82^{o}N), while most other observations began around 1990. The network also includes measurements in the Southern hemisphere (7 stations), however, the number of sites in the Northern hemisphere is much larger (26 sites).

Europe (see Gilge et al., 2010)

Continuous CO measurements are available at *Hohenpeisseneberg* from 1995, at *Jungfraujoch* from 1996, and at *Sonnblick* from 2002. Measurements at *Zugspitze* started in 1995 and ended in 2002. The CO measurements show decreasing mixing ratios in recent years.

At several of the flask sampling network sites CO is also continuously measured (e.g. Mace Head (http://macehead.nuigalway.ie/mace1.html) and Mauna Loa).

3.2.3 PAN (Peroxyacetylnitrate)

Continuous measurements of PAN are sparse. The best information available is included in Fiori et al. (paper in prep.):

Europe

PAN was continuously measured at *Jungfraujoch* from April 1997 to May 1998, as well as from May to October 2008 and May 2009 to June 2010. Campaign measurements are available from 30 Aug. 2005 to 16. Sept. 2005 and during a few campaigns in 2005 (see also Pandey Deolal et al. 2013). At *Zugspitze* PAN has been continuously measured since May 2004. At *Hohenpeissenberg* PAN was measured from January 2003 to December 2008.

North America

PAN was measured at *Mount Bachelor,* USA, (43.979°N, 121.687°W, 2763 m asl.) during campaigns from 3 April – 18 June 2008, 30 August – 7 October 2008, 26 March – 20 May 2009, 23 March – 25 May 2010, and at *Big Hill, USA* (38.84°N, 120.41°W 1860 m asl.) from March 2003 to February 2004. PAN measurements are also available from airborne campaigns along the Pacific Coast of North America (Parrish et al., 2004), as well as more recently from the INTEX-B campaign.

Asia

PAN was measured at *Mt. Waliguan,* China (36.29°N, 100.90°E, 3816 m), July 22 – August 16 2006.

Note that at the European alpine sites the interpretation of the seasonal cycle of PAN measurements needs to take into account upslope wind transport in spring and summer, since the concentration gradient between the polluted planetary boundary layer and the free troposphere is large (particularly in spring, and is much larger than for ozone mixing ratios) (Pandey Deolal et al., 2013). PAN simulations from HTAP show large differences between the individual models (pers. commun. A. Fiori). This is possibly attributable to different parametrisations of NMVOC chemistry, possibly being related to the production rate of acetaldehyde from hydrocarbons (if, for example, NMVOC chemistry of a model is limited to alkanes and alkenes (i.e. no aromatics) production of PAN might be small).

3.2.4 Nitrogen oxides (NO_x) and NO_y

Ambient air concentrations of nitrogen oxides are measured in many urban/suburban air pollutant networks, however, the measurements are not representative for larger areas and the commonly used sensors are not suitable for NO_2/NO_x measurements at rural and remote sites (not specific for NO_2, too high detection limits). Continuous series of NO_x and NO_y from rural and remote sites are sparse.

NO_x is measured at *Jungfraujoch* since 1995 (Pandey Deolal et al., 2012) and at *Hohenpeissenberg* since 1996. (The apparent decreasing trend of NO_x at Junfgraujoch (see Gilge et al., 2010) is most likely an artifact since the mixing ratios are very large during the first two years probably due to construction work at the Jungfraujoch). Continuous NO_y measurements started at *Jungfraujoch* in 1998, the measurements show only recently a small decrease in mixing ratios (Pandey Deolal et al., 2012).

3.2.5 Volatile Organic Compounds (Volatile Organic Compounds)

Continuous measurements of individual hydrocarbons are available from selected sites where NOAA/GMD canisters are collected or by automated instruments (e.g. Jungfraujoch) (Helmig et al., 2009). The ratios of individual VOCs (hydrocarbons) contain valuable information concerning age of air. Many models treat at least ethane and propane explicitly, and some information for direct comparison with numerical simulations is available if one considers species as groups, such as "butanes and higher alkanes". There are a few FTIR measurements of ethane columns. The use of such measurements for comparison with global numerical simulations is, nevertheless, not straight forward, since NM-VOC chemistry is always parametrised in these models.

3.3 Ozonesondes

Ozonesondes are presently launched from many sites (see Fig. 6). Different sensors are used at different sites and the launch frequency strongly differs from site to site as well (ozonesondes are usually launched three times per week at only three European sites). Ozonesonde data are presently stored in different archives (WOUDC; NDACC; SHADOZ; NOAA) using somewhat different formats.

Ozonesonde measurements relevant in this context are performed with the following types of sensors (for further details see Smit et al., 2011):

1. The first regular ozonesonde measurements were performed using **Brewer Mast (BM)** sensors. BM sensors were used at the three long-term European sites (Hohenpeissenberg, Uccle and Payerne) in the earlier decades (as well as in Canada during the earliest period of the measurement record). The data quality of earlier ozonesondes is difficult to assess and concerns regarding the data quality of earlier tropospheric BM ozonesonde measurements (prior to around 1998) remain (see below). These tropospheric measurements should not be used for long-term trend analysis and comparison with numerical simulations.
2. Most stations presently use **Electrochemical Cell (ECC)** sondes. However, two different manufacturers produce ECC sondes (ENSI and SPI) and different solute concentrations have been used. Inappropriate combinations of ECC sondes and solute concentrations can cause deviations of up to 10% in tropospheric ozone readings (Smit et al., 2007). Particularly during the 1990s some stations used the wrong configurations. Because of such changes, the ECC ozonesonde data from several sites require homogenisation. A coordinated effort to homogenise (ECC) ozonesondes is planned as part of the on-going Ozone Profile Initiative of SPARC/IOC/IGACO-O3/UV(GAW)/NDACC (see Harris et al., 2011).

3. Another ozonesonde type is used in the Indian network. These measurements seem less accurate and they must be checked for data quality prior to comparison with numerical simulations.
4. Another type of sensor has been used in the Japanese network

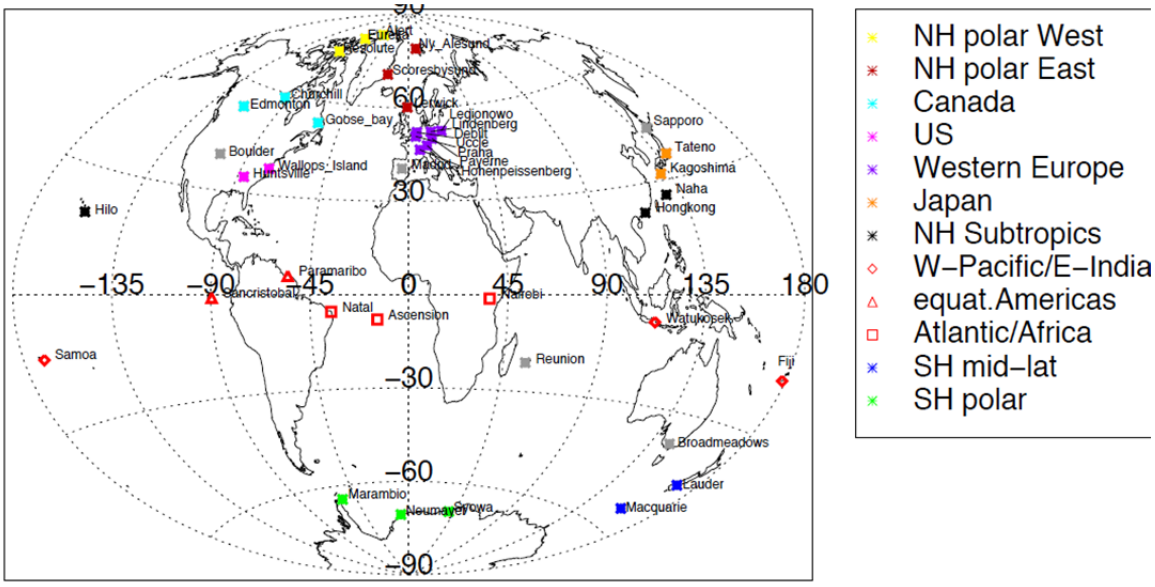

Figure 6: Ozonesonde stations covering at least six years of data from 1995 to 2009. Different colours denote different regions: NH Polar East: brown, NH Polar West: yellow, Canada: cyan, Eastern United States of America (US): pink, Western Europe: purple, Japan: orange, NH Subtropics: black, SH Mid-latitudes: blue, SH Polar: green. The tropics: red, are further divided into: West Pacific / East Indian Ocean (diamonds), equatorial Americas (triangles), and Atlantic / Africa (squares), following the selection of regions by Thompson et al., 2011, submitted. Stations that are not included in selected regions are shown in grey.

Tilmes et al. (2012) combined ozone stations with similar ozone characteristics (see Fig. 6) into regions.

Below follows a short list of available long-term ozonesonde data (based on ECC sondes unless otherwise noted):

Europe

Regular ozone sonde measurements at the following European sites were performed with BM sondes for the earlier part of the measurement period, thereafter with ECC sondes. Three ascents per week were planned from the beginning.

- *Hohenpeissenberg* (Southern Germany, 47.8°N, 11.0°E, 975 m asl.) measurements started in 1966,
- *Uccle* (Belgium, near Brussel, 50.8°N, 4.35°E, 100 m asl.) measurements started in 1966 (at Uccle BM sondes were replaced by ECC sensors in March 1997)
- *Payerne* (Switzerland, 46.80°N, 6.95°E, 491 m asl) started in 1968 (at Payerne ECC sensors are used since August 2002)

All three sites show strong increases in tropospheric ozone between the early 1970s and 1990, however, the temporal evolution in ozone at 500 hPa is different among the three sites (Logan et al., 1999). Logan et al. (2012) studied the coherence between European tropospheric ozone

measurements in the lower free troposphere, including ozonesonde observations, measurements at high mountain sites (Zugspitze, Jungfraujoch and Sonnblick, see 3.2) and central European MOZAIC data. They found coherent variation of the data after 1998, but pronounced deviations and incoherent variations in the early 1990s (see Fig. 5). The analysis revealed that the tropospheric part of the BM sonde observations before 1998 is problematic and should not be used for trend analysis. This supports concerns raised regarding the data quality of the earlier BM measurements by Schnadt Poberaj et al. (2009), who found remarkable differences in long-term ozone changes in the upper troposphere compared to the regular aircraft measurements of GASP and MOZAIC (see 3.4).

Other European sites (ECC sondes, one ascent per week):

- *Sodankyla (Finnland)* (67° 22' N, 26° 38' E, 180 m asl.) from1989,
- *De Built (The Netherlands)* (52.1^{o}N., 5.18^{o}E, 15 m asl.) from 1992
- *Lindenberg (Germany)* (52.2^{o}N, 14.1^{o}E, 110 m asl.) from 1992
- *Legionowo (Poland)* (52.4^{o}N, 20.95^{o}E, 95 m asl.) from 1993
- *Praha (Czech Republic)* (50.0^{o}N, 14.4^{o}E, 305 m asl.) from 1993, only available in winter and spring
- *Madrid (Spain)* (40.4^{o}N, 3.6^{o}W, 640 m asl.) from 1995, for some years less than one ascent per week
- *Lerwic (UK)* (60.^{o}N, 1.0^{o}E), from 1992, in some years less than one measurement per week

North America

The longest time series in the USA have been launched from:

- *Wallops Isl., Virginia*, USA (37.9^{o}N, 75.5^{o}E, 13 m asl.) from 1970
- *Boulder, Colorado*, USA (40.0^{o}N, 105.0^{o}W, 1745 m asl.) from 1979

Long-term ECC series (ignoring first BM measurements) are available from Canadian sites, including:

- *Edmonton*, Alberta, Canada (53.6^{o}N, 114.1^{o}W, 766 m asl.) from 1979
- *Goose*, Newfoundland, Canada (53.3^{o}N, 60.3^{o}W, 44 m asl.) from 1979
- *Churchill*, Manitoba, Canada (58.8^{o}N, 94.1^{o}W, 35 m. asl.) from 1980

Arctic sites

Ozonesondes were launched at least once a week from the following high northern latitude sites:

- *Resolute*, NWT, Canada (74.7^{o}N, 95.0^{o}W, 64 m asl.) from 1980
- *Alert*, Nunavut, Canada (82.5^{o}N, 62.3^{o}W, 62 m asl.), ozonesonde measurements were performed from 1987 to 2004
- *Eureka, Canada* (80. 5^{o}N, 62.3^{o}W) from 1992
- *NyAlesund, Spitzbergen* (78.9^{o}N, 11.9^{o}E) from 1991-2006
- *Scorebysund, Greenland* (70.5^{o}N, 22.0^{o}W) from 1989-2003

Asia

Japan

The long-term Japanese ozonesonde data were measured using a different type of sensor until recently. Presently ECC sensors are used in the Japanese ozone sonde measurements. Ozonesondes were launched from the following sites:

- *Sapporo*, (43.1^{o}N, 141.3^{o}E, 19 m asl.) from 1967
- *Tsukuba* (Tateno) (36.1^{o}N, 140.1^{o}E, 31 m asl.) from 1968
- *Kagoshima* (31.6^{o}N, 130.6^{o}E, 31, m asl.) from1968-2004
- *Naha* (26.2^{o}N, 127.7^{o}E, 27 m asl.) from 1989

Note that the number of ascents was small earlier in the measurement period.

India
Long-term measurements are available from selected sites. It appears, however, that the data require some editing for data quality prior to use.

Sub-Tropics and Tropics. Regular ozonesonde measurements (at least once a week) are available from:

- *Hilo*, Hawaii, USA (19.7°N, 155.1°W, 11 m asl.) from 1982
- *Pago Pago*, American Samoa (14.5°S, 170.5°W, 10 m asl.) from 1986

Observations are available from the following stations at least once a year for at least six years from 1995-2009:

- *Hong Kong*, China (22°N, 114°E) from 2000
- *San Cristobal*, Galapagos, (0.9°N, 110°W) from 1996
- *Paramaribo*, Surinam, (5.95°N, 55.2°W) from 1998
- *Natal,* Brazil (5.8°S, 35.2°E) from 1990-1992 and 1998-present
- *Ascension* (5.4°S, 35.4°W) from 1990-92, 1998-present
- *Nairobi*, Kenya (1.1°S, 36.3°E) from 1996
- *Watukosek*, Jawa Island (7.5°S, 112.7°E) from 1998
- *Fiji* (17.8°S, 177.2°E) from 1998
- *Reunion* (21.2°S, 55.3°E) from 1998

The observations at **extra-tropical sites in the Southern hemisphere** include:

- *Broadmeadows*, Australia (37.7°S, 144.9°E) from 1999-2003 and 2007-2009
- *Lauder*, New Zealand (45.0°S, 169.7°E, 370 m asl.) from 1986
- *Macquarie* (54.5°S, 158.9°E) from 1995-2003 and 2006-2009

High Southern Latitudes:

- *Marambio*, Antarica (64.0°S, 56°W) from 1988-1998 and 2006-2009
- *Syowa*, Antarctica (69.0°S, 39.6°E, 21 m asl.) from 1966
- *South Pole*, Antarctica (90.0°S, 2840 m asl.) from 1986

More ozonesonde measurements are available for a period of less than 6 years, particularly in the southern sub-tropics as part of the network SHADOZ, however, the number of reliable ascents depends on the site.

3.4 Measurements from regular aircraft

Continuous measurements from regular aircraft allow representative sampling along air routes. Aircraft emissions seem to affect trace gas concentrations only marginally, as a result of significant mixing. Cruise altitude is close to the tropopause in the mid-latitudes, whereas flight altitude lies in the middle to upper troposphere in the tropics. Because of the large vertical gradient of ozone close to the tropopause measurements at flight altitude should be used in sigma (θ) coordinates or scaled to tropopause altitude for useful analysis and comparison with numerical models.

3.4.1 Ozone

3.4.1.1. Global Atmosphere Sampling Program

Ozone was measured as part of the Global Atmosphere Sampling Program (GASP) in regular aircraft (from the USA) which were equipped with UV ozone sensors (see Schnadt Poberaj et al., 2007). Measurements were performed from 1975-1979 and flight routes covered large parts of the northern hemisphere, including the Pacific Ocean, Atlantic Ocean, Europe and the Far East (see top Fig. 7). Measurements are available at flight altitudes.

3.4.1.2 Measurement of Ozone and Water Vapour by Airbus In-Service Aircraft Program

Continuous measurements as part of the MOZAIC (Measurement of Ozone and Water Vapour by Airbus In-Service AirCraft)/IAGOS program started in 1994 (http://mozaic.aero.obs-mip.fr/web/). The measurements cover large parts of the northern hemisphere and some parts of the southern hemisphere (see bottom Fig. 7). Most measurements are performed at flight altitude but additional profile data are available close to airports. Measurements are not equally distributed over time, since flight routes of the equipped aircraft change. As a result of the installation of new equipment (new program IAGOS) the frequency of available measurements is presently reduced.

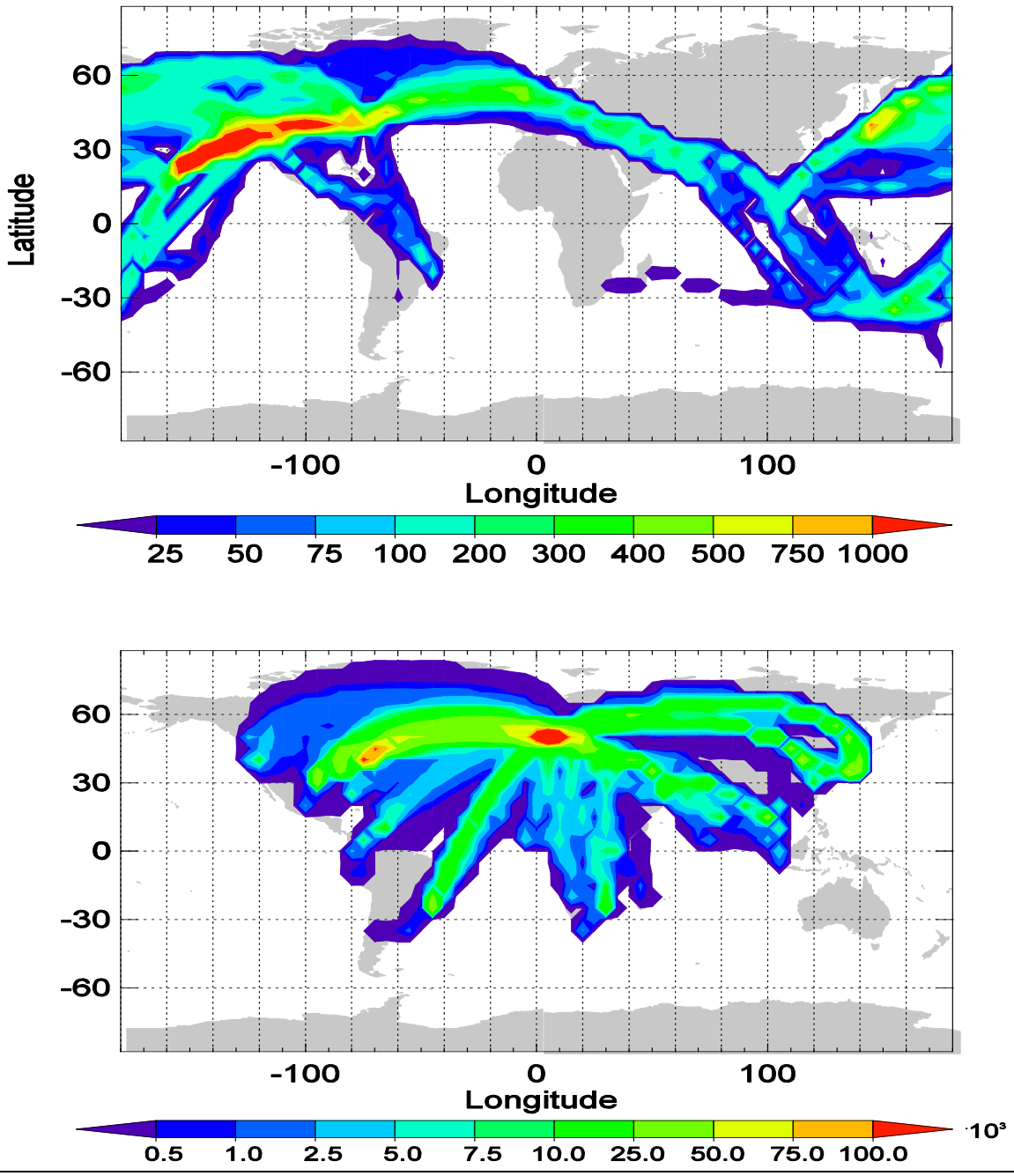

Figure 7. Flight routes (including number of flights) for GASP (top) and MOZAIC (only for the for period 1994-2002), from Schnadt Poberaj et al., 2009.

MOZAIC measurements show an increase of ozone in the upper troposphere – lower stratosphere (UT/LS) over large parts of the northern mid-latitudes approximately up to the year 2000 (see Fig. 8; Thouret et al., 2006; Zbinden et al., 2006).

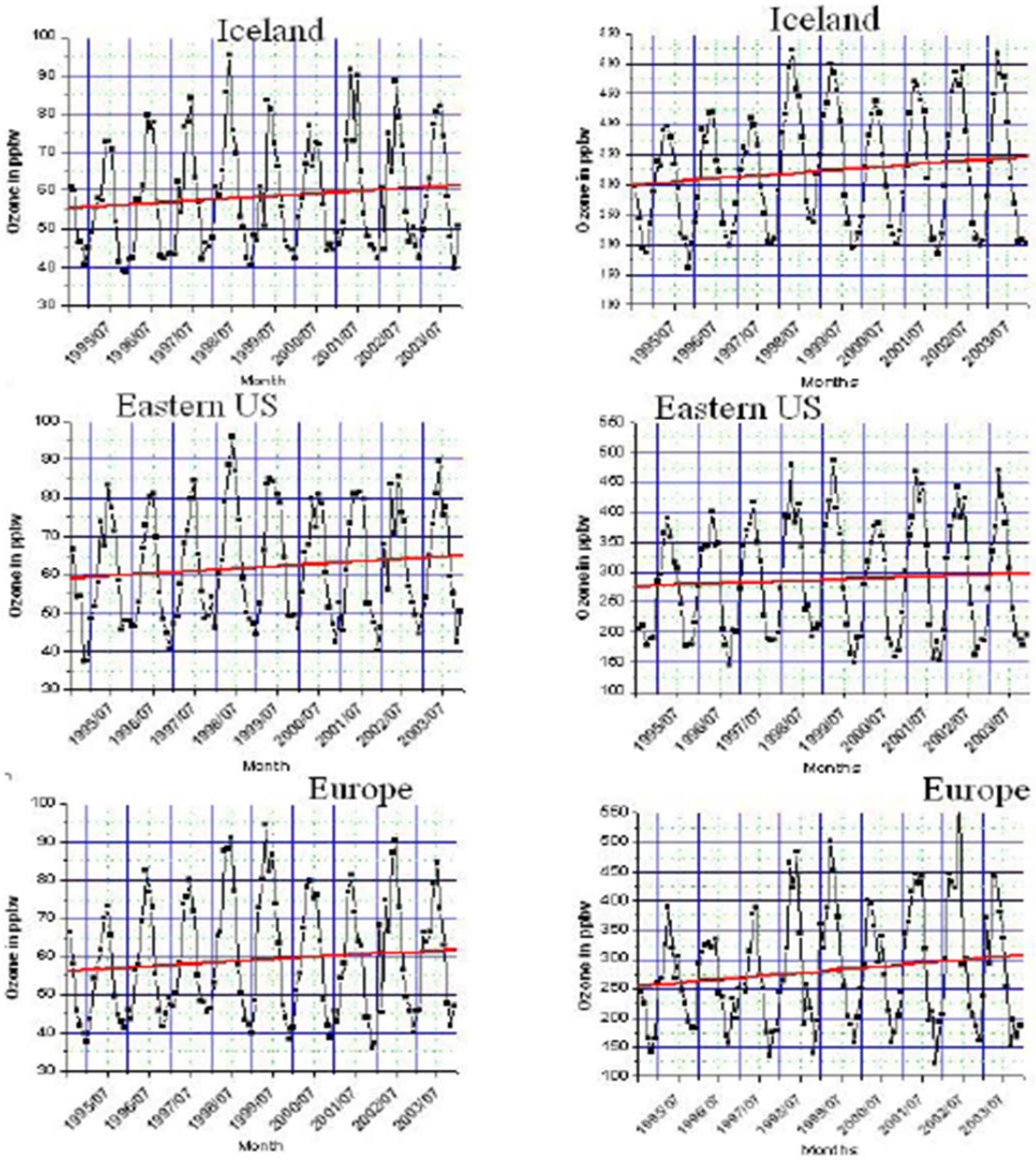

Figure 8: Time series of monthly mean ozone concentrations for the upper troposphere (left panels) and the lower stratosphere (right panels) from MOZAIC ozone measurements covering from the Eastern part of the US to Europe (from Thouret et al., 2006).

Fig. 9 shows results of a comparison of the upper tropospheric ozone climatology of GASP and MOZAIC (1994-2002) indicating large increases over Turkey, India, China and Japan. Unfortunately, the number of measurements is limited, thus large uncertainty regarding the changes in ozone remain. Ozone trends in the upper troposphere deduced from comparison of the

two regular aircraft measurements are, however, smaller than those determined from BM sonde observations at Hohenpeissenberg, Uccle and Payerne (see Schnadt Poberaj et al., 2009) supporting concerns regarding the data quality of the earlier BM ozonesonde measurements (see Fig. 5 and 3.3).

3.4.1.3. Nitrogen OXides along Air Routes

Measurements from the NOXAR (Nitrogen OXides along Air Routes) program were carried out from 1995 to 1997. The sample includes 540 long distance flights carried out between May 1995 and May 1996, including flights from Zürich, Switzerland, and destinations in the USA, as well as from Peking, Bombay and Hong Kong (see Brunner et al., 1998, 2001). In addition, 98 trans-Atlantic flights were carried out from August to November 1997. Measurements are available from cruise altitude.

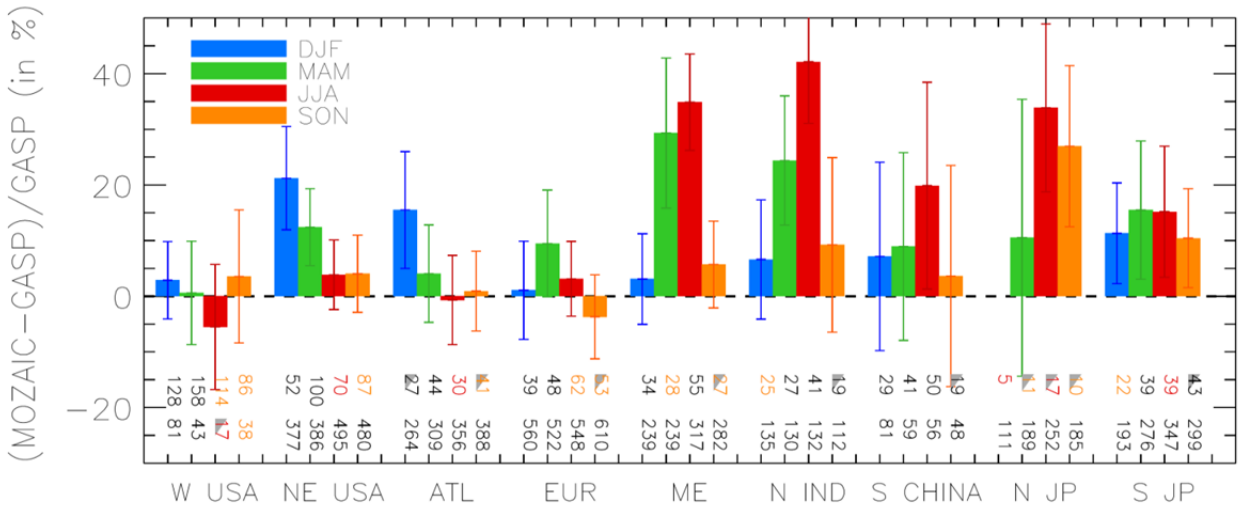

Figure 9: Long-term changes of UT ozone (MOZAIC-GASP/GASP, in %) as a function of season for nine different regions (Western USA (W USA), North-East USA (N.E. USA), Atlantic (ATL), Europe (EUR), Mediterranean (ME), Northern India (N. IND), Northern Japan (N. JP), Southern Japan (S. JP), Southern China (S CHINA)). The values are shown only if 10 or more daily regional averages are available for both GASP and MOZAIC. Vertical bars indicate the 95% confidence intervals of the relative difference. Bottom rows of numbers indicate the number of daily means available for averaging (upper row: GASP, lower row: MOZAIC). Numbers are coloured in orange (red) if the data are biased toward one year: 50–75% of data from one year (>75% from one year). Grey triangles mark regional averages for which data from only three years were available for averaging (from Schnadt Poberaj et al., 2009).

3.4.2 Nitrogen Oxides (NO_x) and NO_y

Extended NO_x measurements at flight altitude are available from the NOXAR program. For flight routes and time of measurements see 3.4.1.3. NO_y has been measured as part of the MOZAIC/IAGOS program since 2001 (see http://mozaic.aero.obs-mip.fr/web/).

3.4.3 Carbon monoxide

Carbon monoxide has been regularly measured as part of MOZAIC/IAGOS since 2001 (http://mozaic.aero.obs-mip.fr/web/).

3.4.4 Water vapour

Water vapour has been measured as part of the MOZAIC/IAGOS program from its establishment in 1994.

3.4.5 Hydrocarbons and other Volatile Organic Compounds

A variety of gas-phase organic species are measured as part of the CARIBIC program (see: http://www.caribic-atmospheric.com/).

3.4.6. Other Measurements

Other compounds which could be relevant for tropospheric model evaluation are measured from Japanese regular aircraft (program CONTRAIL (**C**omprehensive **O**bservation **N**etwork for **Tr**ace gases by **Air**liner), see: http://www.cger.nies.go.jp/contrail/index.html; http://journals.ametsoc.org/doi/full/10.1175/2008JTECHA1082.1)

3.5 Points relevant for validation of tropospheric numerical simulations with available ground-based measurements

Questions of particular relevance to the comparison of tropospheric numerical simulations and trace gas measurements are addressed in this section. Sections 3.5.1 to 3.5.3 refer to the evaluation of hindcast modelling (CCM models used in nudged mode or CTMs) with trace gas measurements, whereas section 3.5.4 includes some more general remarks concerning model evaluation using field measurements. In any case uncertainties of measurements need to be quantified or at least estimated as without defining errors and uncertainties in the measurements, it will be impossible to assess to what extent a deviation of the models from the measured distribution and/or time series constitutes an acceptable or unacceptable agreement. (However note, that exact uncertainty estimates is often difficult for historical measurements.)

3.5.1 Temporal and spatial resolution of comparison

The use of monthly mean trace gas measurements for comparison with CTM models (or simulations running in a mode nudged with meteorological fields) is only justified for continuous surface measurements. Ozonesondes are launched at most sites once per week (except at three European sites) or even less frequently. This frequency is too small to obtain a representative sample for monthly means (see Logan et al., 2012). The comparison with ozone measurements of regular aircraft needs to take into account the **spatial variability** and the data should **be compared in θ coordinates or tropopause altitude**. For comparison with $\mathbf{NO_x}$ the **diurnal variation** also needs to be taken into account. In all cases, the **spatial/temporal** coincidence between measurement location and model needs to be considered.

3.5.2 Time periods

Almost no measurements are available for the 1970s for comparison with numerical simulations (keeping in mind that European BM sonde measurements should be ignored for this purpose, except after 1998, see Fig. 5).

The large increase in ozone from the late 1950s to the early 1990s in the PBL and free troposphere over Europe seems to be a robust feature of all available measurements (see section 3.2.1). This large increase is most likely attributable to the large increase in anthropogenic ozone

precursor emissions in the industrialised world. If time slice experiments are planned, it is suggested to start the comparison with measurements in the second part of the 1950s. Since changes in meteorology are expected to be of minor importance, and keeping in mind the relative uncertainty in meteorological data prior to 1979, one might consider using meteorological fields from the 1980s in combination with anthropogenic emissions from the second part of the 1950s to compare CTM model data and check whether the models can reproduce the **ozone increase from 1960 – 1990** (as a way of a benchmarking).

Modelling the **second part of the 1970s** would be valuable for comparison with the upper tropospheric GASP measurements (including the Pacific area) as well as for comparison with MOZAIC measurements (see 3.4.1).

This report underlines that the number of trace gas measurements useful for comparison with numerical simulation significantly increases during the 1980s.

In order to make full use of the MOZAIC (3.4.1.2) and NOXAR (including upper tropospheric NO_x, see 3.4.2) measurements as well as the large number of ground-based measurements, it would be useful to start model simulations for the most **intensive comparison** with measurements in **1995**.

3.5.3 Evolution of tropospheric ozone in the free tropospheric over Central Europe from 1978

Logan et al. (2012) compared available ozone measurements for receptor sites in the lower free troposphere in the alpine/central European area (at approximately 3 km asl.) from ozonesondes, MOZAIC and several surface monitoring sites. According to this analysis the most likely evolution of ozone in this area was a large increase during the 1980s (7-13 ppb) followed by a smaller increase during the 1990s (3-7 ppb) and a small decrease from 2000-2008 (-1 to-1.5 ppb).

No similar analysis is, to our knowledge, presently available for other regions. At most other sites ECC sondes were used. For this type of sonde the preparation prior to launch seems less critical and their design is simpler for operational use. It is therefore believed, that (homogenised, see Section 3.3) ECC sonde records are useful to test the long-term evolution of tropospheric ozone.

3.5.4 Additional comments concerning evaluation of numerical simulations with field measurements

It is attractive not only to look at time series, but also to study ratios of trace gas mixing ratios (e.g. O_3/CO and PAN/CO) for comparison with numerical simulations and to consider the particular relationships both in models and field measurements (e.g. comparison of the time series of surface O_3 at Jungfraujoch and Zugspitze and ozonesonde values in the lowermost stratosphere in Central Europe during the 1990s (Ordonez et al., 2007), the comparison of ozone changes at Jungfraujoch and Kislovodsk (Caucaus, Tarasova et al. 2009), or the relation between the North Atlantic Oscillation and surface ozone in Europe (Pausata et al., 2012), etc.

4. Conclusions

4.1 General points

The following general points need to be considered when discussing the validation of global models with ground based measurements:

(i) Uncertainties and frequency distributions need to be taken into account in model-measurement comparisons;

(ii) Data quality is an important issue (avoid using the results of Schönbeinpapers, the tropospheric part of BM ozonesondes before 1998, etc.). The dialogue between

scientists with experience in data quality of particular measurements and the scientists doing numerical simulations is important because sometimes conflicting results are published in the literature;

(iii) Direct comparisons between numerical simulations and measurements are important e.g. in hindcast simulations, if the CCMs are run in a nudged mode or if CTMs are considered. In addition, comparisons of relations between variables in numerical simulations and in field measurements are very valuable for learning more about particular model aspects.

4.2 Evaluation of stratospheric numerical simulations

Long-term, ground-based measurements from Dobson spectrophotometers allow the assessment of ozone layer variability at a single site many decades before ODSs reached the stratosphere. Valuable long-term measurements of vertical ozone distributions originate from ozonesondes going back to the late 1960s at several sites, while Umkehr measurements started in the early 1960s at selected sites. NDACC measurements (not only O_3) are available since the early 1990s. All these instruments have different strengths and weaknesses, for example, regarding vertical resolution, which need to be considered when using these data for model validation. One also needs to take into account the averaging kernels to be used in the comparisons between numerical simulations and remote sensing data. Since 1979 quasi-continuous satellite measurements are available. These data are very attractive for the evaluation of numerical simulations. Ground-based measurements with known data quality remain crucial for the validation of these satellite measurements, as well as for the evaluation of their stability over time and in order to create merged datasets.

4.3 Evaluation of tropospheric numerical simulations

It is useful to consider starting hindcast model runs by the end of the 1950s, when reliable and redundant surface ozone measurements from several European receptor sites are available. From the middle of the 1990s onwards a large number of reliable surface ozone measurements from rural and remote sites are available from North America and Europe, while fewer measurements are available from East Asia. Measurements from other parts of the world are still very sparse. Extended surface measurements of carbon monoxide suitable for comparison with CCMs and CTMs go back to the early 1990s, whereas only a few Peroxyacetylnitrate (PAN) datasets exist. The recent publication by Logan et al. (2012) should be used concerning the evolution of ozone in the lowermost free troposphere over Europe since 1978. BM ozonesondes should not be used for tropospheric ozone trend analysis prior to 1998 and it is recommended to use measurements of those ozonesonde stations that have been assessed and homogenised if required. It is also suggested to make use of regular aircraft measurements (GASP, NOXAR, MOZAIC, CARIBIC). However, in this case one needs to carefully consider the spatial/temporal coincidences with the observations. It is also suggested to compare relationships between different features (e.g. O_3/CO, PAN/CO, O_3 at European high mountain sites vs. lower most stratosphere over Europe, etc).

4.4 Recommendations

This document contains an inventory of ground-based measurements, ozonesondes and regular aircraft measurements recommended for comparison between numerical simulations and field measurements. Possible synergies with other projects need to be considered. It might be useful to link the SPARC/IGAC model evaluation activity with AQMEII, which is also coordinating a model evaluation exercise though for regional models. In the second phase of AQMEII the focus will be on the evaluation of fully-integrated chemistry-meteorology models that include feedbacks

between chemistry and climate. The AQMEII models will cover either Europe or North America. In terms of evaluation datasets there will probably be several synergies with CCMI. For example, the harmonised data series of CO, O_3 and NO_2 from the EU-project GEOMON (http://geomon.empa.ch/final_dataset/v2/,.http://www.geomon.eu/science/act2/SciAct2_final_harm_dataset.php) could be attractive for tropospheric CCMI validation. Datasets that currently extend to 2006 could likely be updated rather easily.

Most of the measurements discussed are available from a variety of archives (the addresses of the archives are given in the text, in many cases a data protocol needs to be signed prior to use of the measurements). From a modeller's perspective, it would be more desirable to have one (or possibly a few) archives where the data are stored in a common format and containing only those measurements believed to fulfil the demands of data quality suitable for comparisons with numerical simulations. Several attempts have been made in the past to sample part of the measurements useful for such purposes (e.g. ETHMeg database and others), however, further discussions and a common effort is needed to design such an archive (possibly including in-situ data and satellite measurements as well) in order (i) to fulfil the demand of the modelling community, (ii) not to duplicate efforts and (iii) to plan how the required resources can be mobilised.

Acknowledgement: I would like to thank Fiona Tummon (SPARC Office) for help in editing the manuscript.

References

Brönnimann, S., J. Cain, J. Staehelin, and S.F.G. Farmer, Measurements of total ozone prior to the International Geophysical Year (IGY) 1957, Data and quality, *Quart. J. Roy. Met. Soc*., **129**, 2819-2843, 2003.

Brunner, D., J. Staehelin, and D. Jeker, Large-scale nitrogen oxide plumes in the tropopause region and implications for ozone, *Science*, **282**, 1305-1309, 1998.

Brunner, D., J. Staehelin, D. Jeker, H. Wernli, and U. Schumann, Nitrogen oxides and ozone in the tropopause region of the Northern Hemisphere: Measurements from commercial aircraft in 1995/96 and 1997, *J. Geophys. Res.,* **106**, 27673-27699, 2001.

Brunner, D., J. Staehelin, H. L. Rogers, M. O. Köhler, J. A. Pyle, D. Hauglustaine, L. Jourdain, T. K. Berntsen, M. Gauss, I. S. A. Isaksen, E. Meijer, P. van Velthoven, G. Pitari, E. Mancini, V. Grewe, and R. Sausen, 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, 2003.

Brunner, D., J. Staehelin, H. L. Rogers, M. O. Köhler, J. A. Pyle, D. A. Hauglustaine, L. Jourdain, T. K. Berntsen, M. Gauss, I. S. A. Isaksen, E. Meijer, P. van Velthoven, G. Pitari, E. Mancini, V. Grewe, and R. Sausen, An evaluation of the performance of chemistry transport models, Part 2: detailed comparison with two selected campaigns, *Atmos. Chem. Phys*., **5**, 107-129, 2005.

Chan and R. J. Vet, Baseline levels and trends of ground level ozone, *Atmos. Chem. Phys*., **10**, 8629–8647, 2010.

Cooper, O.R., D. D. Parrish, A. Stohl, M. Trainer, P. Nedelec, V. Thouret, J. P. Cammas, S. J. Oltmans, B. J. Johnson, D. Tarasick, T. Leblanc, I. S. McDermid, D. Jaffe, R. Gao, J. Stith, T. Ryerson, K. Aikin, T. Campos, A. Weinheimer, and M. A. Avery, Increasing springtime ozone mixing ratios in the free troposphere over western North America, *Nature*, **463**, 344-348, doi:10.1038, 2010.

Cui, J., S. Pandey Deolal, M. Sprenger, S. Henne, J. Staehelin, M. Steinbacher, and P. Nedelec, Free Tropospheric Ozone Changes Over Europe as observed at Jungfraujoch (1990-2008), An Analysis based on backward trajectories, *J. Geophys. Res.*, **116**, D10304, doi:10.1029/2010JD015154, 2011.

Derwent, R.G., P.G. Simmonds, A.J. Manning, and T.G. Spain, Trends over a 20 year period from 1987 to 2007 in surface ozone at the atmospheric research station, Mace Head, Ireland, *Atmos. Environ*., **41**, 9091-9098, 2007.

Emmons, L.K., D.A. Hauglustaine, J.-F. Müller, M.A. Carroll, G.P. Brasseur, D. Brunner, J. Staehelin, V. Thouret, and A. Marenco, Data comparison of tropospheric ozone and its precursors from aircraft measurements, *J. Geophys. Res*., **105**, 20'497-20'538, 2000.

Eyring V., I. Cionni, G.E. Bodeker, A.J. Charlton-Perez, D.E. Kinnison, J.F. Scinocca, D.W. Waugh, H. Akiyoshi, S. Bekki, M.P. Chipperfield, M. Dameris, S. Dhomse, S.M. Frith, H. Garny, A. Gettelman, A. Kubin, U. Langematz, E. Mancini, M. Marchand, T. Nakamura, L.D. Oman, S. Pawson, G. Pitari, D.A. Plummer, E. Rozanov, T.G.

Shepherd, K. Shibata, W. Tian, P. Braesicke, S.C. Hardiman, J.F. Lamarque, O. Morgenstern, J.A. Pyle and Y. Yamashita, Multi-model assessment of stratospheric ozone return dates and ozone recovery in CCMVal-2 models, *Atmos. Chem. Phys*., **10**, 9451-9472, 2010.

Fiore., A.M., F. J. Dentener, O. Wild, C. Cuvelier, M. G. Schultz, P. Hess, C. Textor, M. Schulz, R. M. Doherty, L. W. Horowitz, I. A. MacKenzie, M. G. Sanderson, D. T. Shindell, D. S. Stevenson, S. Szopa, R. Van Dingenen, G. Zeng, C. Atherton, D. Bergmann, I. Bey, G. Carmichael, W. J. Collins, B. N. Duncan, G. Faluvegi, G. Folberth, M. Gauss, S. Gong, D. Hauglustaine, T. Holloway, I. S. A. Isaksen, D. J. Jacob, J. E. Jonson, J. W. Kaminski, T. J. Keating, A. Lupu, E. Marmer, V. Montanaro, R. J. Park, G. Pitari, K. J. Pringle, J. A. Pyle, S. Schroeder, M. G. Vivanco, P. Wind, G. Wojcik, S. Wu, and A. Zuber, Multimodel estimates of intercontinental source-receptor relationships for ozone pollution, *J. Geophys. Res*, **114**, D04301 doi:10.1029/20078D010816, 2009.

Goetz, F. W. P., A. R. Meetham, and G. M. B. Dobson, The vertical distribution of ozone in the atmosphere, *Proc. R. Soc. London*, *Ser. A*, **145**, 416-446, 1934.

Gilge, S., C. Plass-Duelmer, W. Fricke, A. Kaiser, L. Ries, B. Buchmann, and M. Steinbacher, Ozone, carbon monoxide and nitrogen oxides time series at four alpine GAW mountain stations in central Europe, *Atmos. Chem. Phys.*, **10**, 12295–12316, 2010.

Harris, N., J. Staehelin, and R. Stolarski, The New Initiative on Past Changes in the Vertical Distribution of Ozone, SPARC Newsletter, **37**, 23-26, 2011.

Helmig, D., J. Bottenheim, I. E. Galbally, A. Lewis, M. J. T. Milton, S. Penkett, C. Plass-Duelmer, S. Reimann, P. Tans, and S. Thiel, Volatile Organic Compounds in the Global Atmosphere, *Eos Trans. AGU*, *90*(52), 513, doi:10.1029/2009EO520001, 2009.

HTAP, 2010, Hemispheric transport of air pollution, Part A: Ozone and Particulate matter, air pollution studies, No. 17**,** Edited by Frank Dentener, Terry Keating, and Hajime Akimoto, prepared by the Task Force on Hemispheric Transport of Air Pollution acting within the framework of the Convention on Long-range Transboundary Air Pollution, Economic Commission for Europe, United Nations, New York and Geneva, 2010.

Kerr, J. B., I.A. Asbridge, and W.F.J. Evans, Intercomparison of total ozone measured by the Brewer and Dobson Spectrophotometers at Toronto, *J. Geophys. Res*., *93(D9)*, **11**, 129– 11,140,1988.

Komhyr, W.D., Operations handbook: Ozone observations with a Dobson spectrophotometer, *Global Ozone Res. and Monit. Proj.*, Rep. 6, WMO, Geneva, Switzerland, 1980.

Komhyr, W.D., R.D. Grass, and R.K. Leonard, Dobson Spectrophotometer 83: A standard for total ozone measurements, 1962-1987, *J. Geophys. Res*., **94**, 9847-9861, 1989.

Logan, J.A., I.A. Megretskaia, A.J. Miller, G.C. Tia, D. Choi, L. Zhang, R.S. Stolarski, G.J. Labow, S.M. Hollandsworth, G.E. Bodecker, H. Claude, D. DeMuer, J.B. Kerr, D.W. Tarasick, S.J. Oltmans, B. Johnson, F. Schmidlin, J. Staehelin, P. Viatte, and O. Uchino: Trends in the vertical distribution of ozone: A comparison of two analyses of ozonesonde data*, J. Geophys. Res*., **104**, 26'373-26'399, 1999.

Logan, J.A., J. Staehelin, I. A. Megretskaia, J.-P. Cammas, V. Thouret, H. Claude, H. De Backer, M. Steinbacher, H. E. Scheel, R. Stübi, M. Fröhlich, and R. Derwent, 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.

Novelli, P.C., K. A. Masari, and P. M. Lang, Distributions and recent changes of carbon monoxide in the lower troposphere, *J. Geophys. Res*., **103**, 19,015-19,033, 1998.

Ordóñez, C., D. Brunner, J. Staehelin, P. Hadjinicolaou, J.A. Pyle, M. Jonas, H. Wernli, and A. S.H. Prévôt, Strong influence of lowermost stratospheric ozone on lower free tropospheric ozone changes over Europe, *Geophys. Res. Lett*. , **34**, L07805, doi:10.1029/2006GL029113, 2007.

Oltmans, S.J., A.S. Lefohn, J.M. Harris, I. Galbally, H.E. Scheele, G.E. Bodeker, E.Brunke, H. Claude, D. Tarasick, B.J. Johnson, P. Simmonds, D. Shadwick, K. Anlauf, K. Hayden, F. Schmidlin, T. Fujimoto, K. Akagi, C. Meyer, S. Nichol, J. Davies, A. Redondas, and E. Cuevas, Long-term changes in tropospheric ozone, *Atmos. Environ*., **40**, 3156–3173, 2006.

Pandey Deolal, S., D. Brunner, M. Steinbacher, U. Weers and J. Staehelin, Long-term in situ measurements of NO_x and NO_y at Jungfraujoch 1998-2009: time series analysis and evaluation, *Atmos. Chem. Phys*., **12**, 2551–2566, 2012.

Pandey Deolal, S., J. Staehelin, D. Brunner, J. Cui, M. Steinbacher, C. Zellweger, S. Henne and M. Vollmer, Transport of PAN and NO_y from different source regions to the Swiss high alpine site Jungfraujoch, *Atmos. Environ.*, **64**, 103-115, 2013.

Parrish, D. D., et al., Changes in the photochemical environment of the temperate North Pacific troposphere in response to increased Asian emissions, *J. Geophys. Res*., **109**, D23S18, doi:10.1029/2004JD004978, 2004.

Parrish, D.D., D. B. Millet, and A. H. Goldstein, increasing ozone in marine boundary layer inflow at the west coasts of North America and Europe, *Atmos. Chem. Phys.*, **9**, 1303–1323, 2009.

Parrish , D.D., K.S. Law , J. Staehelin , R. Derwent , O. Cooper , H. Tanimoto , A. Volz Thomas , S. Gilge , H.-E. Scheel , M. Steinbacher , and E. Chan, Long-term changes in lower tropospheric baseline ozone concentrations at northern mid-latitudes, *Atmos. Chem. Phys.*, in press, 2012.

Pausata, F.S.R., L. Pozzoli, E. Vignati, and F.J. Dentener, North Atlantic Oscillation and tropospheric ozone variability in Europe: model analysis and measurement intercomparison, *Atmos. Chem. Phys.*, **12**, 6357-6376, 2012.

Pommereau, J.-P., and F. Goutail, O_3 and NO_2 ground-based measurements by visible spectrometry during Arctic winter and spring 1988, *Geophys. Res. Lett*., **15**, 891-894, 1988.

Schnadt Poberaj, C., J. Staehelin, D. Brunner, V. Thouret, and V. Mohnen, A UT/LS ozone climatology of the nineteen seventies deduced from the GASP aircraft measurement program, *Atmos. Chem. Phys.*, **7**, 5917-5936, 2007.

Schnadt Poberaj, C., J. Staehelin, D. Brunner, V. Thouret, H. De Backer, and R. Stübi, 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, 2009.

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

Smit, H.G.J and Panel for the Assessment of Standard Operating Procedures for Ozonesondes (ASOPOS), Quality Assurance and Quality Control for Ozonesonde Measurements in GAW, GAW Report No. 201, World Meteorological Organization (WMO), Geneva, 2011.

Staehelin, J., J. Thudium, R. Bühler, A. Volz-Thomas, and W. Graber, Surface ozone trends at Arosa (Switzerland), *Atmos. Environ.*, **28**, 75-87, 1994.

Staehelin, J., A. Renaud, J. Bader, R. McPeters, P. Viatte, B. Högger, V. Bugnion, M. Giroud, and H. Schill, Total ozone series of Arosa (Switzerland). Homogenization and data comparison, *J. Geophys. Res*., **103**, 5827-5841, 1998.

Tanimoto, H., Increase in springtime tropospheric ozone at a mountainous site in Japan for the period 1998-2008, *Atmos. Environ*., **43,** 1358-1363, 2009.

Tanimoto, H., T. Ohara, and I. Uno, Asian anthropogenic emissions and decadal trends in springtime tropospheric ozone over Japan: 1998–2007, *Geophys. Res. Lett*., **36**, L23802, doi:10.1029/2009GL041382, 2009.

Tarasova, O.A., I.A. Senik, M.G. Sosonkin, J. Cui, J. Staehelin, and A.S.H. Prévôt, Surface ozone at the Caucasian site Kislovodsk High Mountain Station and the Swiss Alpine site Jungfaujoch: Data analysis and trends (1990-2006), *Atmos. Chem. Phys.,* **9**, 4157-4175, 2009.

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

Tilmes, S., J.-F. Lamarque, L. K. Emmons, A. Conley, M. G. Schultz, M. Saunois, V. Thouret, A. M. Thompson, S. J. Oltmans, B. Johnson, and D. Tarasick, Technical Note: Ozonesonde climatology between 1995 and 2009: description, evaluation and applications, *Atmos. Chem. Phys.*, **12**, 7475-7497, 2012.

Thompson, A., et al., SHADOZ (Southern Hemisphere Additional Ozonesondes) Ozone climatology. 4. tropospheric and lower stratospheric profiles (2005-2009) with comparisons to OMI total ozone, subm. 2011.

Vogler, C., S. Brönnimann, J. Staehelin, and R.E.M. Griffin, Dobson total ozone series of Oxford: Re-evaluation and applications, *J. Geophys. Res.*, **112**, D20116, doi:10.1029/2007JD008894, 2007.

Wilson, R.,C., Z. L. Fleming, P.S. Monks, R. Clain, S. Henne, I.B. Konovalov, S. Szopa, and L. Menut, Have primary emission reduction measures reduced ozone across Europe ? An analysis of European rural background ozone trends 1996-2005, *Atmos. Chem. Phys*., **12**, 437-454, 2012.

Zanis, P., E. Schuepbach, H. Scheel, M. Baudenbacher, and B. Buchmann, Inhomogeneity and trends in the surface ozone record (1988-1996) at Jungfraujoch in the Swiss Alps, *Atmos. Environ*., **33**, 3777–3786, 1999.

Zbinden, R.M., J.-P. Cammas, V. Thouret, P. Nedelec, F. Karcher, and P. Simon, Mid-latitude tropospheric ozone columns from the MOZAIC program: climatology and interannual variability, *Atmos. Chem. Phys*., **6**, 1053-1073, 2006.