Chapter 4: Climatology evaluations

Within the SPARC Data Initiative, satellite trace gas measurements are compared as monthly zonal mean time series following a 'climatological' approach to data validation, in contrast to the more common approach of using coincident profile measurements. The climatological validation method has the advantages that it is consistent between all instruments, avoids sensitivity to arbitrary coincidence criteria, and generally produces larger sample sizes, which should in theory minimise the random sampling error. At the same time, climatological means may be biased due to non-uniformity of sampling as described in *Section 3.2.1* Another important aspect of our approach is that trace gas climatologies are compared without any modification to account for different resolutions in altitude due to application of the averaging kernels. We consider our simplified approach as justified, because in most cases the vertical resolutions of the limb sounders are quite similar, and the degree to which the *a priori* information influences the retrieved profile is limited. Furthermore, highly structured and transient features that may not be resolved by some instruments will most likely average out in the monthly climatologies. The SPARC Data Initiative evaluations are based on the use of the multi-instrument mean (MIM) as a reference. This choice is not based on the assumption that the MIM is the best climatology available, but is motivated by the need for a reference that does not favour a certain instrument. Evaluations are carried out for time periods that allow for maximum overlap between different instruments in order to yield relatively robust conclusions on instrument performance. All evaluations in the following chapter are based on the climatological validation approach, and the above advantages and disadvantages will be discussed where appropriate.

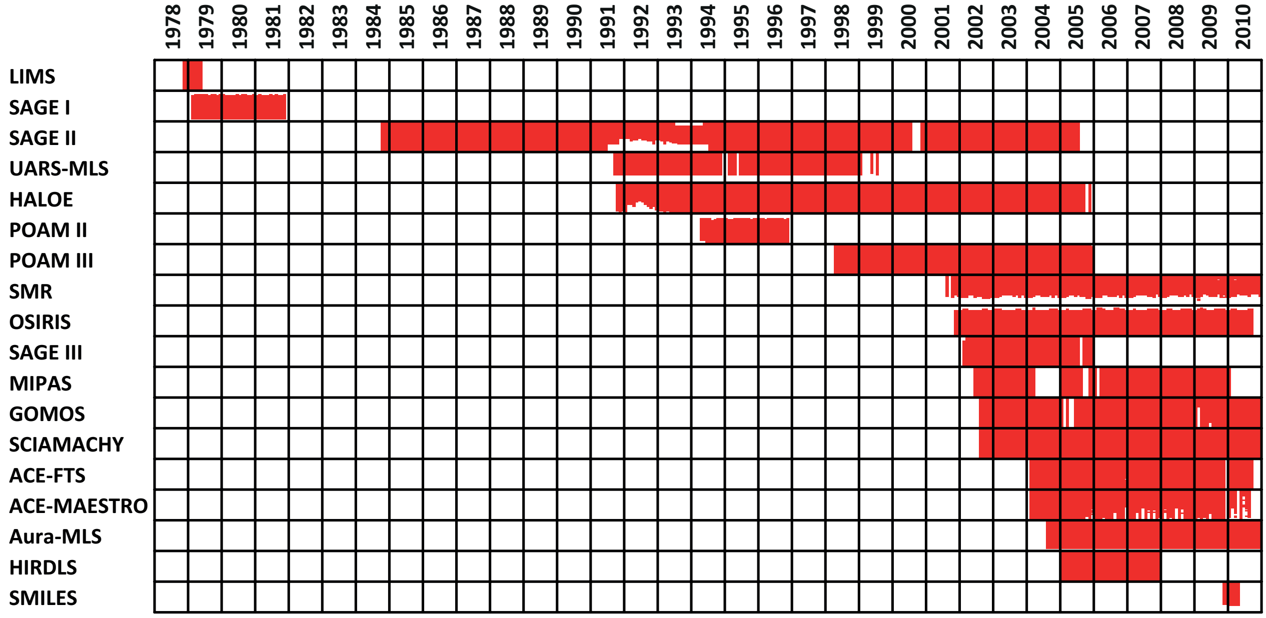
4.1 Ozone – O_3

Ozone is one of the most important trace species in the atmosphere due to its absorption of biologically harmful ultraviolet radiation and its role in determining the temperature structure of the atmosphere. Most ozone (about 90%) is found in the stratosphere, and the region of highest ozone concentration between 20-25 km is commonly known as the ozone layer. The recent depletion of the ozone layer as a result of anthropogenic emissions of halogenated species is expected to decrease and reverse [*Austin and Butchart*, 2003; *SPARC*, 2010; *WMO*, 2014] due to the phase-out of ozone-depleting substances (*e.g.*, CFCs, see *Sections 4.5* and *4.6*) specified by the Montreal Protocol and its subsequent amendments. Detection and attribution of the expected ozone recovery in a future changing climate [*e.g.*, *Newman et al.*, 2006; *Waugh et al.*, 2009] with increasing greenhouse gases and a modified residual circulation will require a comprehensive understanding of short- and long-term ozone changes, and their altitude, latitude and seasonal dependence. Such knowledge can only be derived from high quality, vertically resolved, global, long-term ozone datasets. A large number of satellite instruments have been measuring stratospheric ozone over the past three decades and the resulting datasets will be evaluated in the following section. The spread between the various climatological datasets will be presented and where possible related to instrument characteristics and sampling issues. Additionally, the physical consistency of the datasets will be tested. The systematic comparison presented here, as summarised in *Tegtmeier et al.* [2013], has served as input for other initiatives, such as the SPARC Initiative on Past changes in the Vertical Distribution of Ozone (SI2N), NASA MEaSUREs Global OZone Chemistry And Related trace gas Data records for the Stratosphere (GOZCARDS) project, or the European Ozone Climate Change Initiative (ESA O3-CCI), which aimed to analyse various sources as homogeneous data records suitable for trend studies.

4.1.1 Availability of O_3 measurements

The SAGE II ozone dataset is considered to be the most reliable long-term satellite data source for the detection and quantification of ozone changes in the lower stratosphere between 1984 and 2005. HALOE and UARS-MLS measurements also cover the 1990s, with HALOE providing the second longest record, from 1991 to 2005. Many other satellite instruments have been measuring the vertical ozone distribution since 2000. A thorough assessment of how well the new measurements agree with each other and with older measurements is critical in order to create a merged data record for the investigation of ozone trends. Although the SBUV (Solar Backscatter UltraViolet) and SBUV/2 instruments provide a long-term ozone record with excellent coverage and density, the data are not included here due to their limitations in vertical resolution. The SBUV algorithm retrieves the ozone content for relatively thick (6–8 km) layers between about 30–50 km, and provides only very limited profile information outside this region [*Bhartia et al.*, 2004]. As a result, the amplitude of ozone fluctuations with a fine vertical structure, such as the quasi-biennial oscillation (QBO) signal, are damped in the SBUV(/2) dataset [*McLinden el al.*, 2009]. Independent ozone profile measurements from selected sites are available from ground

***Table 4.1.1:** **Available ozone measurement records between 1978 and 2010 from limb-sounding satellite instruments participating in the SPARC Data Initiative.** The red filling of the grid boxes indicates the temporal (January to December) and vertical (300 to 0.1 hPa) coverage of the respective instruments.*



based ozone monitoring instruments (*e.g.,* ozonesonde, Umkehr, LIDAR and microwave), which are often used for satellite validation and in other investigations. Knowledge derived from such comparisons with independent measurements will be used where available in order to explain identified differences between the satellite datasets.

Tables 4.1.1 and **4.1.2** compile information on the availability of ozone measurements, including data version, time period, vertical range, vertical resolution, and references relevant for the data products used in this report.

4.1.2 O_3 evaluations: Zonal annual mean cross sections, vertical and meridional profiles

Annual zonal mean cross sections are analysed to investigate mean differences between the various datasets. The annual means have been calculated over multiple years as indicated in the section headings. The time periods have been chosen so that a maximum number of instruments can be compared in each case. Differences between individual instruments and the multi-instrument mean (MIM, see *Section 3.3* for definition) are presented. Note that the choice of the MIM is not based on the assumption that the MIM is the best climatology available, but is motivated by the need for a reference that does not favour any particular instrument. For instruments without complete yearly coverage at all latitude bands, the differences can be caused not only by the instrumental bias with respect to the MIM, but also by the fact that not all months of the year are available for the calculation of the annual mean. For such cases, the analyses will refer to monthly zonal mean cross sections, as shown in *Appendix A4.1*. Additionally, monthly mean vertical and meridional profiles are presented to analyse the mean differences in more detail. Profiles are presented together with the standard error of the mean (SEM, see *Section 3.2.3* for definition), an estimate of the statistical uncertainty in the mean value.

In the mesosphere, day- and night-time ozone differences exist due to photodissociation processes within the odd oxygen families [*e.g.*, *Brasseur and Solomon*, 1984]. The resulting diurnal ozone variations are of the order of 10% in the upper stratosphere, 20% at the lowest mesospheric levels (~ 1 hPa) and grow with increasing altitude up to more than 100% for upper mesospheric levels [*e.g.*, *Wang et al.*, 1996; *Schneider et al.*, 2005]. **Figure 4.1.1** shows examples of the diurnal ozone cycle as a function of local solar time (LST) for three different pressure levels as derived with a chemical box model [*McLinden et al.*, 2010]. Depending on the instruments' sampling pattern, the diurnal cycle in ozone may therefore add an additional sampling bias in the LM.

SAGE II, UARS-MLS, HALOE and POAM II (1994-1996)

The annual zonal mean ozone climatologies for 1994-1996 for SAGE II, UARS-MLS, HALOE, POAM II, and their MIM are shown in **Figure 4.1.2**. The maximum ozone mixing ratio is found in the tropics at about 10 hPa, well above the ozone layer at about 50 hPa. Differences of the individual datasets with respect to the MIM are shown in **Figure 4.1.3**. The instruments show excellent agreement within the tropical and mid-latitude MS/US, with differences around ±2.5%; UARS-MLS exhibits positive differences with respect to the MIM, HALOE shows negative differences, while SAGE II shows differences of mixed sign. POAM II, which is restricted to higher latitudes, shows larger annual differences (of up to -20%). In general, relative differences for all instruments are larger in the UTLS, and LM, as well as in the polar regions at all altitude levels (mostly limited to ±20%). In the LM (above 1 hPa), differences between

Table 4.1.2: Data version, time period, vertical range, vertical resolution, references and other comments for ozone datasets participating in the SPARC Data Initiative.

Instrument	Time period	Verical range	Vertical resolution	References	Additional comments
LIMS V6.0	Nov 78 – May 79	cloud top – 0.01 hPa (10 – 80 km)	3.7 km	*Remsberg et al.*, 2007	
SAGE I V5.9	Feb 79 – Nov 81	cloud top – 55 km	1 km	*McCormick et al.*, 1989 *Wang et al.*, 1996	With altitude corrections based on *Wang et. al.* [1996]
SAGE II V6.2	Oct 84 – Aug 05	cloud top – 70 km	0.5 – 1 km	*Chu et al.*, 1989 *Wang et al.*, 2002	
UARS-MLS V5	Oct 91 – Oct 99	100 – 0.02 hPa (16 – 75 km)	3.5 – 5 km (LS-US) 5 – 8 km (LM)	*Livesey et al.*, 2003	Not as good for trends after Jun 1997 (no more MLS retrievals of T), and sparser data.
HALOE V19	Oct 91 – Nov 05	250 – 0.002 hPa (10 – 90 km)	2.5 km	*Grooß and Russell,* 2005	
POAM II V6.0	Oct 93 – Nov 96	15 – 50 km	1 km	*Lumpe et al.*, 1997 *Rusch et al.*, 1997	
POAM III V4.0	Apr 98 – Dec 05	5 - 60 km	1.0 km	*Lumpe et al.*, 2002 *Randall et al.*, 2003	
SMR V2.1	Jul 01 –	Antarctic: 100 – 0.1 hPa 16 – 65 km Tropics: 75 – 0.1 hPa 18 – 65 km	2.5 – 3.5 km	*Urban et al.*, 2005a, 2006	O_3 is measured at the 501.8 GHz band. Several other O_3 products exist which are not used here.
OSIRIS V5.07	Oct 01 –	10 – 60 km	2 km	*Degenstein et al.*, 2009	
SAGE III V4.0	Feb 02 – Dec 05	cloud top – 100 km	0.5 – 1 km	*Wang et al.*, 2006	Separate retrievals for mesospheric ozone exist (not used here)
MIPAS MIPAS(1) V9 MIPAS(2) V220	 Mar 02 – Mar 04 Jan 05 – Apr 12	 cloud top – 70km cloud top – 70 km	 3.5 – 5.0 km 2.7 – 3.5 km	 *Steck et al.*, 2007 *von Clarmann et al.*, 2009a	
GOMOS V5.0	Aug 02 – Apr 12	15 – 100 km	2 – 3 km	*Kyrölä et al.*, 2010a	
SCIAMACHY V2.5	Aug 02 – Apr 12	10 – 60 km	3 – 5 km	*Mieruch et al.*, 2012	
ACE-FTS V2.2	Mar 04 –	5 – 95 km	3 – 4 km	*Dupuy et al.*, 2009	
ACE-MAESTRO V2.1 (VIS)	Mar 04 –	5 – 60 km	2 km	*Dupuy et al.*, 2009	UV ozone product exists (not used here)
Aura-MLS V2.2	Aug 04 –	215 – 0.022 hPa (12 – 75 km)	3 km 4 km above 0.2 hPa	*Froidevaux et al.*, 2008a *Jiang et al.*, 2007 *Livesey et al.*, 2008	
HIRDLS V6.0	Feb 05 – Dec 07	420 – 0.1 hPa (10 – 65 km)	1 km	*Nardi et al.*, 2008	HIRDLS data exist until mid March 2008
SMILES V2.1.5	Oct 09 – April 10	100 – 0.0005 hPa (16 – 96 km)	3 – 5 km	*Baron et al,* 2011	

instruments that measure at different LSTs cannot be easily evaluated since they can be exaggerated or obscured by the effects of the diurnal ozone cycle (see **Figure 4.1.1**).

Monthly mean vertical ozone profiles at the equator, northern mid-latitudes, and northern and southern high latitudes are shown in **Figure 4.1.4** together with their differences from the MIM. Tropical profiles in October and NH mid-latitude profiles in April confirm that between 30 and 1 hPa all available instruments show only small differences (±2.5%). In the tropics, SAGE II and HALOE agree even within their relatively small SEM error bars. At higher southern latitudes in spring, the differences between the datasets are larger, reaching values of ±10%.

The comparison of monthly mean zonal mean data is complicated by the different sampling patterns of the instruments, especially at high latitudes where intra-monthly

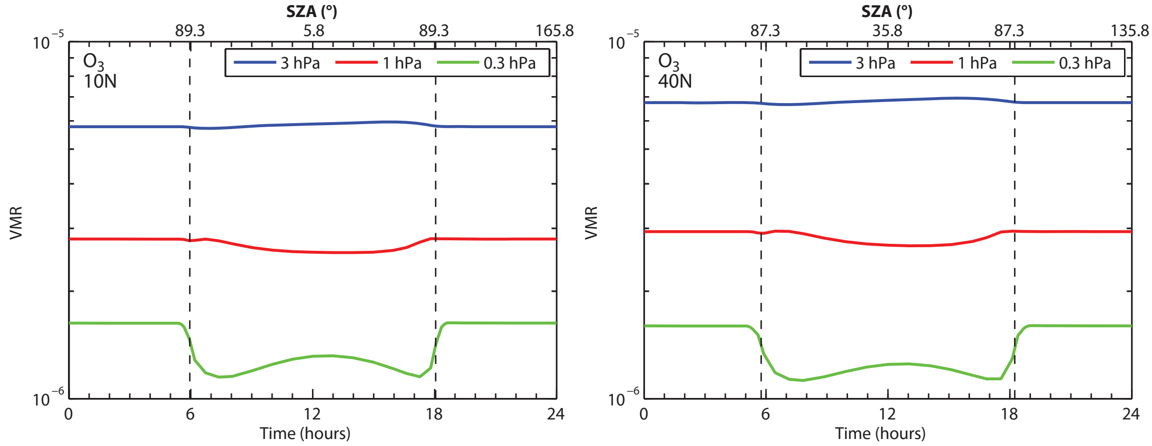


*Figure 4.1.1: **Diurnal ozone cycle.** Ozone variations as function of local solar time are shown at 10°N and 40°N at 3, 1 and 0.3 hPa for March 15.*

and interannual natural variability is strongest. First, when looking at SAGE II, UARS-MLS, HALOE, and POAM II 1994-1996 multi-annual mean values in a given month (*e.g.*, in **Figure 4.1.4**), one needs to keep in mind that the four instruments do not provide data at high latitudes for all years. Additionally, the instruments may measure in a particular latitude band during different times of the month (see *Section 3.2.1* for a detailed discussion of this effect). The comparison of October monthly means at 65°S – 70°S (upper left panel of **Figure 4.1.4**) shows differences of up to ±10% between HALOE and SAGE II with a change of sign at 10 hPa, which can be interpreted as the effect of a small vertical offset between the zonal monthly mean profiles of the two instruments. However, at this latitude band SAGE II provides October values for 1996 only, with measurements mostly at the end of the month, while HALOE provides October values for all three years with measurements at the beginning of the month. As a result, the displayed monthly mean differences may not be representative of instrument biases, and can change from month to month (see **Figures A4.1.1 – A4.1.8** in *Appendix A4*). Looking at annual mean differences adds another level of complication, due to the fact that some instruments do not sample all latitude bands for each month of the year. The vertically oscillating differences observed for the annual mean comparisons at high latitudes (**Figure 4.1.3**) are not present to the same degree for the individual months (see **Figures A4.1.1 – A4.1.8** in *Appendix A4*). As a consequence of the above mentioned sampling effects, differences between climatological datasets at high latitudes can be large even if the actual inter-instrument differences for individual measurements are not.

Figure 4.1.5 shows meridional monthly mean profiles at 1, 10, 50 and 70 hPa. The relative differences at 10 hPa are smallest and only exceed the ±5% threshold poleward of 70°. At 1 hPa the differences are slightly larger than at 10 hPa and show maxima at about 50°N and 50°S. This peak is related to the ozone maximum at mid-latitudes, which is slightly displaced by SAGE II compared to the

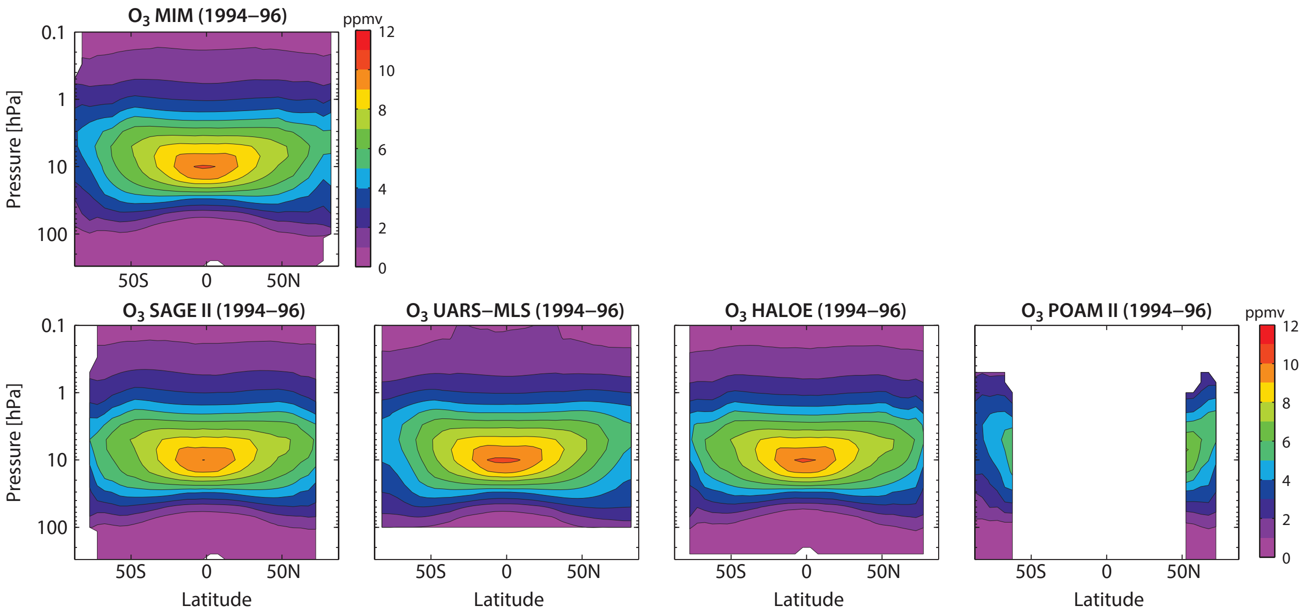


*Figure 4.1.2: **Cross sections of annual zonal mean ozone for 1994-1996.** Annual zonal mean ozone cross sections are shown for the MIM in the upper panel and for SAGE II, UARS-MLS, HALOE, and POAM II in the lower panels.*

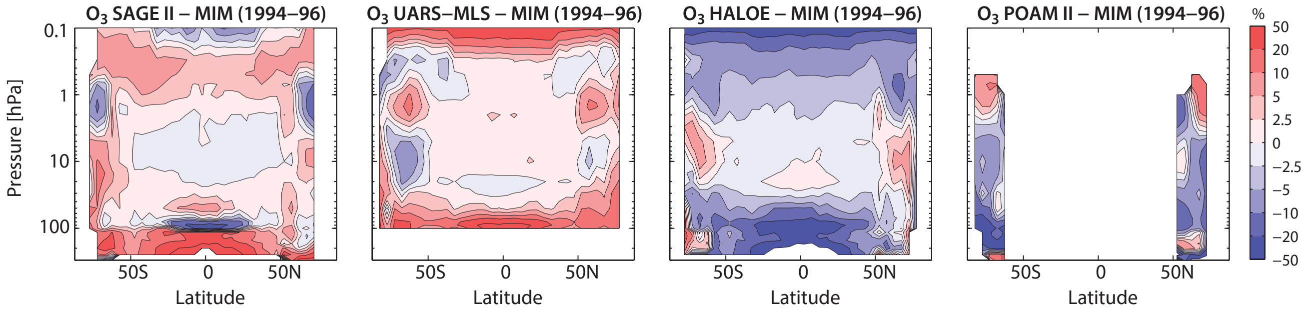


Figure 4.1.3: Cross sections of annual zonal mean ozone differences for 1994-1996. *Annual zonal mean ozone differences between the individual instruments and the MIM are shown.*

other instruments. At the two levels in the LS, the relative differences are largest in the tropics (up to ±15% at 50 hPa and ±30% at 70 hPa), which is likely related to the steep vertical ozone gradient in this region that is resolved in different ways by the various intruments. Also, there are instrumental limitations in this altitude region, *e.g.*, resulting from cloud interference and high extinction, which can vary depending on the spectral regions used for the measurement, and on the measurement mode. While for the upper levels of the LS, HALOE shows a negative difference compared to the other two intruments, for the lower levels HALOE and SAGE II both show negative deviations compared to UARS-MLS. For pressure levels larger than 100 hPa, UARS-MLS ozone values are too large and their use is not recommended [*Livesey et al.*, 2003]. These pressure levels have not been included in the SPARC Data Initiative climatology, as seen in **Figure 4.1.3**. Some high biases (of the order of 10%) were also reported for UARS-MLS at 100 hPa, *versus* sondes and (mainly in the tropics) *versus* SAGE II. Note that 70 hPa UARS-MLS values can be affected by interpolation from the biased high 100 hPa values, which may explain some differences seen in the right panel of **Figure 4.1.5**.

SAGE II, HALOE, POAM III, SMR, OSIRIS, SAGE III, MIPAS(1), GOMOS, SCIAMACHY (2003)

For 2003, a maximum number of instruments overlap including HALOE and SAGE II (which provide the two longest time series), and the newer instruments that measure from 2001 onwards. SMR provides a second ozone product measured at 488.9 GHz, which has very similar characteristics compared to the main SMR ozone product measured at 501.8 GHz, and is therefore not shown in the following evaluations. Note also that an evaluation of 2003-2004 climatologies leads to very similar results as the evaluation of the 2003 climatologies, however, the 2003 climatologies are presented here since MIPAS(1) data are not available for most of 2004. **Figure 4.1.6** shows the annual zonal mean ozone climatologies for all measurements available in 2003. Their differences with respect to the MIM are displayed in **Figure 4.1.7**.

The smallest relative differences are found in the tropical and mid-latitude MS/US. In this region, the comparison of SAGE II, OSIRIS, MIPAS(1), and GOMOS to the MIM

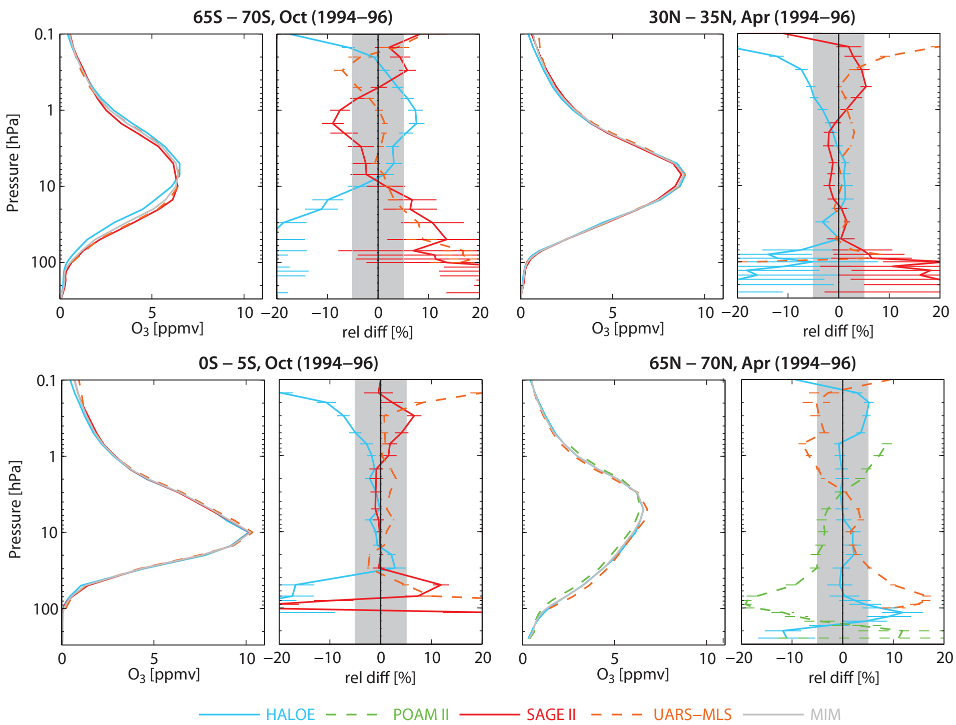


Figure 4.1.4: Vertical Profiles of monthly zonal mean ozone for 1994-1996. *Zonal mean ozone profiles for 65°S-70°S and 0°S-5°S for October in the left column and for 30°N-35°N and 65°N-70°N for April in the right column. Differences between the individual instruments and the MIM profiles are shown on the right side of the panel. Error bars indicate the uncertainties in each climatological mean based on the SEM. The grey shaded area indicates where relative differences are smaller than ±5%.*

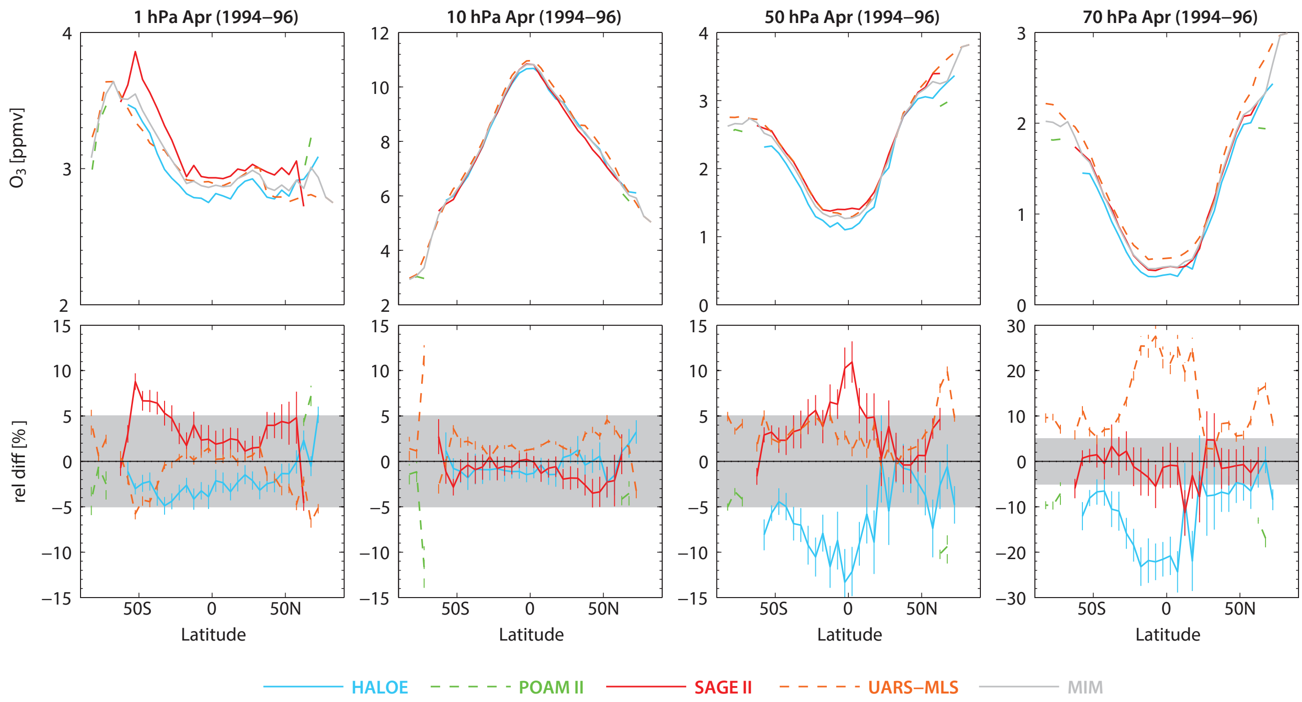


Figure 4.1.5: Meridional profiles of monthly zonal mean ozone for 1994-1996. *Meridional zonal mean ozone profiles at 1, 10, 50 and 70 hPa for April 1994-1996 are shown in the upper row. Differences between the individual instruments and the MIM profiles are shown in the lower row. Error bars indicate the uncertainties in each climatological mean based on the SEM. The grey shaded area indicates where relative differences are smaller than ±5%.*

yields very good agreement, with differences of up to ±5%. HALOE and SMR show good agreement with the other instruments, with negative differences of up to -10%, while SCIAMACHY agrees reasonably well with positive differences of up to +20% covering the MS, US and LM at southern latitudes and the US at northern latitudes. This behaviour is related to the fact that before February 2006 SCIAMACHY retrievals above about 3 hPa suffer from insufficient vertical resolution and limited vertical coverage of the ECMWF temperature data. (Pressure and temperature information from the ECMWF operational dataset are used in the SCIAMACHY retrieval and to convert the originally retrieved number densities into volume mixing ratios.) The relative differences between OSIRIS and the MIM vary with latitude, which is most likely caused by sampling biases introduced by non-uniform sampling over the course of a month or year. In the LS, differences are larger compared to the regions above, however most instruments still agree reasonably well, with differences of up to ±20%. The exception is GOMOS, which shows considerable disagreement of around +50%. Above 70 hPa, GOMOS agrees well to reasonably well with the MIM, but differences to the MIM increase quickly below this pressure level. In the LM, differences increase slowly with increasing altitude, since in this region ozone abundance is decreasing and importance of the ozone diurnal cycle is increasing. The latter effect impedes a direct comparison between instruments measuring at different LSTs. For SAGE III mesospheric ozone, a separate retrieval is available, however, the climatologies evaluated here are based on an algorithm designed for stratospheric and upper tropospheric regions and therefore mesospheric ozone from SAGE III provided with the SPARC Data Initiative should be used with care.

In general, larger differences with respect to the MIM can be observed at higher latitudes compared to tropics and mid-latitudes, in particular for SAGE II, HALOE, OSIRIS (SH), and GOMOS. These differences are partially caused by the effects of non-uniform sampling at high latitudes. The annual mean climatologies from instruments without complete yearly coverage at high latitudes will be biased towards months when measurements are available. This sampling bias will also affect the MIM, and in turn the differences of the instruments with regular sampling patterns to the MIM. The effect of this sampling bias is especially strong at the SH high latitudes. However, large differences at high latitudes observed for some instruments (SAGE II, HALOE, OSIRIS (SH), and GOMOS) are also present in some of the monthly mean comparisons (see **Figures A4.1.9** to **A4.1.16** in *Appendix A4*) and so are not exclusively introduced by the annual averaging. POAM III and SAGE III have limited coverage over the whole year, and provide only data at higher latitudes, where they show differences comparable to those of the other instruments.

Monthly mean ozone profiles at the equator and at high latitudes, together with their differences from the MIM are shown in **Figure 4.1.8**. In the tropics between 50 hPa and 1 hPa, nearly all instruments agree within ±5%, with only SCIAMACHY displaying larger positive differences from the MIM of up to 20%, clearly overestimating the ozone mixing ratio peak at 10 hPa. As already noted for the annual mean comparisons, GOMOS shows some considerable

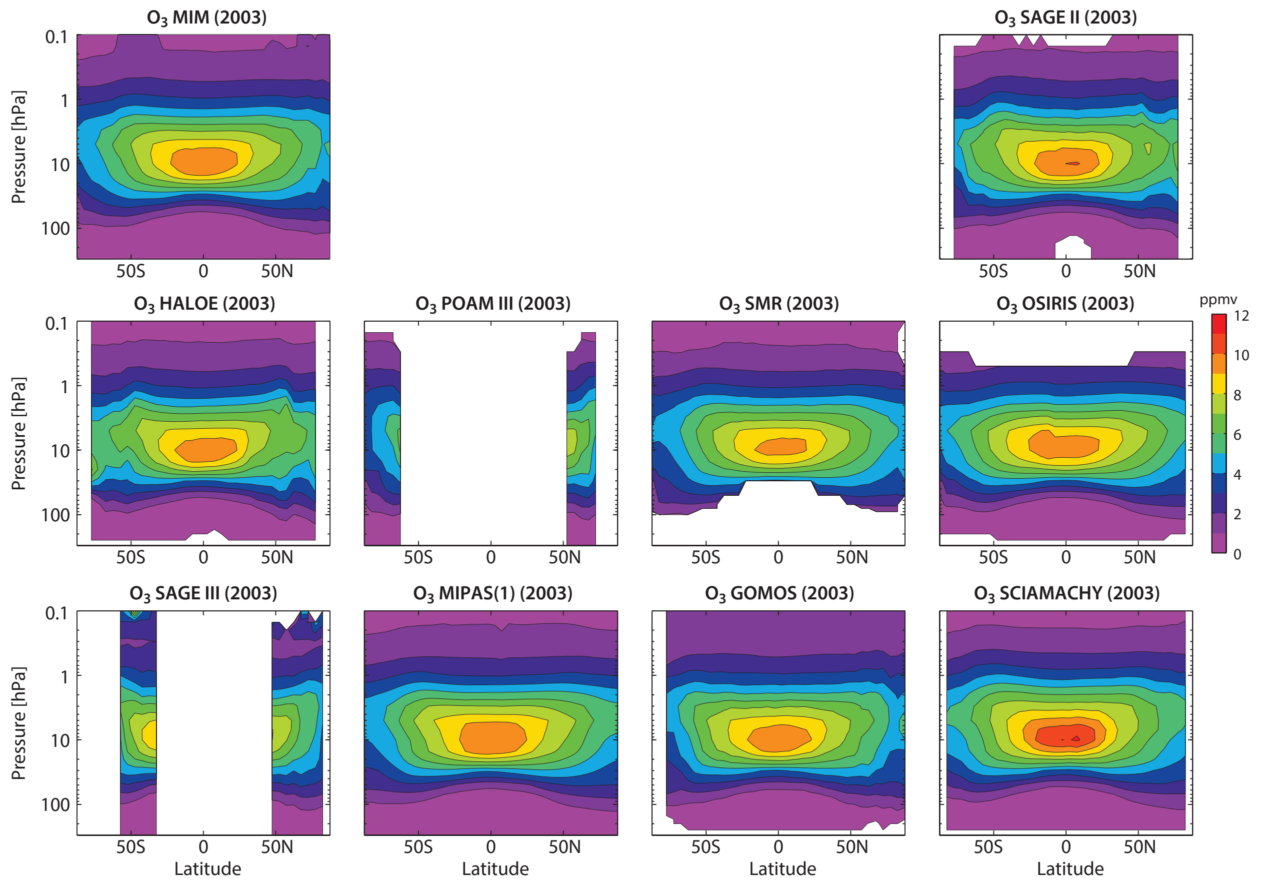


Figure 4.1.6:* *Cross sections of annual zonal mean ozone for 2003. *Annual zonal mean ozone cross sections for 2003 are shown for the MIM, SAGE II, HALOE, POAM III, SMR, OSIRIS, SAGE III, MIPAS(1), GOMOS, and SCIAMACHY.*

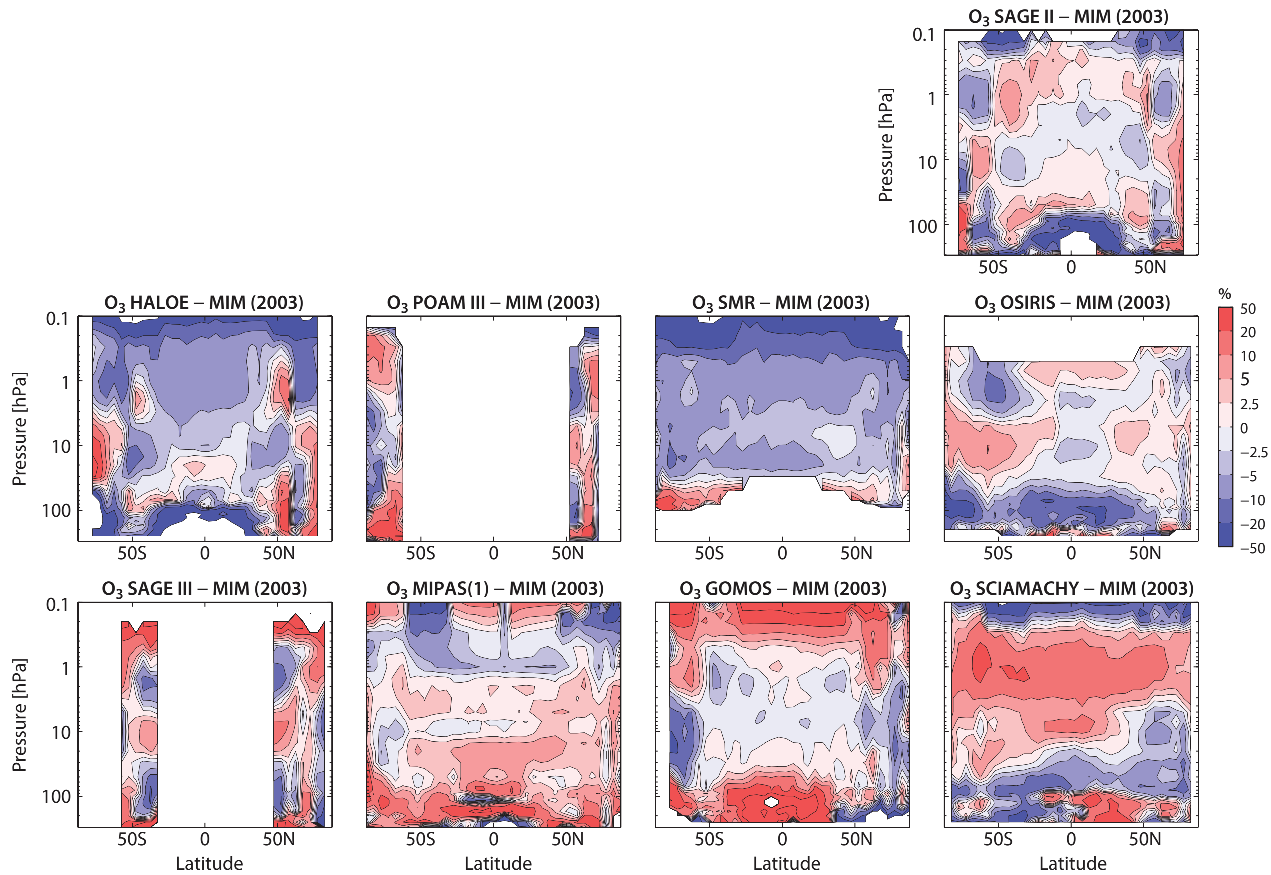


Figure 4.1.7:* *Cross sections of annual zonal mean ozone differences for 2003. *Annual zonal mean ozone differences for 2003 between the individual datasets and the MIM are shown.*

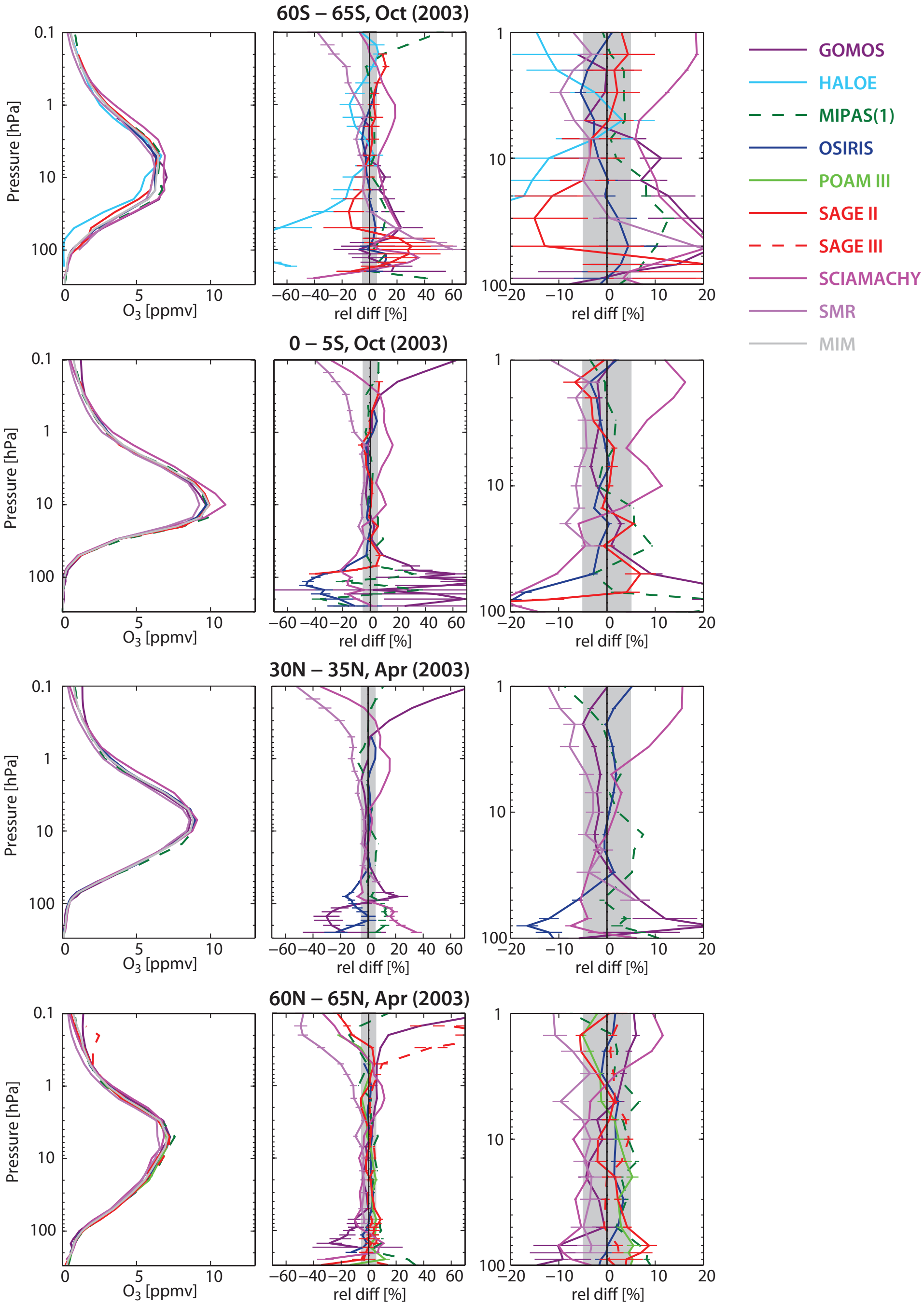


__Figure 4.1.8: Vertical profiles of monthly zonal mean ozone for 2003.__ Monthly zonal mean ozone profiles for 60°S-65°S and 0°S-5°S for October 2003 and for 30°N-35°N and 60°N-65°N for April 2003 are shown in the leftmost column. Differences between the individual instruments and the MIM profiles are shown in the middle column, and the rightmost column provides a magnification of the ±20% difference range between 100 and 1 hPa. Error bars indicate the uncertainties in each climatological mean based on the SEM. The grey shaded area indicates where relative differences are smaller than ±5%.

disagreement in the LS with differences of up to 50%. These deviations are accompanied by large uncertainties in the GOMOS climatological mean values. In the NH at high latitudes, all instruments, including SCIAMACHY, agree very well. However, in the SH at high latitudes, differences are in general larger. Between 10 and 100 hPa, all instruments show considerable disagreement, with differences for most instruments of up to ±50%. In particular, SAGE II (MS) and HALOE (LS/MS) show negative offsets compared to all other datasets, causing the large overall spread. This particular comparison focuses on one month of one year, and could be impacted by different sampling patterns over the course of the month. While this could explain some of the large differences, HALOE and OSIRIS sample very similar parts of the month and still show considerable disagreement. For the October comparison OSIRIS agrees very well with the MIM, with differences below ±5% (see **Figures A4.1.10** and **A4.1.16** in *Appendix A4* for a detailed evaluation of OSIRIS monthly zonal means at SH high latitudes). Note that the SEM is larger at high southern latitudes compared to other regions, indicating a higher uncertainty in the climatological mean values. The rightmost panels of

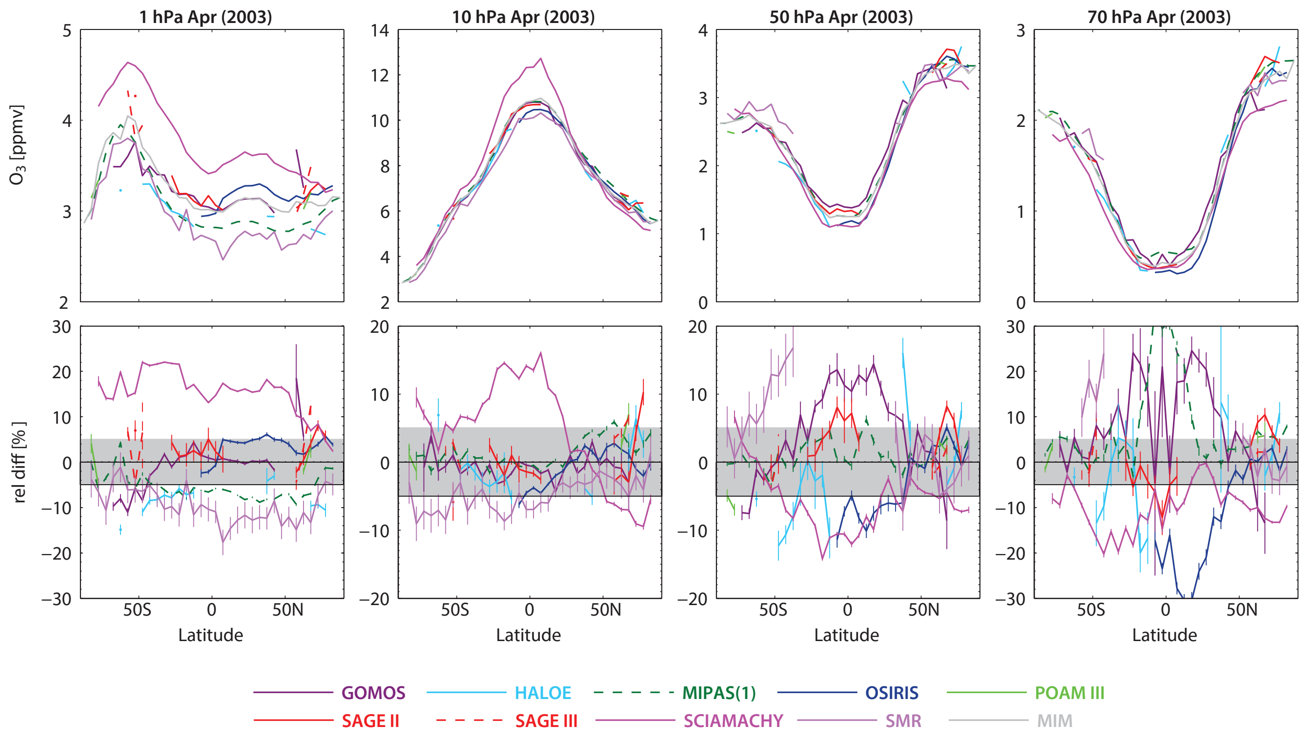


Figure 4.1.9: Meridional profiles of monthly zonal mean ozone for 2003. *Meridional zonal mean ozone profiles at 1, 10, 50 and 70 hPa for April 2003 are shown in the upper row. Differences between the individual instruments and the MIM profiles are shown in the lower row. Error bars indicate the uncertainties in each climatological mean based on the SEM. The grey shaded area indicates where relative differences are smaller than ±5%.*

Figure 4.1.8 shows a magnification of the middle panels, displaying the relative differences in the ±20% range. In the tropics, OSIRIS, SAGE II and GOMOS agree best, while in the polar regions differences among the instruments are spread out more equally.

Meridional profiles at 1, 10, 50 and 70 hPa and their differences with respect to the MIM are compared in **Figure 4.1.9**. At 1 hPa, the instruments show reasonably good agreement, with differences of up to ±20% and with no clear latitudinal structure. At 10 hPa, most measurements agree very well (within ±5%) except for SCIAMACHY, which has an offset in the tropics. At 50 and 70 hPa, the relative differences are smaller in the mid-latitudes than in the tropics.

SMR, OSIRIS, MIPAS(2), GOMOS, SCIAMACHY, ACE-FTS, ACE-MAESTRO, Aura-MLS (2005-2010) and HIRDLS (2005-2007)

Annual zonal mean ozone climatologies for all measurements available for 2005-2010 are shown in **Figure 4.1.10**. For this time period, it is possible to include the more recent instruments ACE-FTS, ACE-MAESTRO, and Aura-MLS. For ACE-MAESTRO, the ozone product derived from the visible spectra is used. HIRDLS is also one of the more recent instruments and covers a time period of three years from 2005 to 2007. HIRDLS is included in the evaluations of the 2005-2010 climatologies but not included in the calculation of the MIM itself. SMR provides a second ozone data product measured at 488.9 GHz, which has very similar characteristics compared to the original SMR ozone product but is not shown in the following evaluations. Differences of all individual climatologies with respect to the MIM are displayed in **Figure 4.1.11**. The instruments show the best overall agreement in the tropical and mid-latitude MS, with OSIRIS, MIPAS(2), GOMOS, ACE-MAESTRO, Aura-MLS and HIRDLS displaying the smallest differences to the MIM of up to ±5%. SMR, SCIAMACHY, and ACE-FTS show good agreement with the other instruments, with positive differences of up to +10% for the latter two and negative differences of up to -10% in case of SMR. While in 2003, four of the instruments yielded very good agreement, not only in the MS but also in the US, in 2005-2010 only HIRDLS data in the US agree very well with the MIM, with differences of less than ±5%. The other instruments show larger differences of up to ±10%, or up to ±20% in case of SMR and SCIAMACHY. Note that the relative differences of SCIAMACHY to the MIM are smaller compared to the earlier time period in 2003, very likely related to the improvement of the vertical resolution of the ECMWF data. Differences of ACE-FTS and ACE-MAESTRO with respect to the MIM show a very similar structure, which is opposite compared to that of OSIRIS, Aura-MLS and GOMOS. The differences between the two instruments ACE-FTS and ACE-MAESTRO and the MIM, including those in the region above 1 hPa, are consistent with a validation study by *Dupuy et al.* [2009]. In general, most instruments display larger differences with respect to the MIM at higher latitudes compared to tropics and mid-latitudes. In particular, this result can be observed for ACE-FTS and ACE-MAESTRO. Instruments agree reasonably well in the LS with differences of up to ±20%, with the exception of OSIRIS, GOMOS, ACE-MAESTRO and HIRDLS, which can differ locally up to ±50% from the MIM. While these

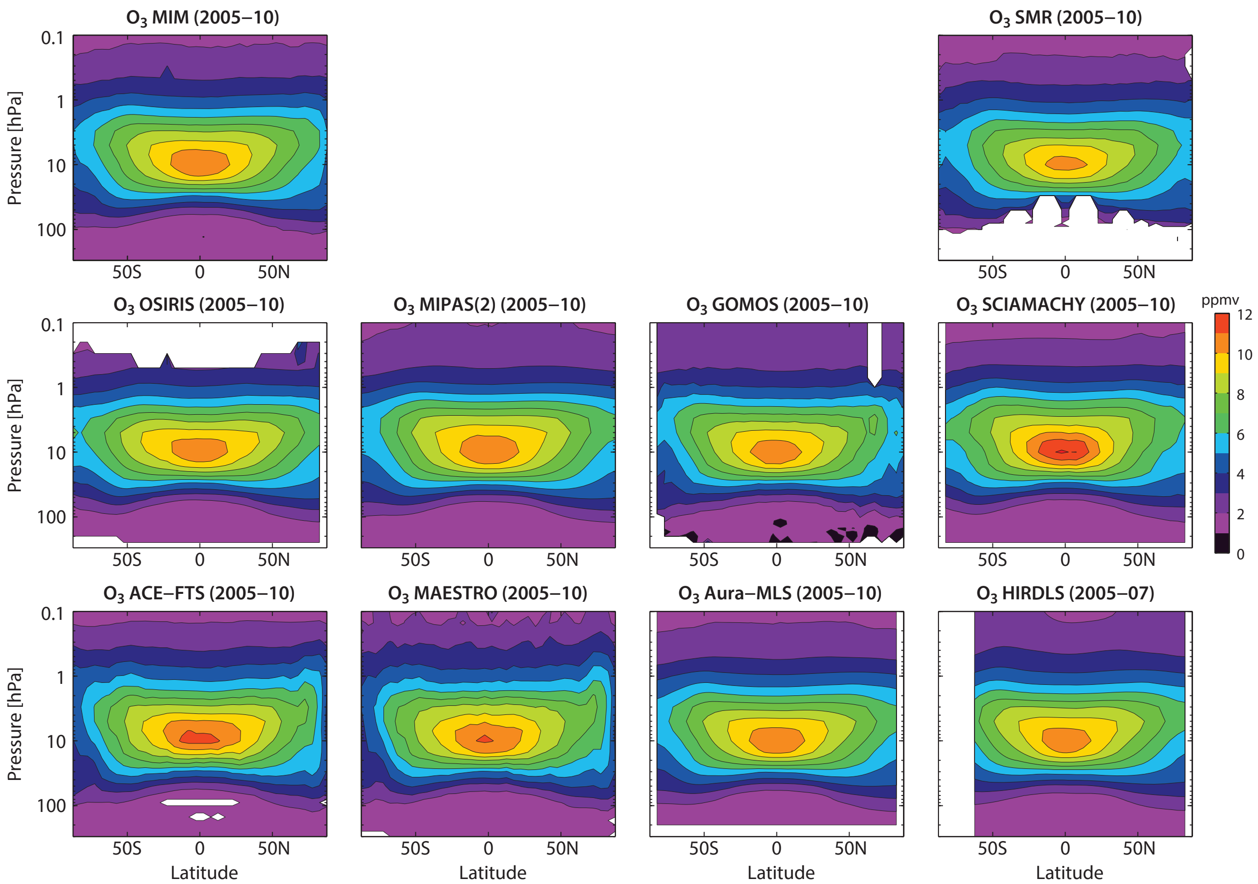


Figure 4.1.10:* *Cross sections of annual zonal mean ozone for 2005-2010. *Annual zonal mean ozone cross sections for 2005-2010 are shown for the MIM, SMR, OSIRIS, MIPAS(2), GOMOS, SCIAMACHY, ACE-FTS, ACE-MAESTRO, Aura-MLS and HIRDLS. Note that HIRDLS covers only 2005-2007 and is not included in the MIM.*

large LS differences are only present in the tropics for most instruments, GOMOS also shows considerable disagreement of up to ±50% in the mid- and high latitude LS.

The monthly mean vertical ozone profiles displayed in **Figure 4.1.12** show a similar picture compared to the results from 1994-1996 and 2003 (see **Figures 4.1.4** and **4.1.8**), with the largest differences in the SH mid-latitude spring. In the NH mid- and high latitudes in the MS, nearly all instruments agree very well (within ±5%), while in the tropics slightly larger differences of up to ±10% can be found. As noted for the 2003 time period, SCIAMACHY clearly overestimates the ozone mixing ratio peak at 10 hPa in the tropics by 10%. Differences in the LS are larger in the tropics (up to ±50%) when compared to NH mid- and high latitudes (up to ±10%), with the exception of GOMOS. In the SH high latitudes, the instruments agree only reasonably well with differences of up to ±20%. In particular, GOMOS shows a large negative offset compared to the other instruments. The magnification plots in the rightmost panels reveal that in most latitude bands Aura-MLS, OSIRIS, and HIRDLS measurements are very close to each other (including GOMOS in the tropics and mid-latitude MS/US).

Meridional ozone profiles are shown in **Figure 4.1.13**. At 1 hPa, differences are similar to the 2003 time period, with no clear meridional structure. While in 2003 only SCIAMACHY shows a large positive deviation from all other datasets at 1 hPa, now ACE-FTS and ACE-MAESTRO also disagree strongly with the MIM by up to +20%. Relative differences at 10 hPa are comparable to the 2003 time period with the most prominent feature being the overestimation of the ozone peak by SCIAMACHY, leading to steeper latitudinal gradients for SCIAMACHY compared to all other instruments. At 50 and 70 hPa, relative differences are larger than for the upper levels, especially for ACE-MAESTRO, HIRDLS, and GOMOS. The latter shows a noisy meridional profile with spikes, which was not observed in its 2003 climatology.

LIMS and SAGE-I (1979)

The oldest satellite measurements of ozone profiles are available from the LIMS and SAGE I instruments. They only overlap for 4 months from February to May 1979 and the monthly mean ozone climatologies for March and April 1979 are shown in **Figure 4.1.14**. The monthly mean differences of both datasets with respect to their mean are displayed in **Figure 4.1.15**. In the MS, both instruments show excellent agreement with differences to their MIM mostly within ±2.5% for all latitude bands (corresponding to a direct difference between the two instruments of less than 5%). In the LS, differences are larger; up to ±20% in the tropics and ±10% in the mid-latitudes. For both months shown here (and also for February) LIMS has mostly negative

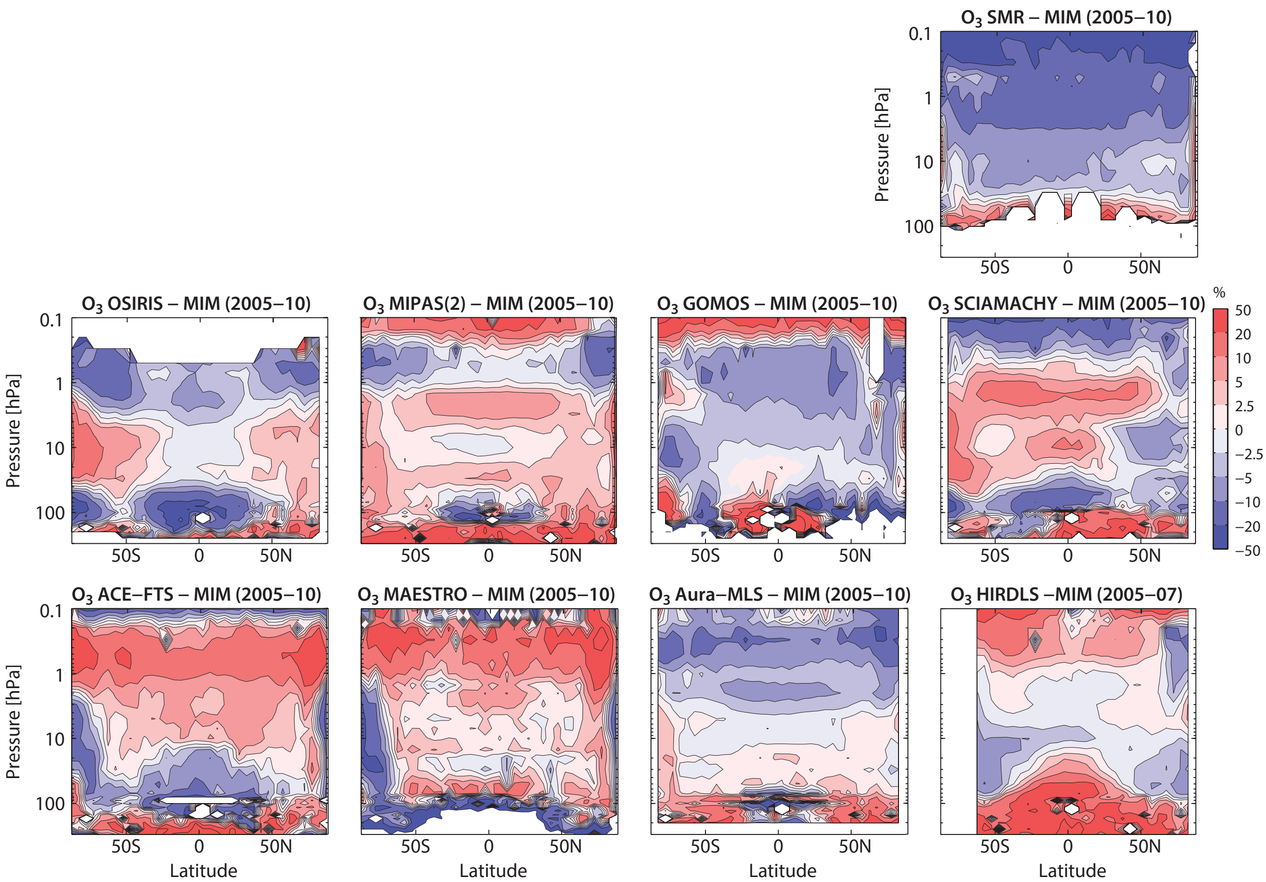


Figure 4.1.11: Cross sections of annual zonal mean ozone differences for 2005-2010. *Annual zonal mean ozone differences for 2005-2010 between the individual datasets and the MIM are shown. Note that HIRDLS covers only 2005-2007 and is not included in the MIM.*

deviations when compared to SAGE I. Note that the differences are reversed in May when LIMS has a mostly positive deviation from SAGE I (see **Figures A4.1.25** – **A4.1.26** in *Appendix A4*). This difference is very likely related to SAGE I sampling issues in May (when its sunrise measurements occur only in early May at NH latitudes and its sunset measurements occur only in the SH), and therefore the monthly mean SAGE I values yield results that are not representative of the true monthly zonal mean differences between the two instruments.

SMILES (2010)

Figure 4.1.16 shows zonal mean ozone cross sections averaged from January until April 2010 for SMILES and the MIM of all available instruments (ACE-FTS, ACE-MAESTRO, Aura-MLS, GOMOS, MIPAS(2), OSIRIS, SCIAMACHY, SMR, and SMILES) for the same time period. The corresponding relative differences can be seen in **Figure 4.1.17**. In the MS, SMILES shows differences of up to ±5% with a positive deviation to the MIM between 5 and 20 hPa and a negative deviation above and below this region. While in the MS, SMILES agrees very well with the other instruments, in the US differences of up to -20% are found, yielding only a reasonable agreement.

SAGE II (1991-2005)

Comparisons of SAGE II to ozonesondes show generally very good agreement, with small biases only at the lowest altitudes [*Wang et al.*, 2002]. Since SAGE II has a very long data record and has been used extensively in validation and long-term studies (it is often referred to as the "gold standard"), it is of interest to use SAGE II as a reference for comparisons with other satellite measurements. While the comparison of an instrument to the MIM provides information on how a respective dataset is related to all other available measurements, the comparison to SAGE II can identify those instruments closest to the longest available data record in any given region, and therefore best able to extend the SAGE II record. **Figure 4.1.18** shows the difference between each individual dataset available and SAGE II. The comparisons are derived for the maximum overlap time period for each individual instrument, *i.e.*, each comparison is based on a different time period varying from 15 years for the comparison to HALOE to 6 months for the comparison to HIRDLS. Note that SAGE II measurements stop in August 2005, thereby marking the end of the comparison time period.

In the tropical and mid-latitude MS, GOMOS and Aura-MLS show excellent agreement, with differences below ±2.5%, while UARS-MLS, HALOE, OSIRIS, SAGE III, and

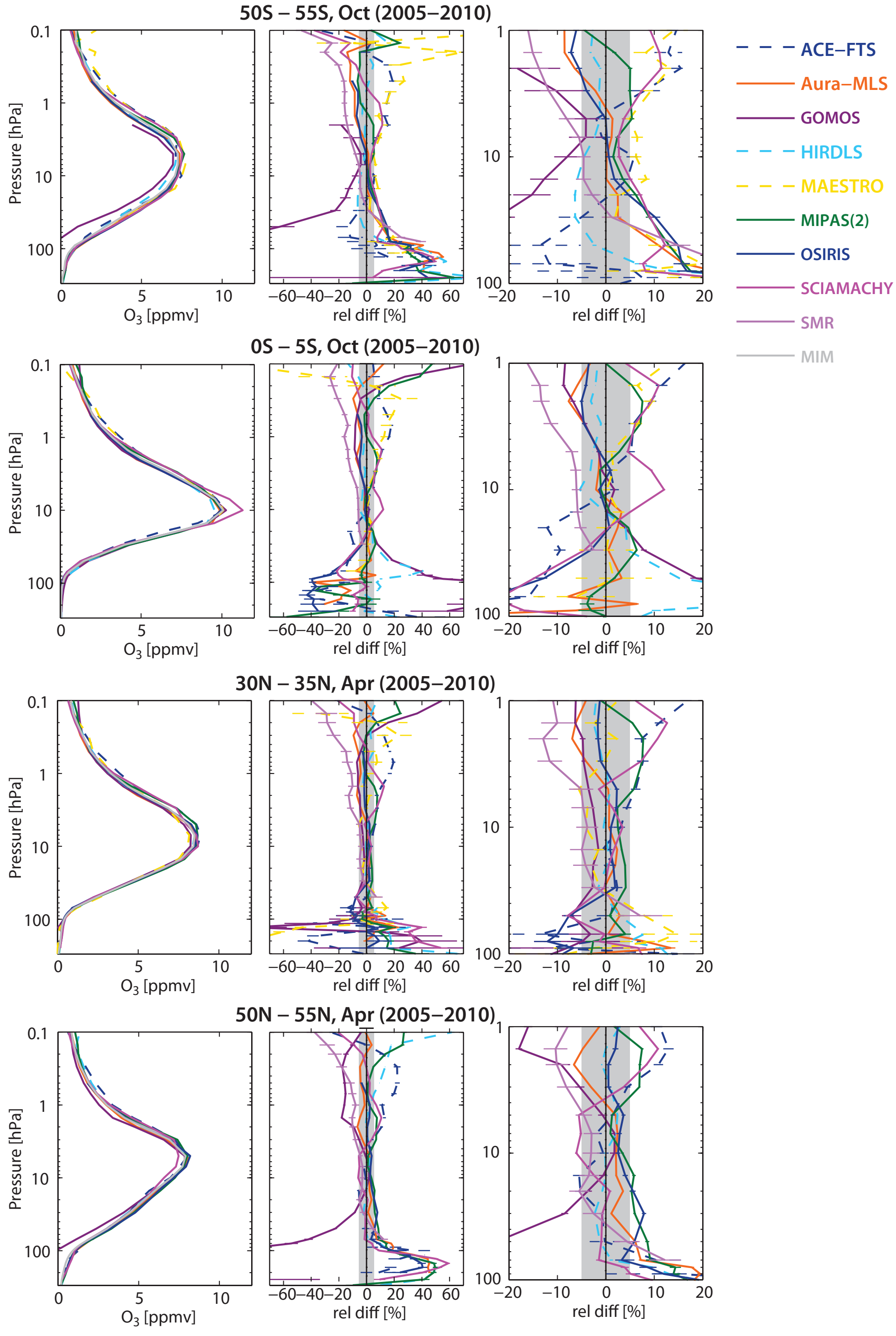


Figure 4.1.12: Vertical profiles of monthly zonal mean ozone for 2005-2010. *Zonal mean ozone profiles for 50°S-55°S and 0°S-5°S for October 2005-2010 and for 30°N-35°N and 50°N-55°N for April 2005-2010 are shown in the leftmost column. Differences between the individual instruments and the MIM profiles are shown in the middle column, and the rightmost column provides a magnification of the ±20% range between 100 and 1 hPa. Error bars indicate the uncertainties in each climatological mean based on the SEM. The grey shaded area indicates where relative differences are smaller than ±5%. Note that HIRDLS covers only 2005-2007 and is not included in the MIM.*

MIPAS(1) have only slightly larger deviations to SAGE II, often up to ±5%. The largest departure from SAGE II can be found for ACE-MAESTRO, ACE-FTS, SCIAMACHY and MIPAS(2) with differences of up to ±20%. For MIPAS(2), this discrepancy is a known characteristic that has already been identified in an earlier data version [*Stiller et al.*, 2012], as well as in all existing MIPAS data processors (A. Laeng, pers. comm.). Thus, it suggests a problem in MIPAS level-1 data rather than at a peculiarity in one of the retrieval processors. For all instruments, differences in the absorption cross sections will account for some of the differences between the climatologies. For example, the ozone cross section used in the SAGE II retrieval (V6.2) is about 2% lower compared to the one used by GOMOS (Chappuis region). Neglecting other potential systematic differences, we would then expect SAGE II to be about 2% larger than

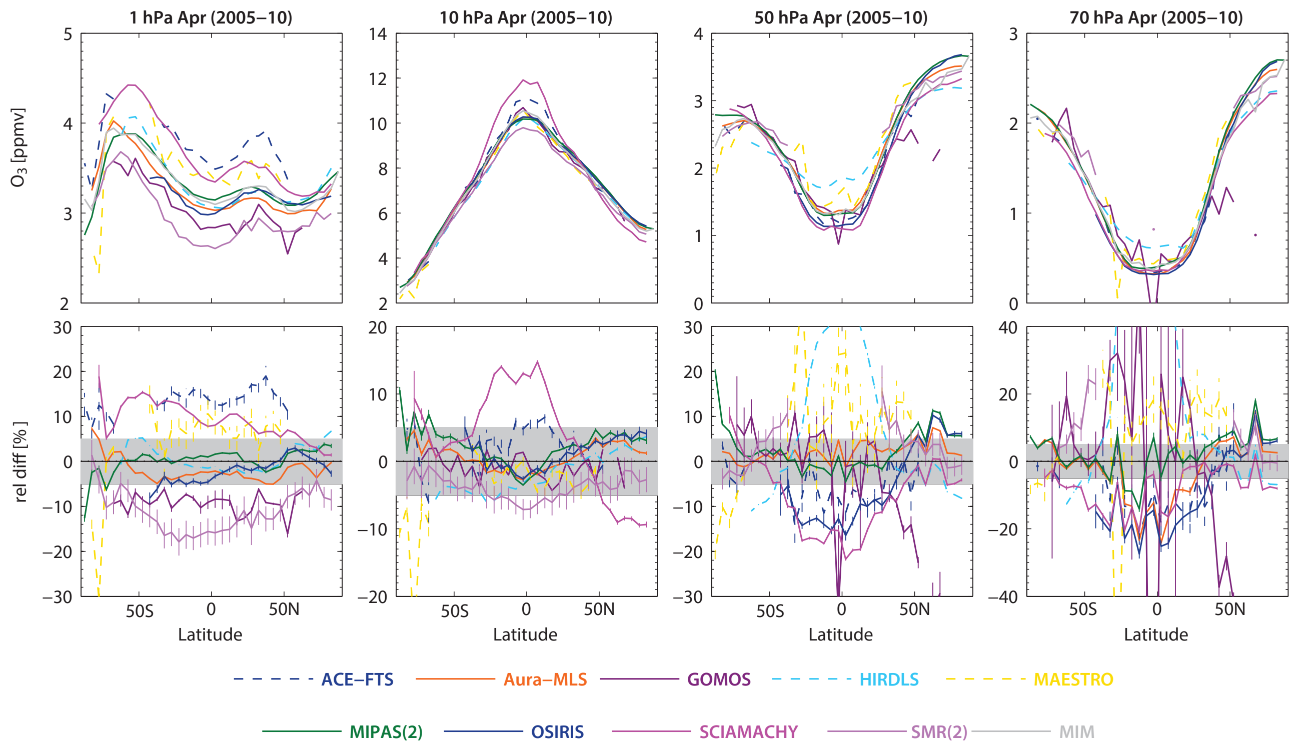


Figure 4.1.13: Meridional profiles of monthly zonal mean ozone for 2005-2010. *Meridional zonal mean ozone profiles at 1, 10, 50 and 70 hPa for April 2005-2010 are shown in the upper row. Differences between the individual instruments and the MIM profiles are shown in the lower row. Error bars indicate the uncertainties in each climatological mean based on the SEM. The grey shaded area indicates where relative differences are smaller than ±5%. Note that HIRDLS covers only 2005-2007 and is not included in the MIM.*

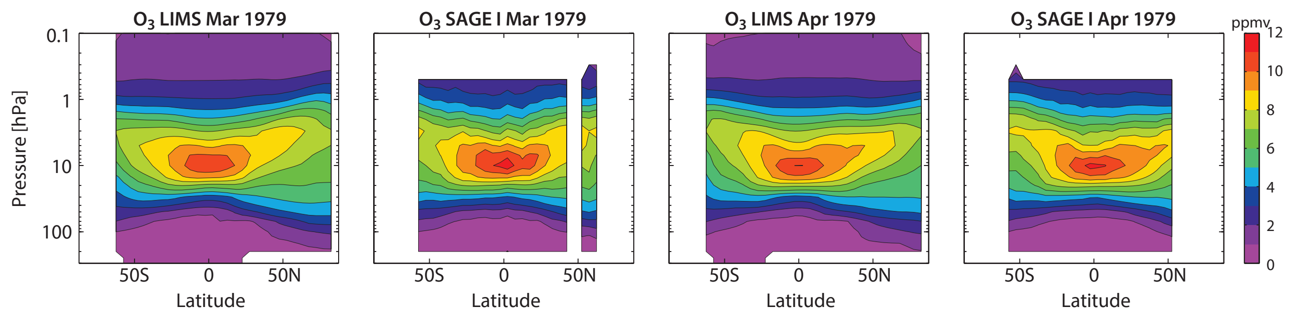


Figure 4.1.14: Cross sections of monthly zonal mean ozone for 1979. *Monthly zonal mean ozone cross sections for 1979 are shown for LIMS and SAGE I for March and April.*

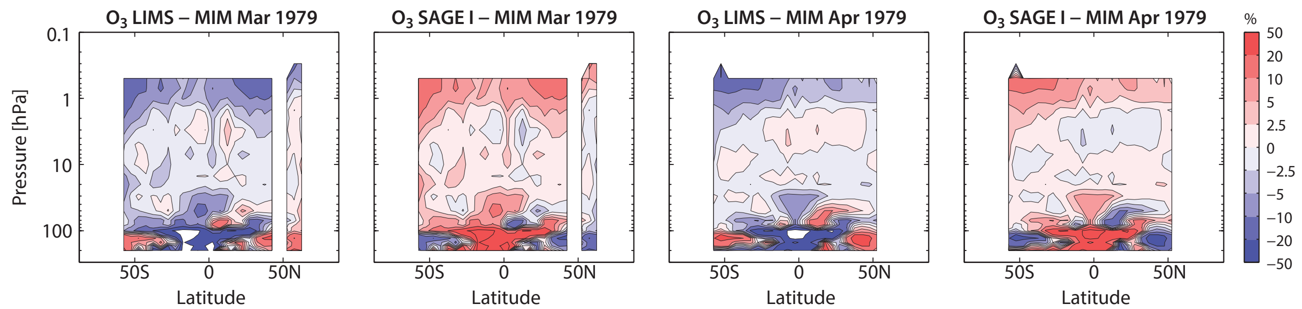


Figure 4.1.15: Cross sections of monthly zonal mean ozone differences for 1979. *Monthly zonal mean ozone differences for 1979 between LIMS, SAGE I and the MIM are shown for March and April.*

GOMOS due to the different ozone cross sections, which is in fact the case in the MS.

In the tropical LS, Aura-MLS, OSIRIS, and MIPAS(1/2) display the best agreement with SAGE II data. In the tropical UT, nearly all datasets (except HALOE and ACE-MAESTRO) show larger values than SAGE II. This result might be related to a low bias with respect to ozone-sondes that SAGE II has in this region [*Wang et al.*, 2002]. An intriguing feature is that nearly all datasets show larger differences

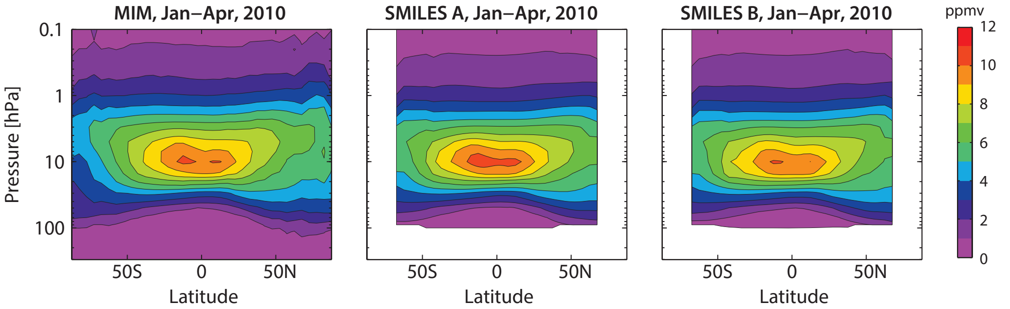


Figure 4.1.16: Cross sections of zonal mean ozone for 2010. *Zonal mean ozone cross sections for January-April 2010 are shown for the MIM, SMILES-A and SMILES-B.*

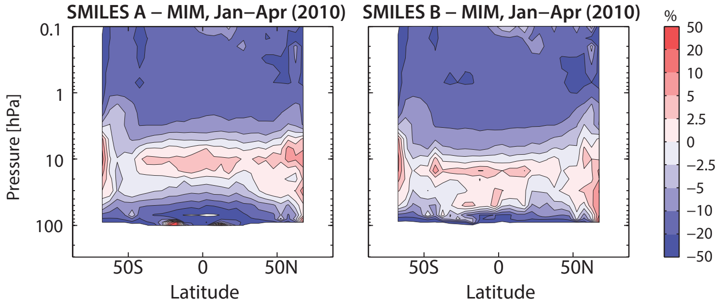


Figure 4.1.17: Cross sections of zonal mean ozone differences for 2010. *Zonal mean ozone differences for January-April 2010 between SMILES-A, SMILES-B and the MIM are shown.*

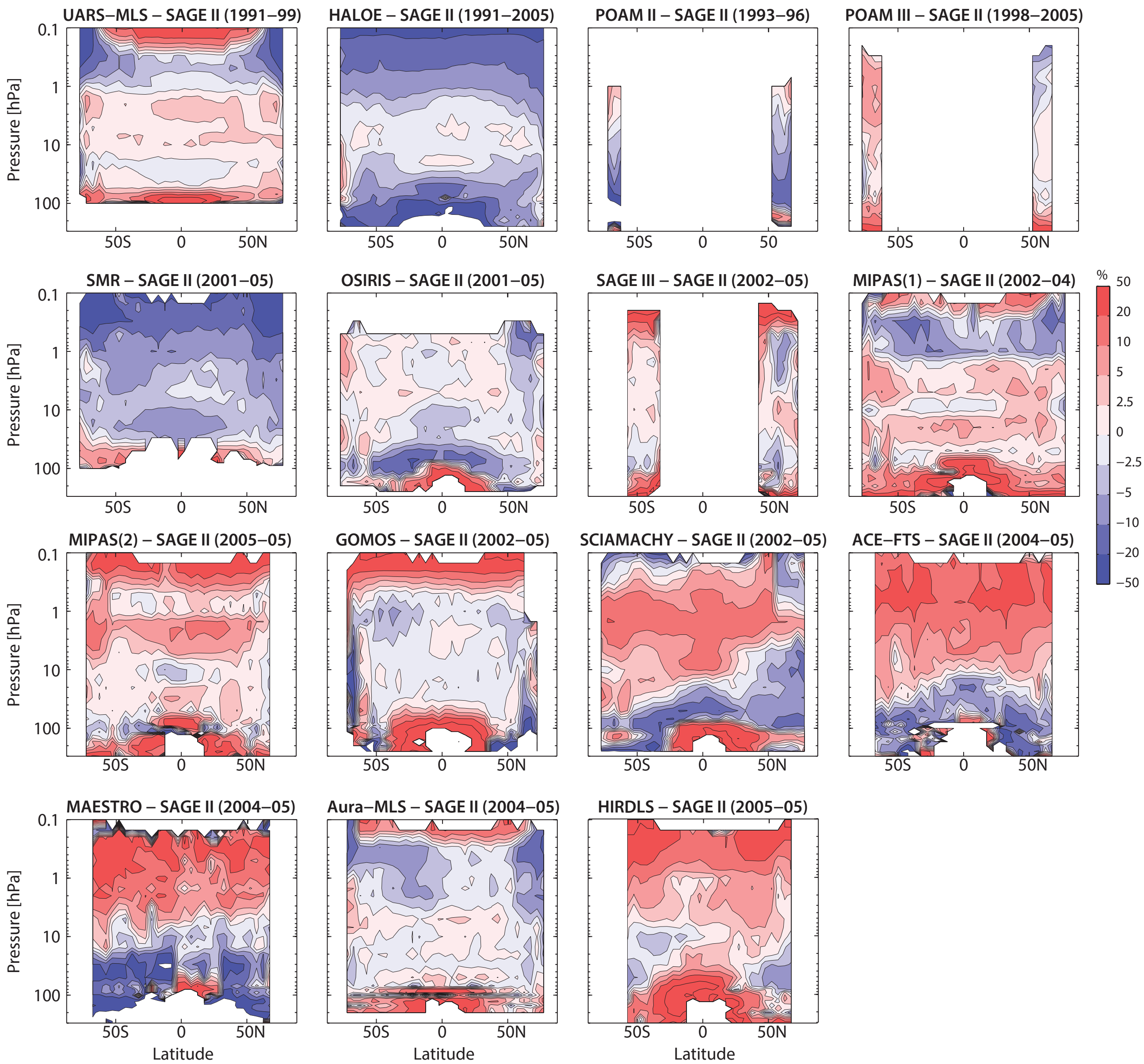


Figure 4.1.18: Cross sections of zonal mean ozone differences to SAGE II. *Zonal mean ozone differences between the individual instruments and SAGE II are shown for time periods of maximum overlap.*

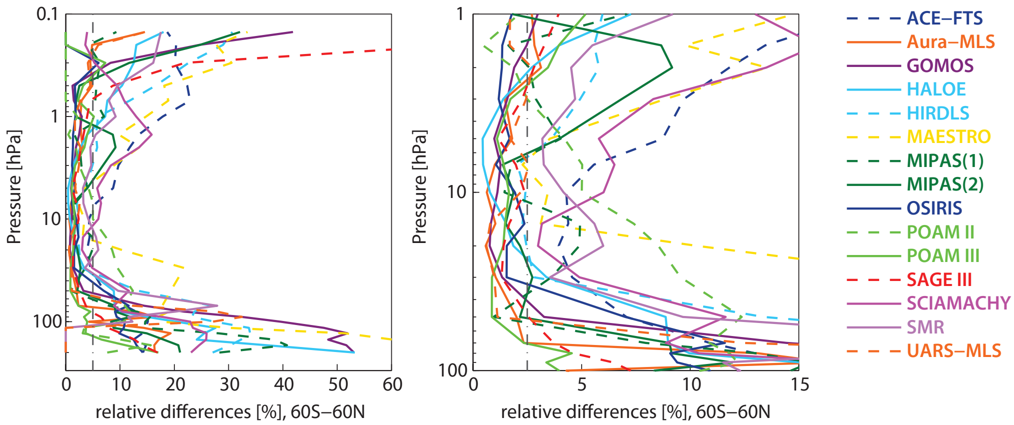


Figure 4.1.19: Vertical profiles of mean ozone differences to SAGE II between 60°S and 60°N. *Absolute values of the relative ozone differences averaged between 60°N and 60°S for time periods of maximum overlap between the individual instruments (HALOE, POAM III, SAGE III, ACE-FTS, Aura-MLS, GOMOS, MIPAS, OSIRIS, SCIAMACHY, SMR, ACE-MAESTRO, and HIRDLS) and SAGE II are shown. The right panel provides a magnification of the 100 to 1 hPa region.*

poleward of 60°S compared to other latitudes. In the SH, Aura-MLS, OSIRIS, HALOE, SMR, and UARS-MLS have only small offsets compared to SAGE II of up to ±10%, while the other instruments show larger differences of up to ±20% or even ±50% in the case of GOMOS. In the NH polar latitudes, HIRDLS, UARS-MLS, OSIRIS, and MIPAS(1) agree well and POAM III agrees very well with SAGE II.

Figure 4.1.19 shows profiles of the absolute values of the differences averaged between 60°N and 60°S. The magnification shown in the right panel demonstrates that between 1 and 50 hPa more than half of the instruments agree very well with SAGE II, showing mean deviations of less than 5%. Below 50 hPa and above 1 hPa, a large spread between the individual instruments can be found, with differences as small as 10-20% or more than 50% in some cases. The overall best agreement to SAGE II in the tropics and mid-latitudes is observed for Aura-MLS, OSIRIS, and MIPAS(2) in the LS, Aura-MLS and GOMOS in the MS, and OSIRIS and POAM II in the US.

4.1.3 O_3 evaluations: Seasonal cycles

Tropical ozone exhibits a large annual cycle near and above the tropopause that is related to seasonal changes in vertical transport acting on the strong vertical ozone gradient in this region [*Folkins et al.*, 2006; *Randel et al.*, 2007]. Although the annual cycle extends over only a narrow vertical range, from approximately 100 to 50 hPa, it is an important characteristic of tropical ozone in the LS and can be used to analyse the seasonal cycle in tropical upwelling. Ozone above 10 hPa exhibits a strong semi-annual cycle associated with the tropical semiannual oscillation (SAO) in zonal wind and temperature [*Ray et al.*, 1994].

The upper panels of **Figure 4.1.20** show the seasonal cycle of tropical monthly mean ozone at 10 hPa for the three time periods. All instruments display the semiannual cycle, which is characterised by a stronger amplitude during the first half of the calendar year. For the time period 1994-1996, the available instruments agree quite well and display very similar phase and amplitude. However, for the later time periods (2003 and 2005-2010), larger differences can be observed in terms of the absolute mean values as well as amplitude and phase of the seasonal cycle. (Note that only amplitude and phase of the seasonal cycle and not the mean values are evaluated by the Taylor diagrams, see *Section 3.3.3*) OSIRIS and SCIAMACHY display larger amplitudes than the other instruments and show deviations from the phase of the MIM seasonal cycle. GOMOS, SAGE II, MIPAS(1) and Aura-MLS are closest to the MIM, with only small differences in the phase and hence yield the best skill scores. SMR agrees well with this group of instruments for the 2003 time period, but has a lower amplitude for the seasonal cycle in the later time period (2005-2010). While HIRDLS agrees quite well during the first half of the year, its amplitude and mean values during the second half of the year are too low. Similarly, HALOE does not capture the second maximum in the seasonal cycle and fails to display the semiannual signal. Due to their limited temporal sampling in the tropics ACE-MAESTRO and ACE-FTS climatologies provide only weak constraints for the seasonal cycle in this region and interpretation of the Taylor diagram has to account for this issue. ACE-MAESTRO's monthly mean values are very close to the MIM, except for the June value, which is much higher than expected and as a consequence prevents fitting a seasonal cycle for ACE-MAESTRO. The monthly mean values derived for ACE-FTS fluctuate slightly more about the MIM values than those from ACE-MAESTRO, however, they still allow for a reasonably well defined seasonal cycle. SCIAMACHY values are much larger than the other climatologies year-round, and well above the 1σ multi-instrument standard deviation. On the other side of the range, SMR, HALOE, and OSIRIS values are lower than the MIM but within 1σ during most of the year. In general, the instruments during the earlier time period (SAGE II, HALOE and UARS-MLS) agree better in terms of absolute values and in terms of the seasonal cycle than the instruments during the later time periods. Instruments that show strong deviations from the MIM in terms of absolute values also have trouble capturing the seasonal cycle.

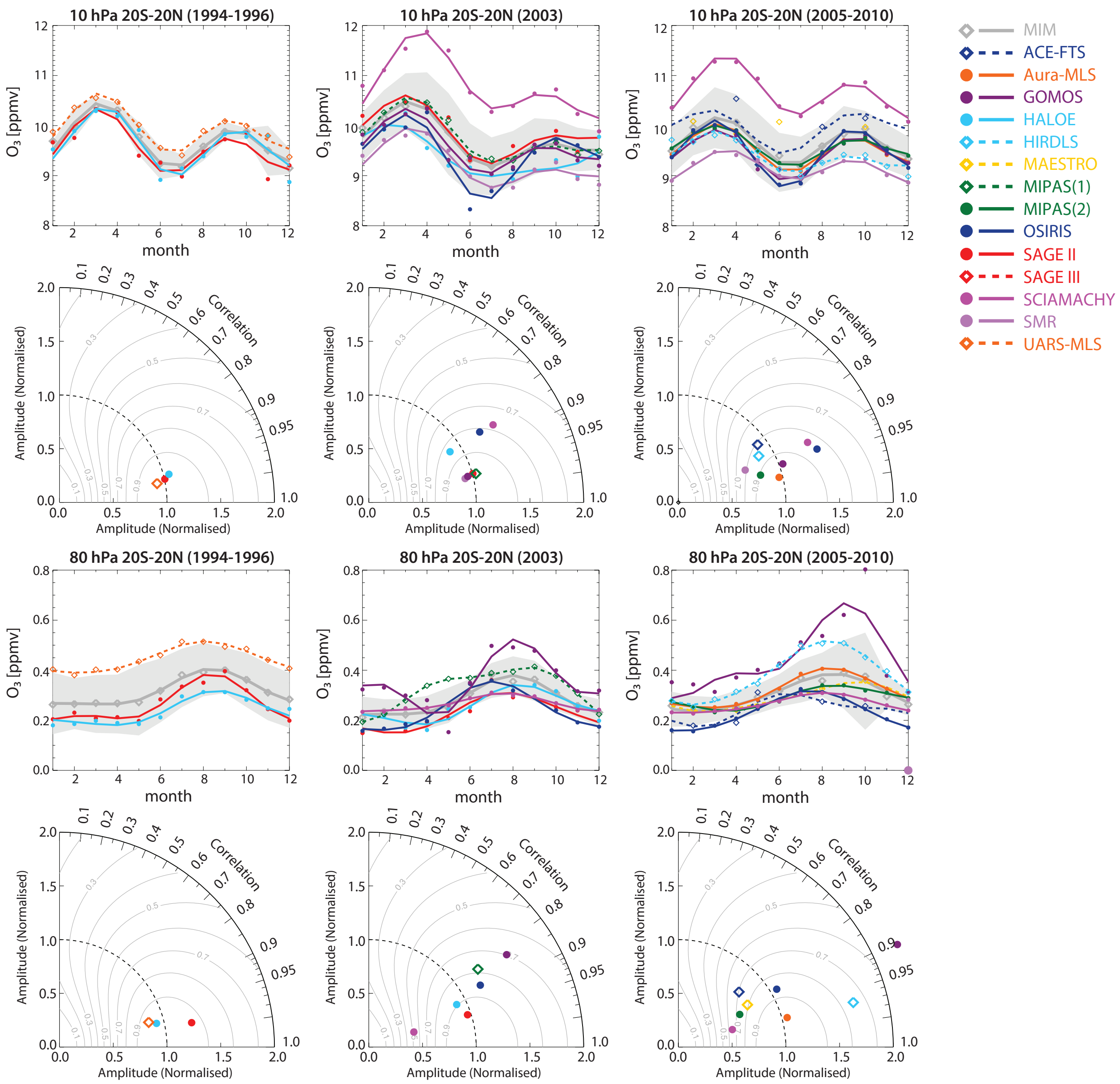


Figure 4.1.20: Seasonal cycle of ozone in the tropics at 10 and 80 hPa. *Seasonal cycle and corresponding Taylor diagram of monthly zonal mean ozone for 20°S-20°N at 10 hPa and 80 hPa. The seasonal cycle is shown for 1994-1996 (left column), 2003 (middle column) and 2005-2010 (right column). The grey shading indicates the MIM±1σ multi-instrument standard deviation.*

The strong annual cycle at 80 hPa observed by ozonesondes [*Randel et al.*, 2007] is more difficult for the satellite measurements to reproduce, and the instruments show a less consistent picture compared to 10 hPa (**Figure 4.1.20**). For all three time periods, large differences in the amplitude of the annual cycle can be observed. The tropical ozone values from UARS-MLS are significantly larger than the SAGE II [*Livesey et al.*, 2003] and HALOE data. However, the amplitude of the seasonal cycle is very similar for UARS-MLS and HALOE, while SAGE II displays a considerably larger amplitude, possibly because of its better vertical resolution (*i.e.*, a version of SAGE II values with degraded vertical resolution would have a smaller amplitude). Despite the differences in the amplitude and absolute mean values, all three datasets show very similar phase, with maximum values between July and September. For the later time periods 2003 and 2005-2010, all instruments show elevated values in NH summer. There is, however, no agreement between the instruments regarding the amplitude or phase of the annual cycle. SAGE II, HALOE and Aura-MLS agree best with the MIM and hence yield the highest skill scores. While the phase of the SCIAMACHY and HIRDLS seasonal cycle is very similar to the MIM, they show a much smaller (SCIAMACHY) or much larger (HIRDLS) amplitude of the signal. The larger amplitude seen by HIRDLS may result from its higher vertical resolution (similar to that noted above for SAGE II), which can be important when observing a feature with a small vertical extent such as the LS ozone annual cycle. GOMOS and MIPAS(1) values are above the 1σ multi-instrument standard deviation

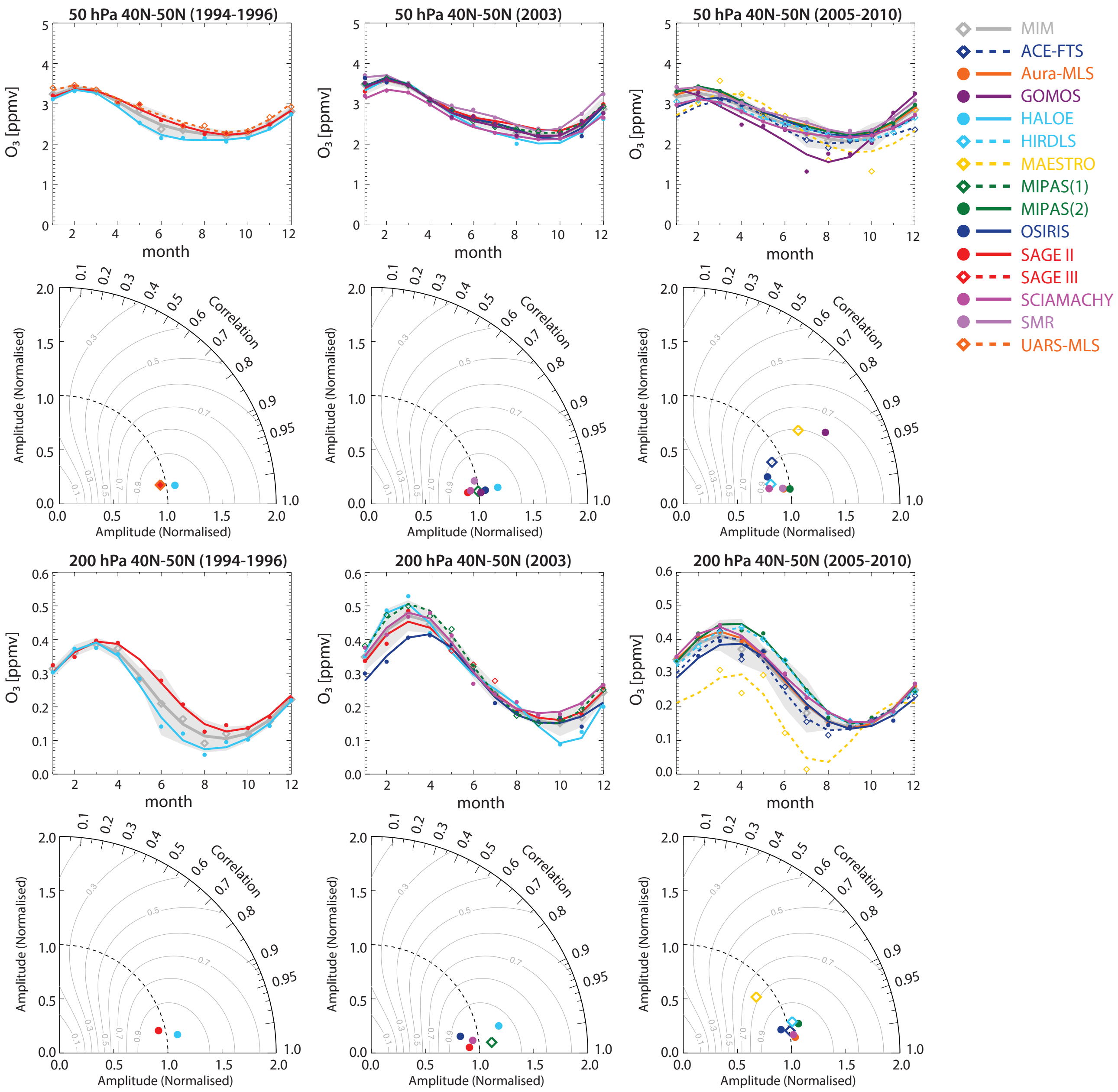


Figure 4.1.21: Seasonal cycle of ozone in the NH mid-latitudes at 50 and 200 hPa. *Seasonal cycle and corresponding Taylor diagram of monthly zonal mean ozone for 40°N-50°N at 50 hPa and 200 hPa. The seasonal cycle is shown for 1994-1996 (left column), 2003 (middle column) and 2005-2010 (right column). The grey shading indicates the MIM±1σ multi-instrument standard deviation.*

for some parts of the year, and additionally show a different seasonal cycle pattern than the other instruments with a second peak in winter or spring, respectively. MIPAS(2), ACE-FTS, and ACE-MAESTRO exhibit a small amplitude seasonal cycle, with the latter two showing considerable differences in their phase with maximum values in summer or late autumn. Due to the large deviations of UARS-MLS on this level, the range of the absolute mean values is better constrained for the later time periods, as opposed to the 10 hPa level. The difficulties of reproducing the annual cycle in ozone at the tropical tropopause are related to the strong vertical gradient in ozone in this region, and the narrow vertical region over which the annual cycle extends [*Randel et al.*, 2007], which require high vertical resolution measurements to be adequately resolved. Also, instrumental limitations resulting from cloud interference and high extinction exist in this altitude region.

The ozone seasonal cycle in NH mid-latitudes at 50 and 200 hPa is shown in **Figure 4.1.21**. For both pressure levels we find a clear annual cycle, with a maximum in early spring and a minimum in late summer/fall related to the changes in transport of the large scale stratospheric circulation. At 50 hPa, the annual cycle is well reproduced by all instruments for all three time periods with the exception of GOMOS and ACE-MAESTRO. Both overestimate the amplitude of the seasonal cycle and do not reproduce the timing of the signal correctly. While the absolute mean values agree very well for all time periods and instruments, GOMOS and especially ACE-MAESTRO have values well

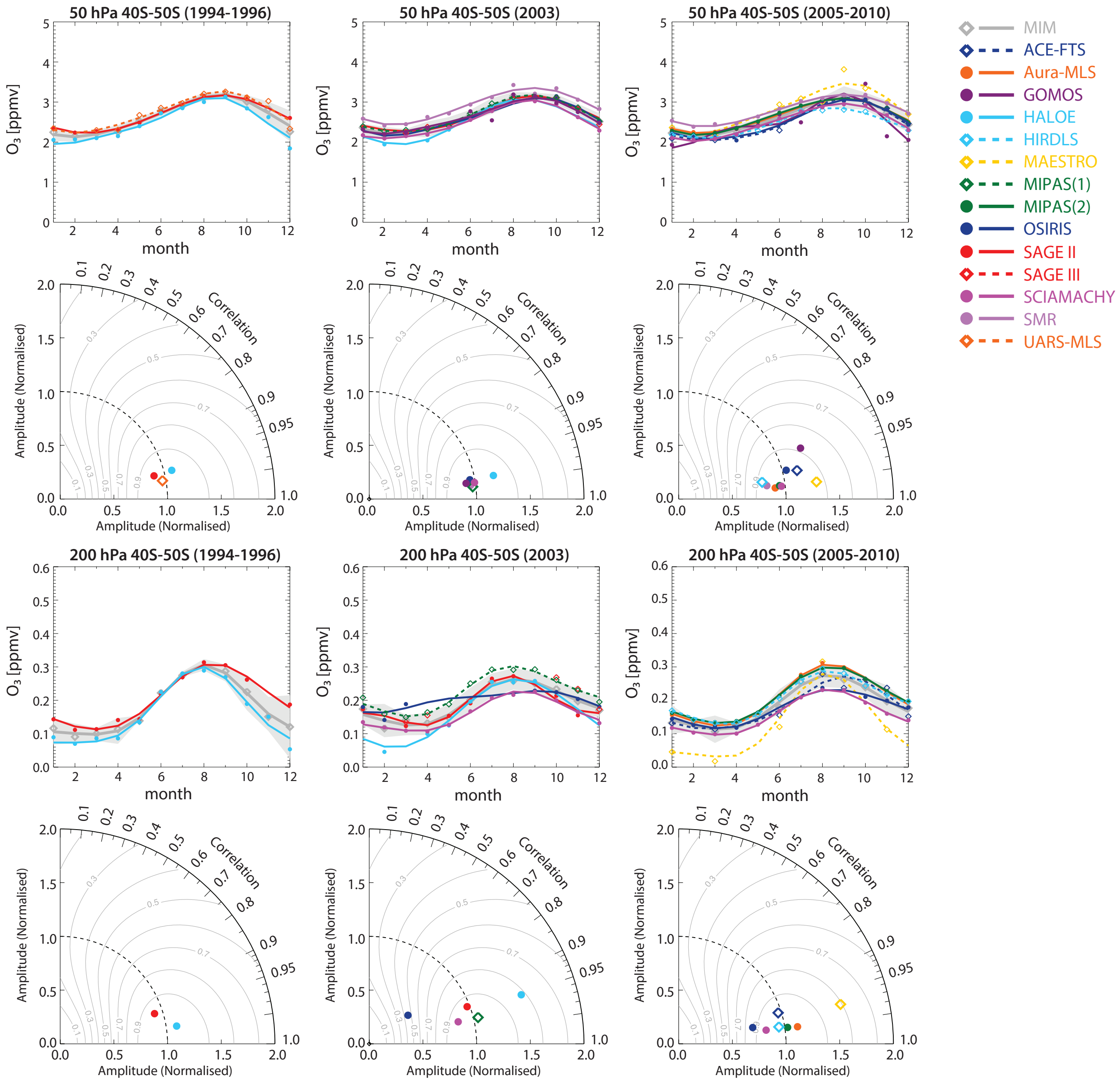


Figure 4.1.22: Seasonal cycle of ozone in the SH mid-latitudes at 50 and 200 hPa. *Seasonal cycle and corresponding Taylor diagram of monthly zonal mean ozone for 40°S-50°S at 50 hPa and 200 hPa. The seasonal cycle is shown for 1994-1996 (left column), 2003 (middle column) and 2005-2010 (right column). The grey shading indicates the MIM±1σ multi-instrument standard deviation.*

below the 1σ range for some parts of the year. At 200 hPa, results are similar and nearly all instruments capture the pronounced seasonal cycle. While in 1994-1996 differences between the instruments mean values are well constrained in autumn and winter and large in spring and summer, the situation is reversed for 2003, where the 1σ range over all instruments is small in spring and summer but large during the rest of the year. In 2003, OSIRIS slightly under-estimates the amplitude of the seasonal cycle while HALOE has a too large amplitude. In 2005-2010, ACE-MAESTRO has a negative offset compared to all other climatologies and the minimum of the seasonal cycle in summer rather than autumn, which results in a low score skill. Note that GOMOS is not displayed here since it shows large deviations from all other instruments at this lower level.

Figure 4.1.22 shows the seasonal cycle for the SH mid-latitudes at 50 and 200 hPa. The dominant signal is an annual cycle with a maximum in SH late summer/fall related to the transport processes of the large scale stratospheric circulation and shifted by half a year *versus* the corresponding NH signal. At 50 hPa, the mean values are well constrained and the annual cycle is well reproduced by all instruments. Small deviations are found for HALOE, ACE-MAESTRO and GOMOS (2005-2010), which exhibit slightly too high amplitudes, as well as HIRDLS and SMR (2005-2010) with slightly too low amplitudes. While in the NH the instruments agree well at both levels (50 and 200 hPa), in the SH there is a larger spread in the seasonal cycle at 200 hPa. The biggest discrepancies are found for HALOE, OSIRIS and ACE-MAESTRO. In particular, OSIRIS in 2003 does not capture the signal and displays a nearly flat line, yielding very low skill scores, likely related to its limited sampling with no measurements in the winter hemisphere. As already noted for the NH differences between SAGE II and HALOE, absolute mean values are well constrained in autumn and winter and very large in spring and summer. Again the seasonal cycle from GOMOS at the 200 hPa level is not shown due to large deviations.

4.1.4 O_3 evaluations: Interannual variability

Ozone is characterised by strong interannual variability related to a number of chemical and dynamical processes. These processes include the QBO signal, variations in the Brewer-Dobson circulation, solar cycle, strong volcanic eruptions and the variability of the polar vortex strength. The impact of the individual processes on the interannual ozone variability changes with altitude, latitude and time. Evaluating time series of deseasonalised ozone anomalies helps to understand how well the sensitivity of the ozone abundance to the various processes is captured by the individual satellite instruments.

Figure 4.1.23 shows the time series of deseasonalised ozone anomalies at 10 hPa from 2000 to 2010. The variability of ozone anomalies in the tropics is dominated by an approximately two-year long cycle that is linked to the QBO. Most instruments successfully reproduce the QBO-ozone cycle. Interannual anomalies from Aura-MLS, GOMOS, MIPAS(1), SCIAMACHY, and SMR agree best. Also, SAGE II shows stronger month-to-month fluctuations than the other instruments. In the polar regions at 10 hPa, there is no periodic cycle governing ozone variability, but strong month-to-month fluctuations. The largest anomalies for NH polar ozone can be found in winter/early spring and are related to the strength of the polar vortex. For some years (*e.g.*, 2006, 2007, 2009) there is a considerable spread between the individual instruments. This might be related to the choice of latitude band, which can be partially inside or outside the polar vortex. As a result, differences in the longitudinal satellite sampling patterns can

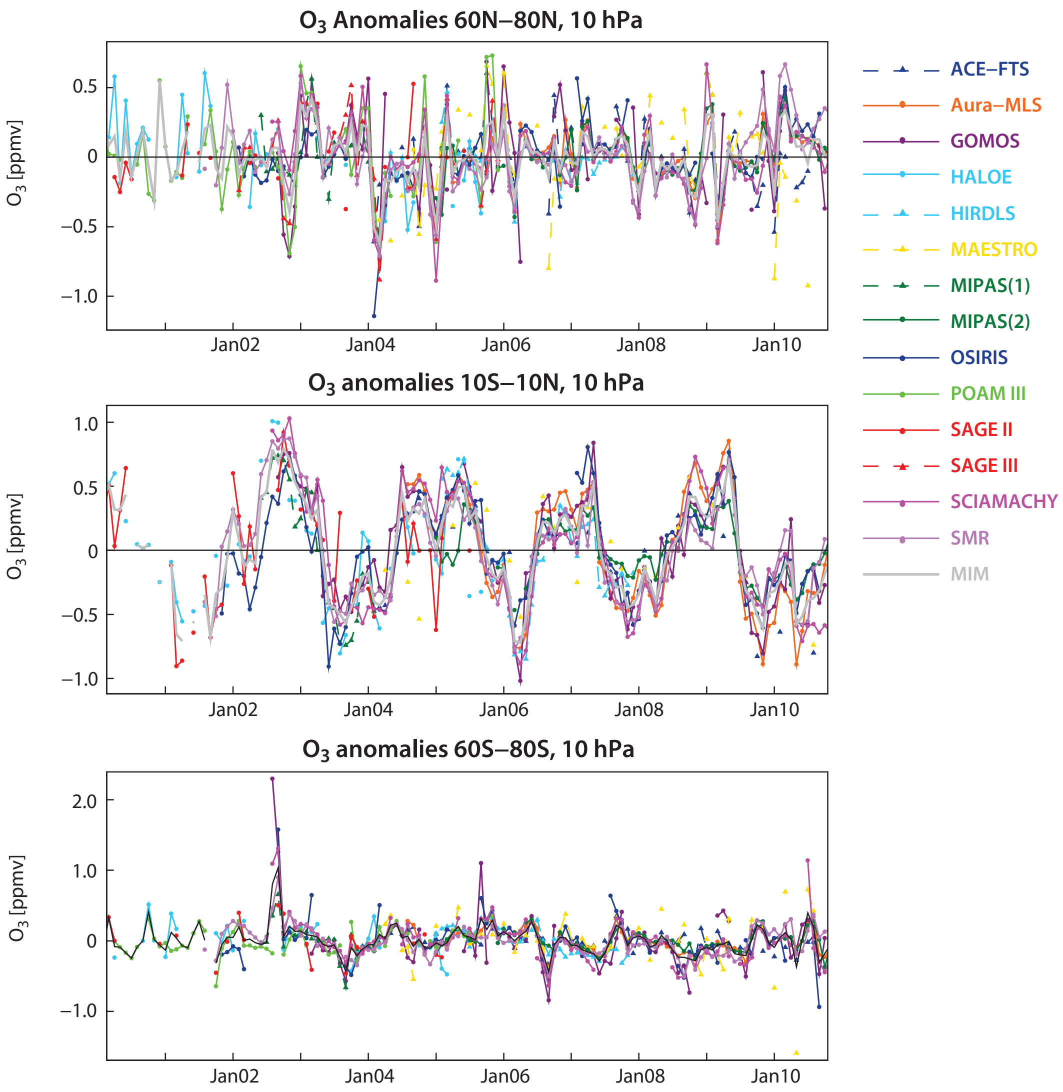


Figure 4.1.23: Time series of deseasonalised ozone anomalies at 10 hPa. *Deseasonalised ozone anomalies at 10 hPa between 60°N – 80°N (upper panel), 10°S – 10°N (middle panel) and 60°S – 80°S (lower panel).*

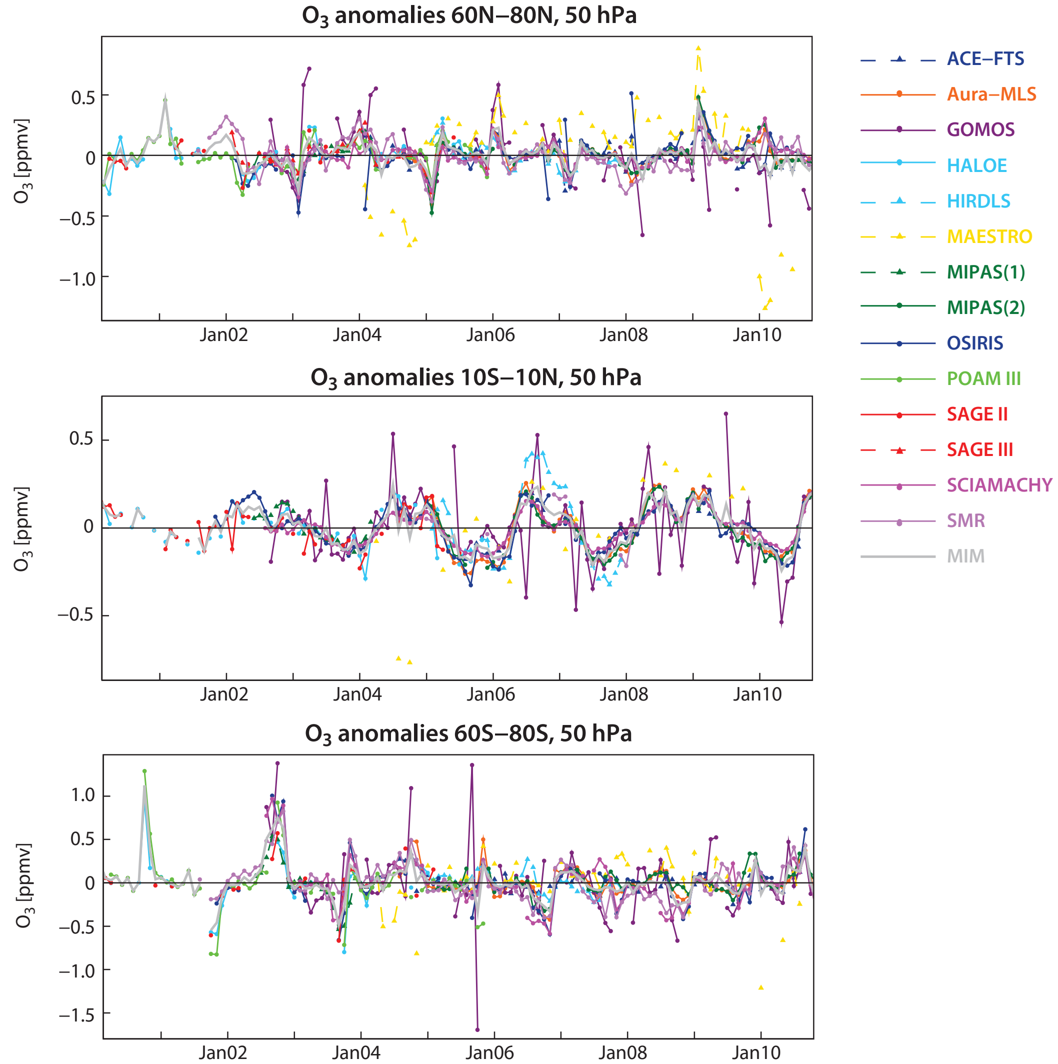


Figure 4.1.24: Time series of deseasonalised ozone anomalies at 50 hPa. *Deseasonalised ozone anomalies at 50 hPa between 60°N – 80°N (upper panel), 10°S – 10°N (middle panel) and 60°S – 80°S (lower panel).*

have a large influence on the zonal mean values depending on the winter. In general the signal of interannual variability from SCIAMACHY, SAGE III, POAM III and Aura-MLS agrees best in the NH MS while SAGE II, ACE-MAESTRO and HIRDLS show deviations. The interannual ozone variability in the SH polar region is again dominated by month-to-month fluctuations and like the signal in the NH, it is characterised by a seasonal cycle with a maximum in late winter/early spring. The interannual ozone variability at high latitudes of both hemispheres is for most of the year (late spring to autumn) smaller than the variability in the tropics. In the SH, some years show a strong signal in late winter/early spring including the large positive anomaly in SH winter 2002 that is related to the major warming of the SH polar vortex [*Krüger et al.*, 2005] and is resolved by all instruments. For positive anomalies found in spring (*e.g.*, 2002 and 2005) the spread between the instruments is considerably larger compared to other years.

The corresponding time series of ozone anomalies at 50 hPa are displayed in **Figure 4.1.24**. The variability of tropical ozone is again dominated by the QBO signal, with a small amplitude during the first 3 years. After 2003, the ozone-QBO signal exhibits stronger amplitudes that are well resolved by most instruments. HIRDLS slightly overestimates the amplitude in 2006/2007 compared to other datasets, perhaps resulting from its higher vertical resolution. Unrealistic spikes caused by large month-to-month fluctuations are apparent in the GOMOS time series. The signal-to-noise ratio in GOMOS measurements varies considerably from star to star. Occultations with low signal-to-noise ratio often lead to outliers in the profile dataset, and are excluded from the GOMOS climatologies based on specific profile inspecting filters and the median (instead of mean) average. Notwithstanding these precautions, outliers or spikes can still be detected in GOMOS climatological estimates when the number of measurements averaged is small.

In the NH polar region at 50 hPa, the late winter anomalies are now clearly the dominant signal. While for some winters the instruments agree rather well on sign and strength of the anomaly (*e.g.*, 2002, 2005), for other winters there is a large spread between the instruments, with disagreement not only on the amplitude but also on the sign of the anomalies (*e.g.*, 2008, 2010). For individual months, large

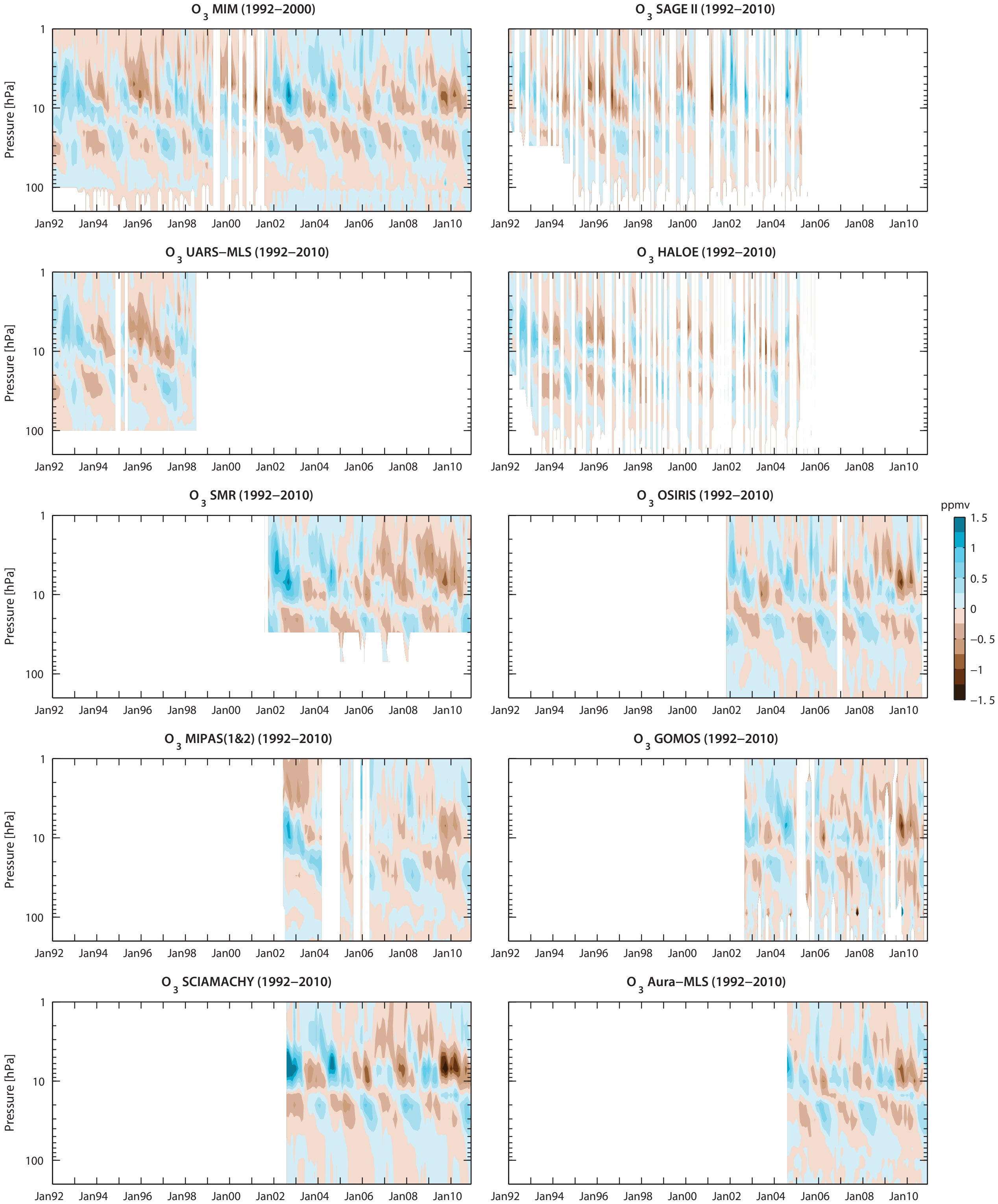


Figure 4.1.25: QBO signal for 1992-2010. *Altitude-time sections of deseasonalised ozone anomalies for 10°S-10°N from 1992 to 2010 are shown. The MIM is based on all displayed datasets.*

deviations of GOMOS and OSIRIS data can be observed. The ACE-MAESTRO dataset deviates from all other climatologies for most of the time period. Late winter ozone anomalies in the SH polar region at 50 hPa are dominated by spring variability, which can be small in some years but very pronounced in other years (*e.g.*, 2002 and 2003). For spring periods with large interannual anomalies also the spread between the instruments is larger.

4.1.5 O_3 evaluations: QBO

The Quasi-Biennial Oscillation (QBO) of the tropical zonal wind is one of the dominant influences on the interannual variability of equatorial ozone exhibiting a double peaked structure in the vertical with maxima in the lower (50-20 hPa) and middle/upper (10-2 hPa) stratosphere

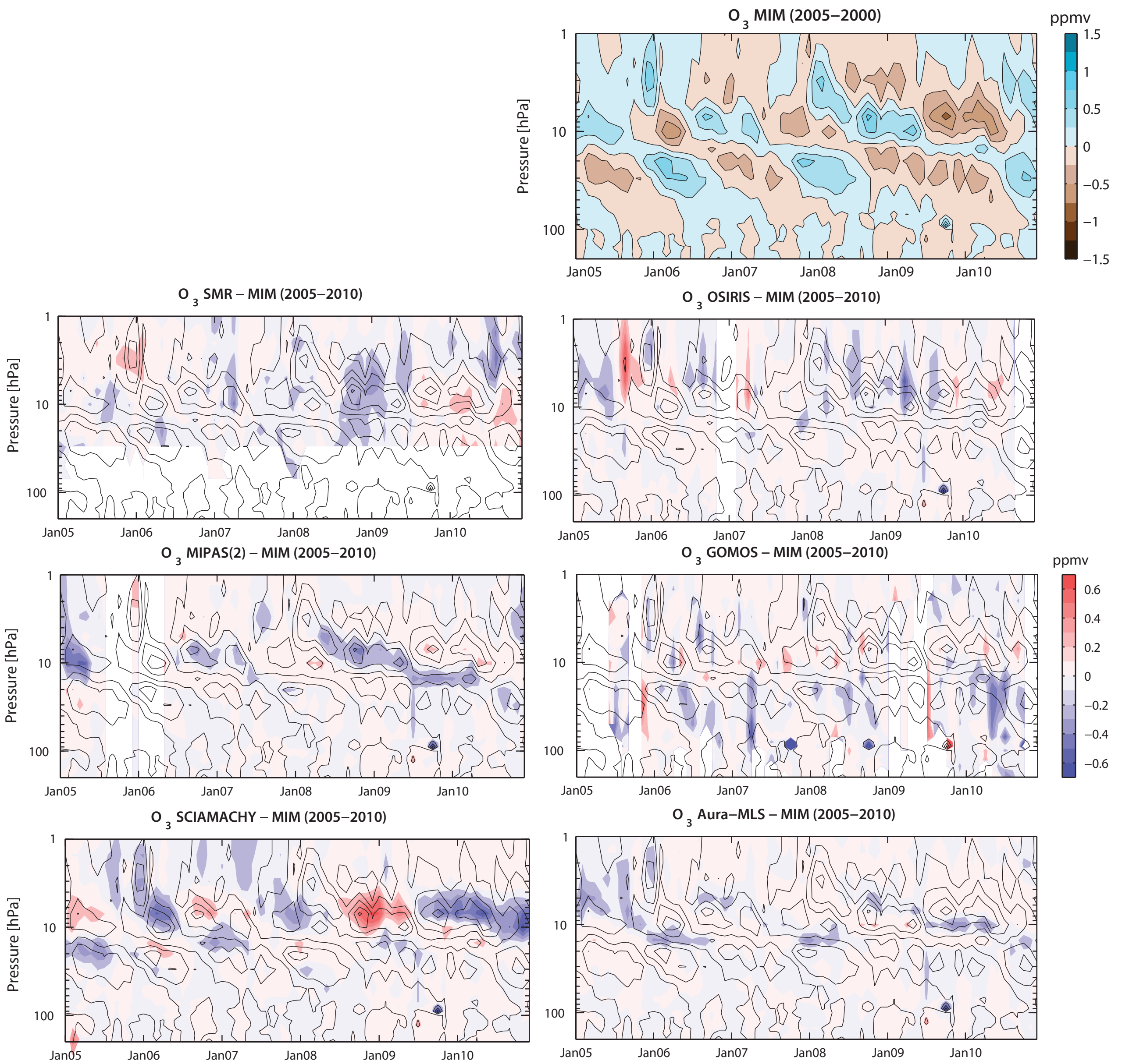


Figure 4.1.26: Differences with respect to MIM for QBO signal for 2005-2010. *Altitude-time section of MIM deseasonalised ozone anomalies for 10°S-10°N from 2005 to 2010 (upper left panel). Ozone anomaly differences between the individual datasets and the MIM are shown in the other panels by colour contours. The black contours present the MIM ozone anomalies from the upper left panel. The MIM is based on all displayed datasets.*

[*Zawodny and McCormick*, 1991; *Hasebe*, 1994]. Below 15 hPa, ozone is mainly under dynamical control and the QBO signal in lower stratospheric ozone results from the transport of ozone by the QBO-induced residual circulation. Above 15 hPa, on the other hand, ozone is under photochemical control and the QBO signal in middle/upper stratospheric ozone is thus understood to arise from QBO-induced temperature variations [*Ling and London*, 1986; *Zawodny and McCormick*, 1991] together with QBO-induced variability in the transport of NO_y which affects ozone chemically through NO_x [*Chipperfield et al.*, 1994]. A realistic characterisation of the altitude-time QBO structure by satellite measurements is an important aspect of the physical consistency of the dataset.

Figure 4.1.25 shows altitude-time sections of deseasonalised ozone anomalies from 1992 to 2010. Note that the tropical coverage from the SCISAT instruments is not sufficient for this analysis and therefore ACE-FTS and ACE-MAESTRO are not shown. While all instruments included in the QBO evaluation display the downward-propagating QBO ozone signal, there are some differences in the evolution and amplitude of the anomalies. One example of this disagreement is the negative ozone anomaly propagating downward from 1 to 10 hPa during 2002 and 2003. While MIPAS displays a strong amplitude for this negative signal, other instruments such as HALOE and OSIRIS observe a weak amplitude, and SMR and SCIAMACHY only detect the signal below 5 hPa. In order to analyse these deviations in more detail, the differences between each instrument's ozone anomalies and the MIM anomalies are calculated. Note that for a changing background ozone, an offset between the ozone anomalies will occur if the anomalies for the various instruments are calculated over different reference time periods. In order

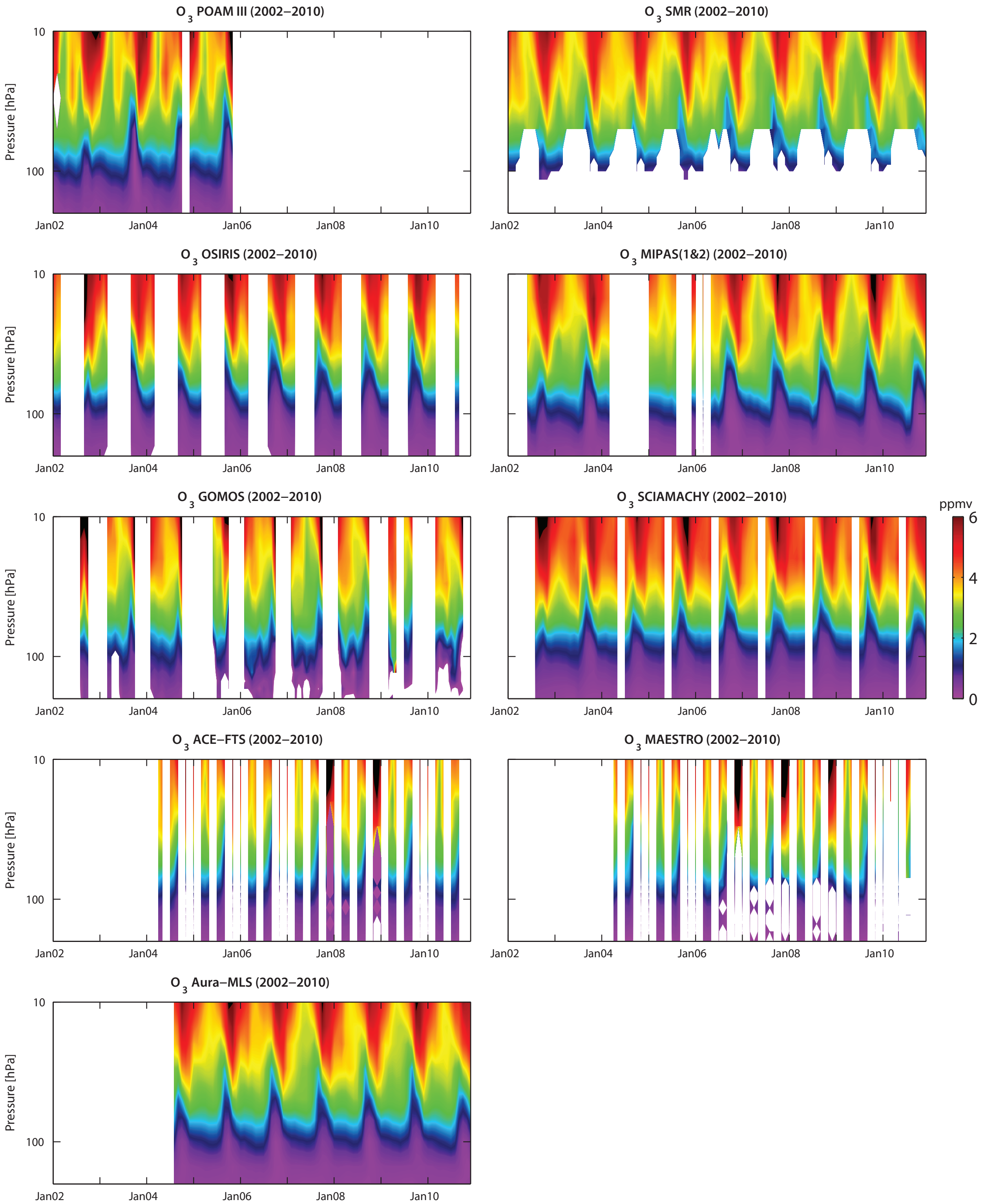


Figure 4.1.27: ***Antarctic ozone hole for 2002-2010.*** *Altitude-time sections of monthly zonal mean ozone for 60°S-90°S from 2002 to 2010 are shown.*

to avoid such an offset only one time period covered by a maximum number of instruments is chosen. The anomalies for six instruments are calculated and subtracted from the deseasonalised MIM for the time period 2005-2010. The differences between the instrument's and the MIM's anomalies are presented in **Figure 4.1.26** together with the contour lines of the MIM anomalies. Overall the anomalies agree well, with differences to the MIM often smaller than ±0.1 ppmv (corresponding to ±10%). Most of the instruments agree better below 15 hPa and show larger differences above 15 hPa where ozone is under photochemical control. Aura-MLS shows the best agreement with the MIM with strongest negative deviations during times of a QBO ozone phase shift from positive to negative anomalies. Deviations of GOMOS or OSIRIS to the MIM last only a few months and are independent of the QBO phase. In contrast, MIPAS(2) and SCIAMACHY deviations from the MIM last over longer time periods of up to 2 years while they propagate downwards in phase with the underlying QBO ozone signal. While MIPAS(2) under-estimates the positive QBO phase (2005 and 2008/09) compared to the MIM, SCIAMACHY shows the opposite behaviour, over-estimating the positive QBO ozone anomalies and under-estimating the negative anomalies (2006, 2009/2010).

4.1.6 O_3 evaluations: Antarctic ozone hole

Stratospheric ozone depletion in the Antarctic and Arctic regions through catalytic chemistry has been one of the major environmental issues of the last decades [*e.g.*, *Solomon*, 1999; *WMO*, 2014]. Ozone depletion in the polar lower stratosphere is linked to the activation of chlorine from its longer-lived reservoir species into reactive forms on the surfaces of polar stratospheric clouds at cold winter time temperatures [*Solomon et al.*, 1986; *Molina and Molina*, 1987]. In the Antarctic, reactive chlorine can be present for 4–5 months [*Waters et al.*, 1993; *Santee et al.*, 2003], during which time most of the ozone in the lower stratosphere is destroyed, resulting in reduction of total ozone by as much as two-thirds [*WMO*, 2011].

Figure 4.1.27 shows altitude-time sections of monthly zonal mean ozone averaged over 60°S-90°S (referred to as the polar cap average in the following) from 2002 to 2010. All instruments show the nearly complete removal of ozone in the lower stratosphere during Antarctic late winter/early spring. Usually, at the end of the year the ozone hole disappears as a result of the increasing polar stratospheric temperatures and the exchange of air between polar and lower latitudes. Severe differences exist in the vertical and temporal extent of the ozone hole as it is quantified by the polar cap averages from the different satellite instrument datasets. While POAM III polar cap averages show evidence of the ozone hole for only 1 to 2 months, polar cap averages for other instruments display longer periods of ozone reduction. Also visible in the ozone altitude-time section is the diabatic descent of air masses with higher ozone mixing ratios from the US during winter and spring.

Figure 4.1.28 displays the relative differences between the MIM and the individual instruments for the time evolution of the polar cap Antarctic ozone. The instruments show considerable disagreement, which is especially pronounced during the peak of the Antarctic ozone depletion when the mixing ratios are low (as indicated by the underlying MIM ozone field) and when temporal and spatial gradients are strongest. **Figures 4.1.29** and **4.1.30** show time series of the relative differences between the MIM and the individual instruments at 30, 50, 80 and 100 hPa for the two latitude bins 65°S-70°S and 80°S-85°S. The breakdown of the polar cap average into the individual latitude bins allows the quantification of how much of the large differences mentioned above are caused by spatial sampling effects (*i.e.*, for some instruments the polar cap average does not include all latitude bins), and allows the examination of those parts of the differences that are also present in the individual latitude bin comparisons. Note that additional sampling effects can result from non-uniformity in day-of-month sampling, which can cause differences for the individual latitude bins of up to ±20% and in some instances above 20% (see *Section 3.2.1* for a detailed discussion).

Reasonably good agreement is found between Aura-MLS, MIPAS(1/2) and OSIRIS with polar cap average differences from the MIM of up to ±20% (**Figure 4.1.28**). Aura-MLS (OSIRIS) observes mostly higher (lower) ozone values except during very short phases around the onset of the ozone hole. MIPAS(1/2) differences to the MIM are negative during the time of the ozone hole and positive during the rest of the year. These characteristics are generally confirmed by the comparisons performed for the individual latitude bins (**Figures 4.1.29** and **4.1.30**) with some exceptions found for individual cases. At the higher latitude bin (80°S-85°S), Aura-MLS shows at 100 hPa larger deviations to the MIM in the range of -50% while differences for the level above and below (with the latter not shown here) are in the range of ±20%. Some cases of larger deviations of up ±50% can also be found for OSIRIS between 30 to 80 hPa at 80°S-85°S.

ACE-FTS, ACE-MAESTRO, GOMOS, POAM III, SCIAMACHY, and SMR show a considerable disagreement with differences often up to ±50% and sometimes exceeding ±100%. ACE-FTS and ACE-MAESTRO do not sample at all latitudes in a given month and therefore the comparison of individual latitude bins is more representative than the polar cap average. For both instruments relative differences are enhanced during times of ozone depletion with large positive deviations found for the vortex inner latitude bins (80°S-85°S) and large negative deviations in the vortex outer latitude bins (65°S-70°S). POAM III and SCIAMACHY polar cap average differences to the MIM are linked to the seasonal cycle, with enhanced differences in winter and spring. POAM III observes more ozone than most other instruments (+20%) except during the peak of the ozone depletion at the end of winter when it under-estimates the ozone abundance (-50%). SCIAMACHY shows the opposite behaviour, with negative deviations during summer and autumn (-20%) and large positive deviations during the time of the ozone hole in late winter and spring (+50%). The detailed analysis for two latitude bins reveals that POAM III agrees reasonably well with the MIM in the outer vortex (differences up to ±20%) but shows large deviations in the inner vortex, which can be either positive or negative depending on the month and latitude bin. For SCIAMACHY, the deviations in the outer vortex area are mostly below ±50% but can be as large as ±100% in the inner vortex. GOMOS deviations from the MIM are not coupled with the seasonal cycle and the appearance of the ozone hole. While the polar-cap-averaged picture shows large deviations for GOMOS in all months, the evaluation of the individual latitude bins reveals that for the upper levels (above 80 hPa) this difference results from the averaging process and deviations are mostly within ±20%. However, for levels equal or lower than 80 hPa, deviations become very large, exceeding ±100%. SMR shows small deviations to the MIM during times with no ozone depletion (smaller

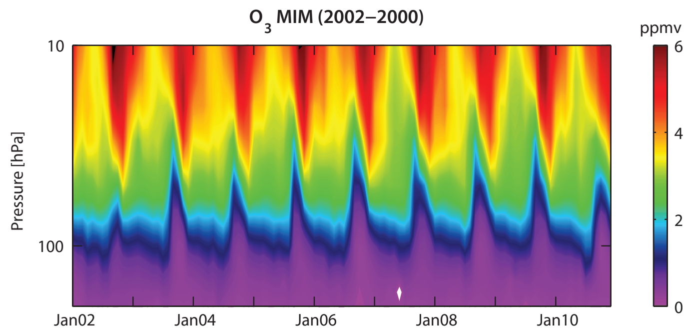
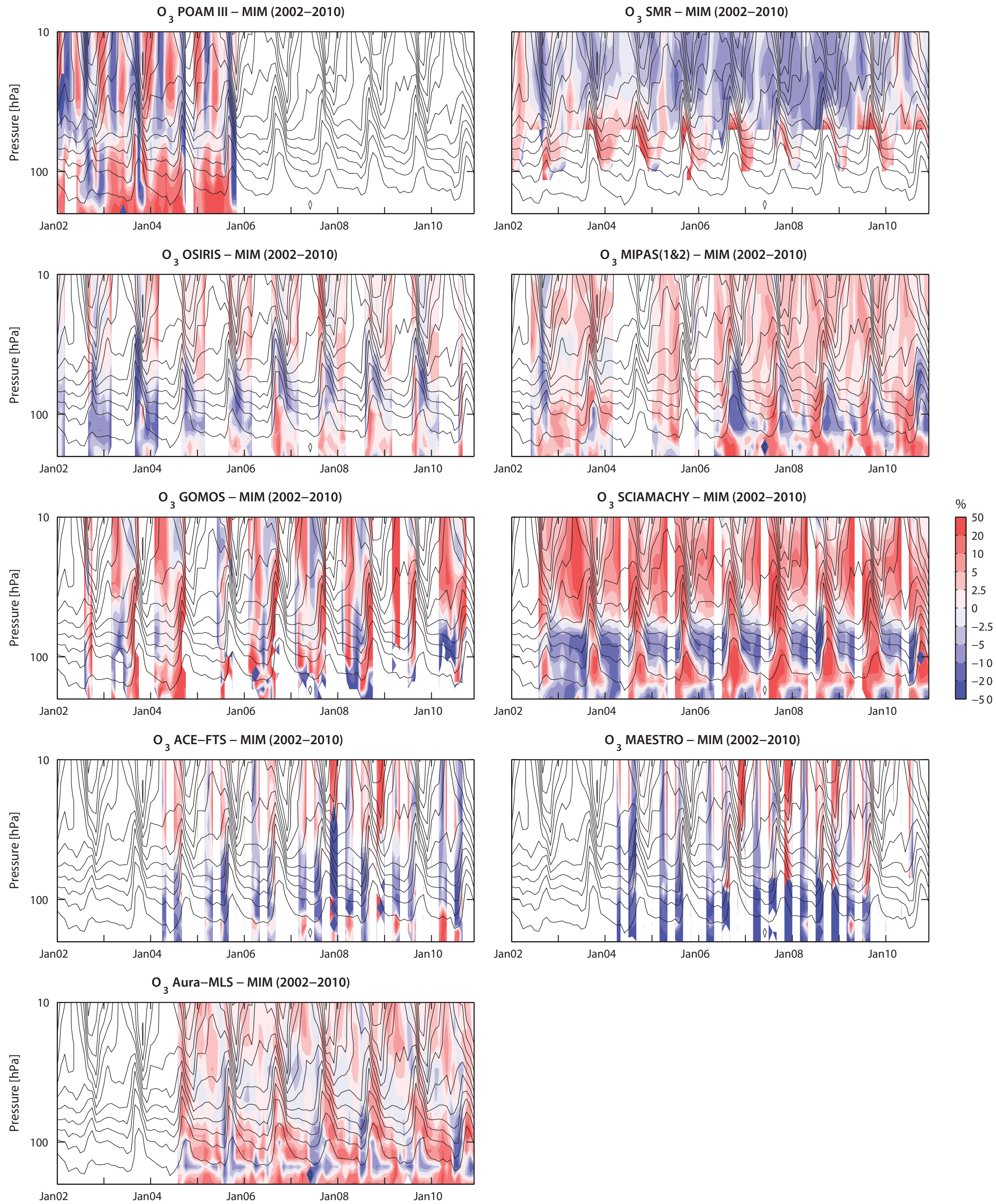


Figure 4.1.28: Differences with respect to MIM for Antarctic ozone hole for 2002-2010. *Altitude-time sections of MIM ozone (calculated based on displayed datasets) for 60°S-90°S from 2002 to 2010 is shown in upper left panel. Differences between the individual datasets and the MIM are shown in the other panels by colour contours. The black contours repeat the MIM ozone field from the upper left panel. Note that the MIM contours are shown for entire time series irrespective of the individual instrument's coverage.*



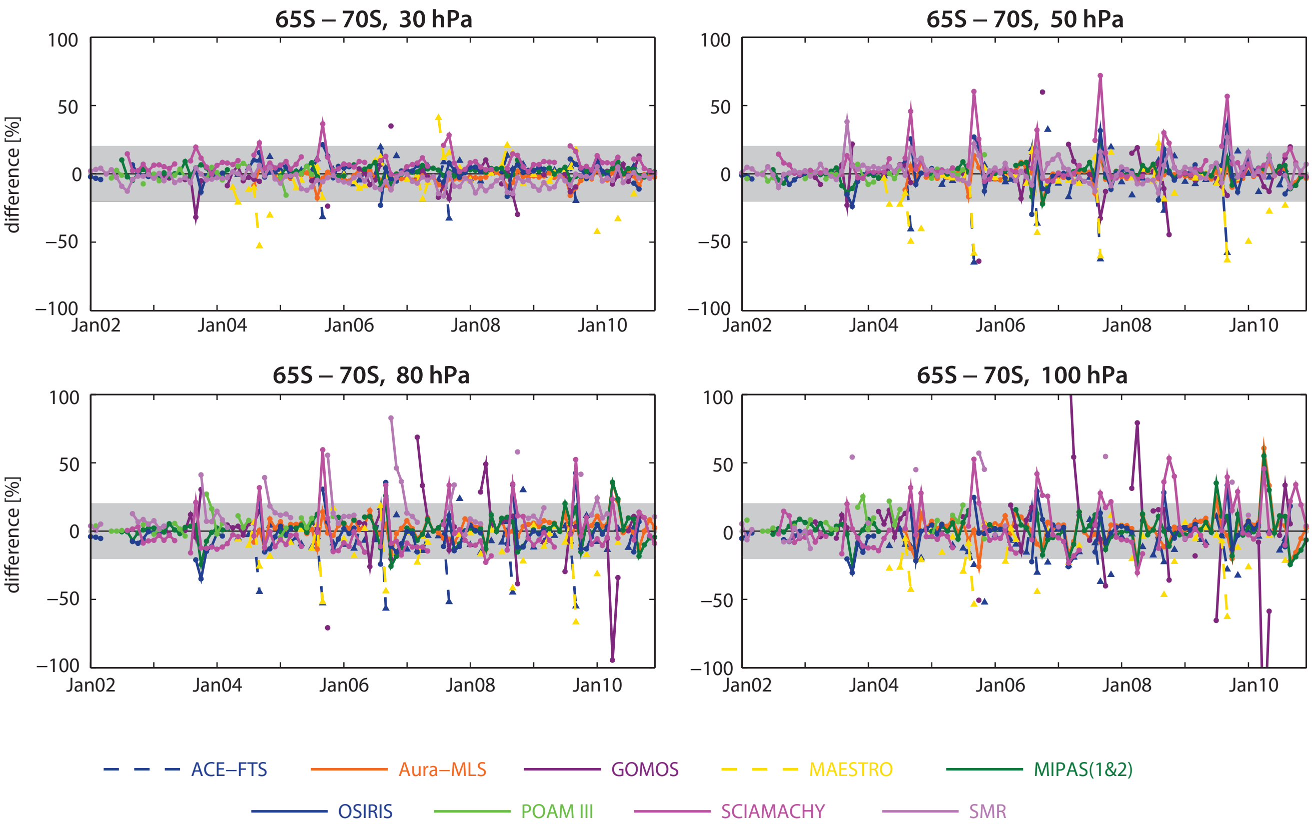


Figure 4.1.29:* *Time series of relative differences with respect to MIM for ozone at 65°S-70°S. *Time series of the relative differences between the individual instruments and the MIM at 30, 50, 80 and 100 hPa for 65°S-70°S are shown. The grey shaded area indicates where relative differences are smaller than ±20%.*

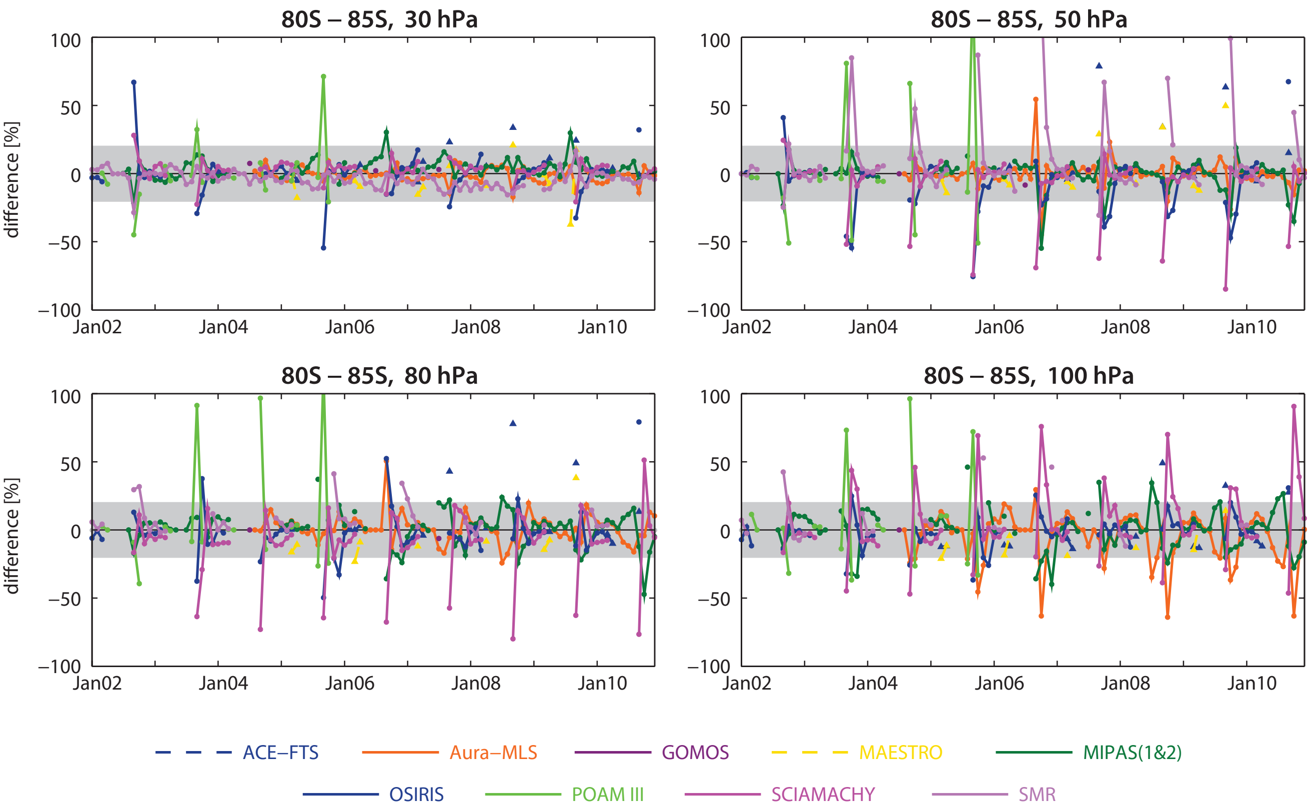


Figure 4.1.30:* *Time series of relative differences with respect to MIM for ozone at 80°S-85°S. *Time series of the relative differences between the individual instruments and the MIM at 30, 50, 80 and 100 hPa for 80°S-85°S are shown. The grey shaded area indicates where relative differences are smaller than ±20%.*

than ±20%) and large positive deviations during the Antarctic ozone hole (up to +100%).

For most of the instruments, the deviations from the MIM change sign during the springtime (during ozone depletion), and are opposite during the rest of the year. The polar-cap-average ozone deviations are influenced by the sampling patterns of the individual instruments and are in some cases (*e.g.*, GOMOS at levels above 80 hPa) larger than differences derived for individual latitude bands. Overall, however, deviations similar to the ones found for the polar-cap-average ozone field are apparent in 5° wide latitude bins that are completely inside the polar vortex over several months and therefore should be less impacted by spatial sampling effects. Note that the magnitude of the large relative differences observed during the time of severe ozone depletion is partially related to the low ozone abundance. However, the evaluation of the absolute difference time series also shows enhanced deviations during the time of the ozone hole (see **Figures A4.1.27-A4.1.28** in *Appendix A4*).

4.1.7 Summary and conclusions: O_3

A comprehensive comparison of 20 ozone profile climatologies from 18 satellite instruments (LIMS, SAGE I, SAGE II, UARS-MLS, HALOE, POAM II, POAM III, SMR, OSIRIS, SAGE III, MIPAS, GOMOS, SCIAMACHY, ACE-FTS, ACE-MAESTRO, Aura-MLS, HIRLDS, and SMILES) has been carried out. Overall findings on the systematic uncertainty in our knowledge of the ozone mean state and important characteristics of the individual datasets are presented in the following summary, including two synopsis plots. The first summary plot (**Figure 4.1.31**) provides information on the ozone mean state and its uncertainty derived from the spread between the datasets. The second summary plot (**Figure 4.1.32**) shows specific inter-instrument differences as deviations of the instrument climatologies to the MIM climatology. For each instrument and selected region, the deviation to the MIM is given as the median (mean) difference over all grid points in this region. Additionally, for each instrument the spread of the differences over all grid points in this region is presented. Note that both pieces of information (average deviation and spread) are important for a meaningful assessment of inter-instrument differences. See *Section 3.3.5* for more detailed information on the summary plots.

Atmospheric mean state

- The uncertainty in our knowledge of the atmospheric ozone annual mean state is smallest in the tropical MS and mid-latitude LS/MS. The evaluation of 13 datasets for the time period 2003-2008 reveals a 1σ multi-instrument spread in this region of less than ±5% (**Figure 4.1.31**, lower right panel).
- Maximum ozone mixing ratios are found in the tropical MS around 10 hPa. Here, the absolute values of the various climatologies show the largest spread for the tropical and extra-tropical stratosphere, with variations between 10 and 12 ppmv (**Figure 4.1.31**, left panel in the middle row).
- In the tropical LS, the spread between the datasets increases quickly with decreasing altitude, reaching ±30% at the tropical tropopause. In the mid-latitude LS, where the average ozone values are similar to those at the tropical tropopause, the various datasets show closer agreement regarding the ozone mean state, with a 1σ of ±10% (**Figure 4.1.31**, lower right panel).
- At polar latitudes, the climatologies give a larger spread of the ozone mean state (1σ of ±15%) compared to lower latitudes (1σ of ±5%). Maximum variations (up to 1σ of ±30%) are found in the Antarctic LS, resulting from large relative differences in the observations of the ozone hole (**Figure 4.1.31**, lower right panel).

Performance by region

Middle stratosphere (30-5 hPa)

The MS is characterised by the lowest spread between the datasets. In the tropical and mid-latitude MS, nearly all instruments show very good agreement with relative differences smaller than ±5% (**Figure 4.1.32**, second row). Exceptions are SMR in the tropics and mid-latitudes, with negative deviations to the MIM of around -5±2% (regional mean ± 1σ) and SCIAMACHY in the tropics with positive deviations of around +5±5%. Note that some datasets (*e.g.*, SCIAMACHY, ACE-FTS, SMILES) show relatively large standard deviations and MADs indicating a wider regional spread of the relative differences, while other instruments (*e.g.*, SMR, Aura-MLS) have small standard deviations and MADs indicating a narrow distribution of the relative differences around their mean. Such narrow distributions together with a small mean difference describe an excellent agreement (differences smaller than ±2.5%) of these datasets (*e.g.*, OSIRIS, GOMOS, Aura-MLS). In the polar regions, all instruments display larger relative differences compared to lower latitudes, with differences up to ±20% in the Antarctic and up to ±10% in the Arctic.

Lower stratosphere (100-30 hPa)

In the LS, there is a clear difference between the performance of the instruments in the tropics and mid-latitudes, with a much better agreement of the datasets in the mid-latitudes. Here average differences are mostly in the range of ±10% with the exception of SMILES displaying an average deviation of +15%. For some instruments a relatively wide regional spread (over all LS mid-latitude grid points) of the differences is found, indicating individual monthly mean differences larger than +20% for UARS-MLS, SMR, and GOMOS and smaller than -20% for GOMOS and SMILES.

In the tropics, the inter-instrument differences are considerably larger and instruments agree only reasonable well, with average differences mostly in the range of ±20% (HIRLDS up to +25%). For some instruments, a large

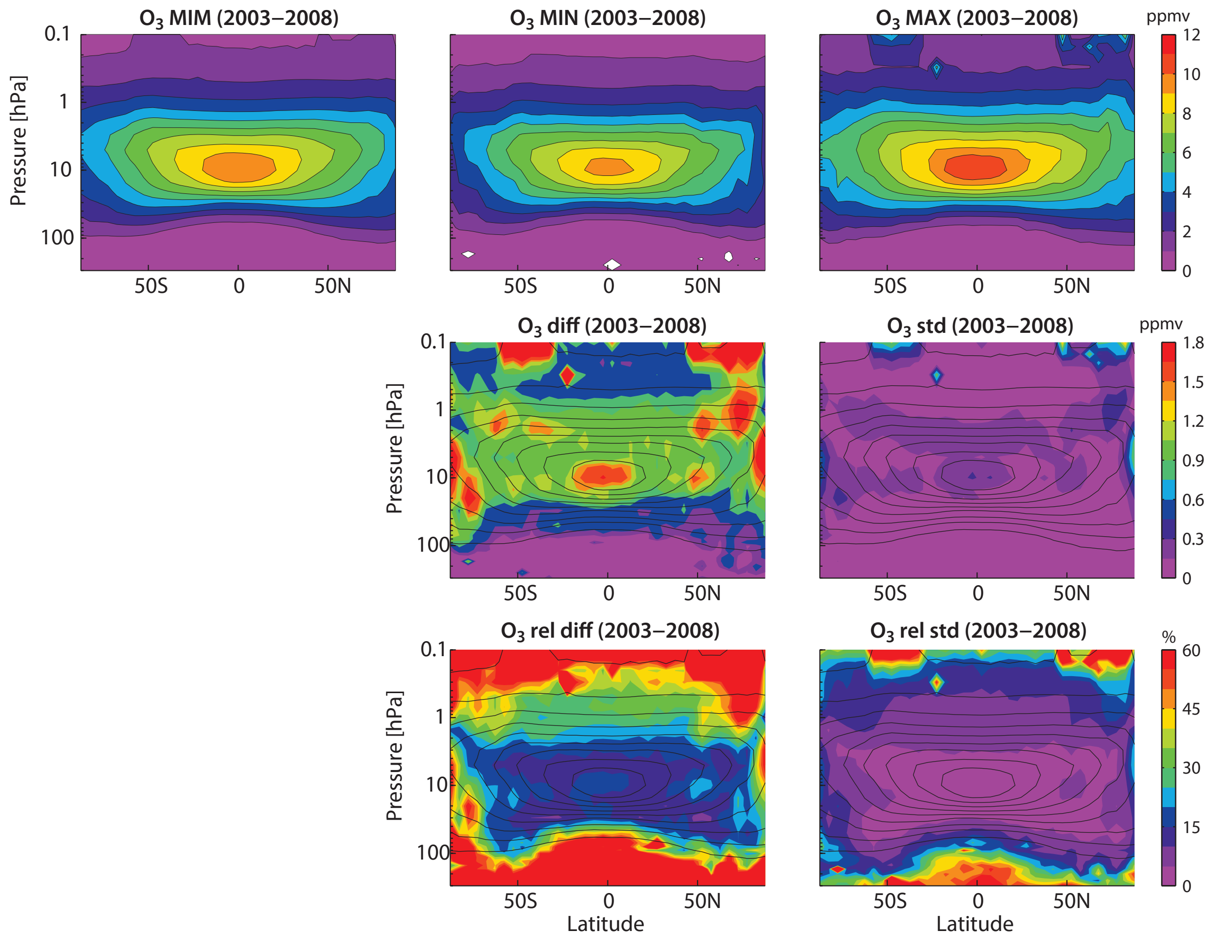


Figure 4.1.31: Summary of ozone annual zonal mean state for 2003-2008. *Shown are the annual zonal mean cross sections of the MIM, minimum (MIN), and maximum (MAX) ozone values in the upper row, the maximum differences over all instruments (MAX-MIN) and the standard deviation over all instruments in the middle row, and relative differences and relative standard deviations with respect to the MIM in the lower row. Black contour lines in the lower rows give the MIM distribution. Instruments included are SAGE II, HALOE, POAM III, SMR, OSIRIS, SAGE III, MIPAS, GOMOS, SCIAMACHY, ACE-FTS, ACE-MAESTRO, Aura-MLS, and HIRDLS.*

regional spread is found reaching values below -30% for SAGE II, HALOE and OSIRIS and well above +30% for UARS-MLS, SMR, GOMOS, and HIRDLS. The poor agreement of the mean values and the larger spread are related to the difficulties the instruments encounter when measuring the small ozone abundances in this altitude region where instrumental limitations (*e.g.*, resulting from cloud interference and high extinction) play a role. Note that SMR, MIPAS and Aura-MLS show excellent agreement, with differences to the MIM of less than ±5%. Furthermore, inter-instrument differences are less than 5% between the datasets from SCIAMACHY, ACE-FTS, ACE-MAESTRO and SMILES (mean deviations to the MIM of ~-10%) and between the datasets from SAGE II, HALOE and OSIRIS (mean deviations to the MIM of ~-20%). At high latitudes, differences are mostly in the range of ±30% for the SH and ±10% for the NH similar to the MS.

Upper troposphere/lower stratosphere (300-100 hPa)

Most instruments achieve good agreement in the mid-latitude UT (average differences up to ±10%) with two small exceptions (up to ±15% for HALOE and MIPAS(1)) and one evident outlier (-40% for ACE-MAESTRO). A large regional spread of up to ±75% exists for GOMOS, ACE-MAESTRO, and SAGE III. The good agreement observed at the mid-latitudes is not found in the tropics, where most instruments show differences of ±20% or larger. Maximum deviations are observed for HALOE, GOMOS and ACE-MAESTRO (with average differences of above ±60%). All datasets have a larger regional spread than in the mid-latitude UT with maximum values of well above +100% for GOMOS and below -100% for ACE-MAESTRO.

Upper stratosphere (5-1 hPa)

In the US, similar differences between the datasets exist in the tropics and at mid-latitudes. In both regions the datasets SAGE II, UARS-MLS, POAM III, OSIRIS, SAGE III, MIPAS(1), GOMOS, UARS-MLS, and HIRDLS agree very well, with average difference around ±5%. Datasets on the low side, with average deviations around -10%, are HALOE, SMR, and SMILES, while datasets on the high side with average deviations around +10% are MIPAS(2), SCIAMACHY, ACE-FTS, and ACE-MAESTRO.

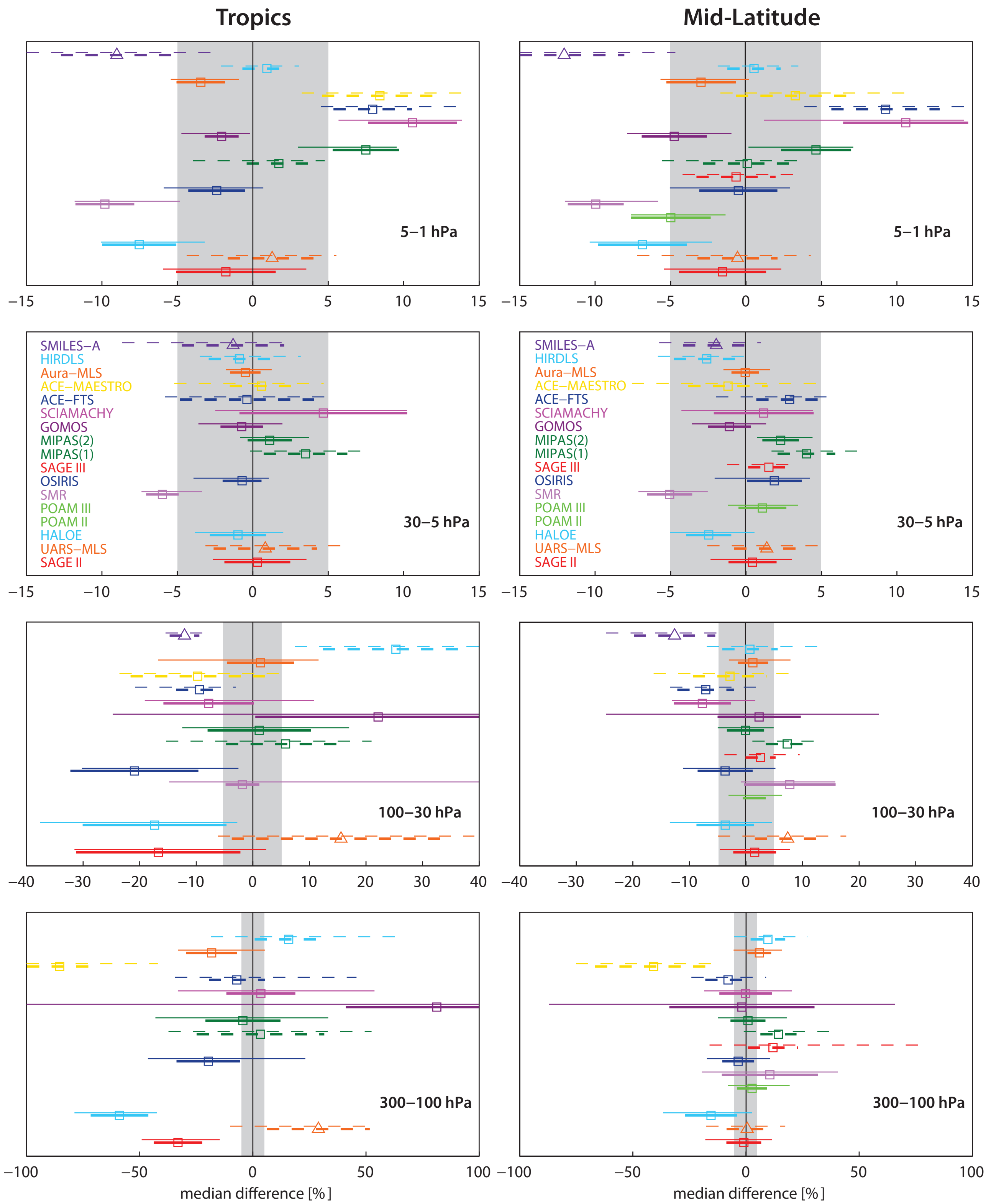


Figure 4.1.32: Summary plot of ozone differences for 2003-2008. *Over a given latitude and altitude region the median (squares), median absolute deviation (MAD, thick lines), and the standard deviation (thin lines) of the monthly mean relative differences between an individual instrument-climatology and the MIM are shown. Results are shown for the tropics (20°S-20°N) and mid-latitudes (30°S-60°S and 30°N-60°N) and for 4 different altitude regions from the UT to the US between 300 and 1 hPa for the reference period 2003-2008. Triangles indicate medians of instruments that are obtained outside of the reference period (UARS-MLS and SMILES-A), which are shown with respect to SAGE II and SMR based on comparison results for the time periods 1994-1996 and 2010, respectively.*

Lower mesosphere (1-0.1 hPa)

In the LM, the spread between the instruments increases with increasing altitude for decreasing ozone mixing ratios. The agreement is reasonably good at mid-latitudes, with relative differences around ±20%. In the tropics, inter-instrument differences are slightly larger (~±30%). The importance of the ozone diurnal cycle increases with altitude above 1 hPa and impedes a direct comparison between instruments measuring at different LSTs. Therefore, the inter-instrument differences mentioned above can not necessarily be considered as representative as the actual instrument offsets and are not shown in **Figure 4.1.32**.

Instrument-specific conclusions

LIMS and **SAGE I** provide the earliest ozone measurements and their climatologies agree very well in the MS, with differences mostly within ±2.5% for all latitude bands. In the LS, differences are larger (up to ±20%).

SAGE II provides the longest data record with climatological ozone values in the tropics and mid-latitudes being in the middle of the measurement range given by the spread of all climatologies. Exceptions are the tropical LS and UT, where SAGE II data shows too low values compared to the other datasets, which is consistent with the SAGE II low bias in this region with respect to ozonesondes [*Wang et al.*, 2002]. In the tropical and mid-latitude MS, GOMOS, and Aura-MLS climatologies show excellent agreement with the SAGE II climatology (differences below ±2.5%) while UARS-MLS, HALOE, OSIRIS, SAGE III, and MIPAS(2) agree very well with SAGE II with slightly larger differences (up to ±5%).

HALOE and **UARS-MLS** observation periods overlap with SAGE II from 1991 to 2005 and 1999, respectively. The HALOE ozone climatology is in general low compared to the other datasets. The negative deviations of the HALOE climatology to the MIM are small in the MS and mid-latitude LS (around -5%), larger but still in the climatological range in the US (-10%) and the tropical LS (-30%) and very large in the Antarctic UTLS in spring (-100%). The UARS-MLS climatology shows the opposite behaviour compared to that of HALOE, with positive deviations from the MIM.

POAM II, **POAM III,** and **SAGE III** mainly observe ozone at higher latitudes with a limited temporal coverage for some latitude bins which leads to larger biases in the annual means than in the monthly means. The SAGE III climatology agrees very well with most other datasets, with only small differences from the MIM with a narrow distribution. The POAM II climatology has a negative offset compared to other datasets which is particularly strong in the NH LS and SH UT. The POAM III climatology shows a positive offset compared to the MIM, which is small in the stratosphere (≤ 5%) and larger in the UT (~ 20%). Its sampling pattern allows POAM III to provide continuous solar occultation observations of the Antarctic ozone hole, where it reports more ozone than most other instruments (+20%) except during the peak of the ozone depletion at the end of SH winter when it under-estimates the ozone abundance (-50%).

Among the newer datasets **OSIRIS**, **GOMOS**, **Aura-MLS**, and **HIRDLS**, climatologies in the MS/US are consistent and show only small deviations (*e.g.*, average differences for the tropical MS of less than 1%). Aura-MLS performs exceptionally well in most regions, being in the middle of the range of all climatologies, and providing a realistic characterisation of ozone variability. While the other datasets perform also very well they have some limitations. OSIRIS data in the SH is impacted by its limited sampling pattern and shows somewhat larger differences from the MIM, as well as an unrealistic seasonal cycle with no amplitude in the UTLS. The GOMOS climatology shows considerable disagreement to all other datasets below 30 hPa, including an unrealistic seasonal cycle and unrealistic spikes in the deseasonalised time series. The HIRDLS climatology agrees well with the MIM in most atmospheric regions except in the tropical LS where it displays the strongest average deviation among all datasets of around +25%.

SMR and **SMILES** provide the lowest climatological ozone values in the stratosphere. While SMILES agrees very well with the other instruments in the MS, differences of up to -20% are found in the LS and US. The SMR climatology agrees well with the other climatologies in the UTLS. However, above 30 hPa it displays a negative offset which determines the lower boundary of the range of the climatological ozone data from all instruments. During Antarctic ozone hole events, SMR severely overestimates the ozone abundance by up to +100%.

The **ACE-FTS** and **ACE-MAESTRO** climatologies agree well with those of the other instruments in the LS and MS. Both datasets have a positive offset in the US (+10%) and ACE-MAESTRO has a strong negative offset in the UT (-50 to -100%). In general, the differences of the two instruments' climatologies with respect to the MIM show very similar structures, which are opposite to that of the OSIRIS, Aura-MLS and GOMOS climatologies. As a result of their limited temporal sampling, they show larger differences at higher latitudes than most other instruments.

The **SCIAMACHY** climatology agrees well with the other datasets in the UT and LS. However, in the tropical MS/US and mid-latitude US it shows in the early years a positive difference of up to +20% which might be related to the vertical resolution of the ECMWF temperature data used in the SCIAMACHY retrieval and climatology construction. The differences are smaller after 2006, with maximum differences of up to +10%. SCIAMACHY provides a physically consistent dataset but overestimates the QBO signal and the Antarctic ozone during the time of the ozone hole (+50%).

MIPAS measured with a different spectral and spatial resolution after 2005 and therefore provides two data products; MIPAS(1) and MIPAS(2). While the MIPAS(2) climatology shows mostly very small differences with respect to the MIM, the MIPAS(1) climatology has a positive offset up to 10% in the stratosphere and 20% in the troposphere. An exception to this classification is the US, where the MIPAS(1) climatology differences are smaller than ±5% and MIPAS(2) has a positive bias of around 10%. Due to the jump between the MIPAS datasets, analysis of time series from the complete MIPAS data requires a method that is immune against such discontinuities [*e.g.*, *von Clarmann et al.*, 2010].

4.1.8 Recommendations: O_3

- The evaluation of 20 ozone profile climatologies shows that our knowledge of the ozone atmospheric mean state is good in the tropical MS and in the midlatitude LS, MS, and US. However, a large climatological spread in the tropical UTLS demonstrates the need for further evaluation activities in this region including *in situ* measurements from balloon or aircraft platforms, as well as datasets from nadir sounders.
- Identified inter-instrument deviations in the LM are not necessarily representative for real climatological differences due to the growing importance of the ozone diurnal cycle above 1 hPa. A SPARC Data Initiative follow-on activity taking into account the effects of ozone variations with the diurnal cycle is recommended.
- Our findings show large inter-instrument differences for monthly zonal mean ozone at high latitudes (compared to tropics and mid-latitudes), which might be related to the different sampling patterns of the individual instruments. More detailed evaluations of high latitude ozone (especially for ozone hole conditions) will require the use of coincident measurement comparisons, polar vortex coordinates and the incorporation of *in situ* measurements.
- Nearly all instruments agree very well on the representation of ozone interannual variability and can be recommended for studies of climate variability. Note that some instruments show unrealistic spikes (month-to-month fluctuations) in some regions (*e.g.*, GOMOS and ACE-MAESTRO).
- SAGE II has been used extensively in validation and long-term studies and it is of interest to extend the time series through merging activities. As a result of their excellent agreement with SAGE II, the datasets from Aura-MLS, GOMOS (only in the tropical and mid-latitude MS), OSIRIS and MIPAS(2) (not above 10 hPa) are recommended for such merging activities.
- For future model-measurement comparison activities, the evaluations of natural variability presented here (seasonal cycle, QBO signal, and Antarctic ozone hole) are recommended. Depending on the evaluation, individual instruments should be excluded from the comparison. Caution should be used when evaluating the seasonal cycle in the tropical LS, which is seen to vary in magnitude between the different instrumental climatologies, probably due to the different vertical resolutions of the instruments and the large vertical gradient of O_3 in this region. A further comparison with ozonesonde measurements is recommended, possibly as part of a SPARC Data Initiative follow-on activity with a focus on the UTLS.

4.2 Water vapour – H_2O

Water vapour (H_2O) is a key greenhouse gas in the atmosphere, and changes in its abundance impact radiative forcing most effectively in the UTLS where strong gradients across the tropopause region are found [*e.g.*, *Gettelman et al.*, 2011]. H_2O is also a key constituent in atmospheric chemistry. It is the source of the cleansing agent of the atmosphere, hydroxyl (OH, see *Section 4.22*), which controls the lifetime of shorter-lived pollutants, tropospheric and stratospheric ozone, and other longer-lived greenhouse gases such as CH_4 [*Seinfeld and Pandis*, 2006]. Furthermore, its presence in the stratosphere has an important influence on stratospheric chemistry through its ability to form ice, thereby offering a surface for heterogeneous chemical reactions, which are involved in the destruction of stratospheric ozone [*Solomon*, 1999]. Accurate knowledge of the distribution and trends of H_2O from the UT up to the mesosphere is therefore crucial for understanding climate and chemical forcings. However, it is not trivial to accurately measure H_2O, and satellite measurements, as well as *in situ* correlative data, have been shown to exhibit large relative differences [*SPARC,* 2000]. In particular, the current lack of an accepted standard from *in situ* correlative data is preventing the community from coming to a conclusive assessment of the performance of available satellite H_2O measurements (see Weinstock et al. [2009]). It is not possible to determine the 'best' instrument for measuring H_2O in the stratosphere. Instead, the results presented here are intended to give an overview of the spread and relative differences between the available satellite measurements, and to determine whether and where the datasets show physically consistent behaviour. The results presented here are summarised in *Hegglin et al.* [2013]. WAVAS II - the second phase of the SPARC water vapour activity - is preparing a complementary study of H_2O based on the classical validation approach using coincident profiles, and includes comparisons with *in situ* correlative measurements.

4.2.1 Availability of H_2O measurements

The first vertically resolved H_2O satellite measurements were provided by LIMS in 1978-1979. The longest dataset is available from the SAGE II instrument, which measured H_2O between 1984 and 2005. However, due to a channel shift and its correction in the SAGE II V6.2 data, which may have impacted the spatio-temporal consistency in the retrievals, SAGE II V6.2 H_2O should only be used with caution for trend studies [*Thomason et al.*, 2004]. Also, note that SAGE II H_2O data exhibit a known high bias above 3 hPa [*cf.*, *SPARC*, 2000]; data above this level were not included in the SAGE II monthly zonal mean climatologies here. This bias has been attributed to the decreasing H_2O

Table 4.2.1: Available H_2O data records between 1978 and 2010 from limb-sounding satellite instruments participating in the SPARC Data Initiative. *The red filling of the grid boxes indicates the temporal (January to December) and vertical coverage (300 to 0.1 hPa) of the respective instrument in a given year.*

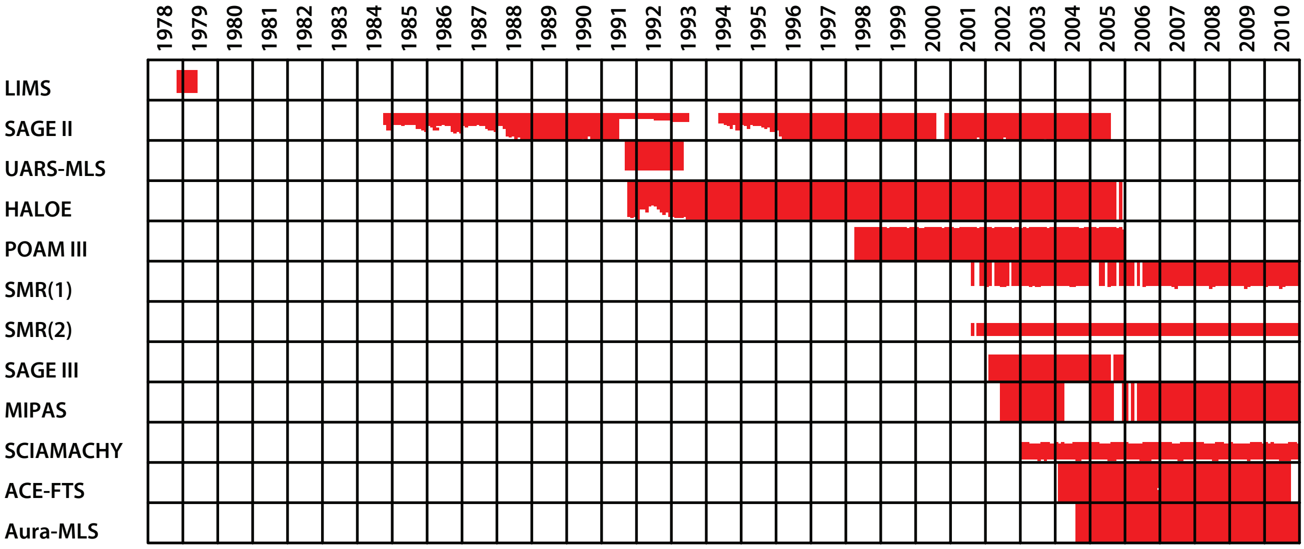


Table 4.2.2: Data version, time period, vertical range, vertical resolution, references and other comments for H_2O profile measurements used to generate the SPARC Data Initiative monthly zonal mean climatologies.

Instrument	Time period	Vertical Range	Vertical resolution	References	Additional comments
LIMS V6.0	Nov 78 – May 79	cloud top – 1 hPa	3.7 km	*Remsberg et al.*, 2009	
SAGE II V6.2	Oct 84 – Aug 05	cloud top – 50 km < 25 km > 30 km	1 – 2.5 km ~ 1 km ~ 2.5 km	*Thomason et al.*, 2004 *Taha et al.*, 2004	Data above 3 hPa are excluded / use for trend studies not recommended
UARS-MLS V6	Oct 91 – Mar 93	~ 18 – 50 km > 50 km	3 – 4 km 5 – 7 km	*Pumphrey*, 1999	H_2O stops early due to radiometer failure
HALOE V19	Oct 91 – Nov 05	cloud top – 90 km	2.5 km	*Grooß and Russell*, 2005	Data below tropopause are excluded
SAGE III V4.0	May 02 – Dec 05	cloud top – 50 km	~1.5 km	*Thomason et al.*, 2010	Only solar products are used here
POAM III V4.0	Apr 98 – Dec 05	5 – 45 km	1 – 2 km	*Lumpe et al.*, 2006	
SMR SMR(2) V2.0 SMR(1) V2.1	Jul 01 –	16 – 75 km 16 – 20 km 20 – 75 km	 3 – 4 km ~3 km	 *Urban*, 2008 *Urban et al.*, 2007	 544 GHz-band 489 GHz-band
MIPAS MIPAS(1) V13 MIPAS(2) V220	 Mar 02 – Mar 04 Jan 05 – Apr 12	 cloud top – 70 km cloud top – 70 km	 4.5 – 6.5 km 2.5 – 6.9 km	 *Milz et al.*, 2005 *Milz et al.*, 2009 *von Clarmann et al.*, 2009a	Measurement mode switched in 2005 from high spectral to high vertical resolution
SCIAMACHY V3.0	Sep 02 – Apr 12	11 – 25 km	3 – 5 km	*Rozanov et al.*, 2011b	New data product
ACE-FTS V2.2	Mar 04 –	5 – 89 km	3 – 4 km	*Carleer et al.*, 2008 *Hegglin et al.*, 2008	
Aura-MLS V3.3	Aug 04 –	316 – 100 hPa 100 – 0.2 hPa < ~0.1 hPa	2 – 3 km 3 – 4 km 6 – 12 km	*Read et al.*, 2007 *Lambert et al.*, 2007 *Livesey et al.*, 2011	

signal at higher altitudes, and the small contribution of H_2O to the total slant path optical depth [*Taha et al.*, 2004]. Hitherto, the most frequently used dataset for water vapour trend analyses is therefore from HALOE, which measured H_2O between 1991 and 2005. However, a newer version of SAGE II, V7.0 [*Damadeo et al.*, 2013], which became available after the finalisation of this chapter improves on some of the issues SAGE II V6.2 exhibited. The V7.0 dataset was shown to yield promising results in data merging activities [*Hegglin et al.*, 2014]. From the early 2000's onwards, H_2O measurements became available from a whole suite of new satellite instruments.

Tables 4.2.1 and **4.2.2** contain information on the H_2O data products available to the SPARC Data Initiative, including time period, height range, vertical resolution, and references. Note that MIPAS measured with different spectral and spatial resolution before and after 2005, and the data products, evaluated sperarately, are denoted MIPAS(1) and MIPAS(2), respectively. SMR provides two H_2O data products derived from two different bands at 489 GHz (here named SMR(1)) and 544 GHz (here named SMR(2)), which yield data above and below ~20 km, respectively.

Due to a lack of available resources, observations available from SAMS on Nimbus 7 [*Jones et al.*, 1986], ISAMS [*Taylor et al.*, 1993] and CLAES [*Roche et al.*, 1993] on UARS, ATMOS [*Gunson et al.*, 1996] and MAS [*Hartmann et al.*, 1996] on the ATLAS Space Shuttle missions, and ILAS on ADEOS [*Sasano et al.*, 1999] could not be included in this report.

4.2.2 H_2O evaluations: Zonal mean cross sections, vertical and meridional profiles

In this section, monthly or annual zonal mean cross sections are analysed to investigate differences between the various datasets. Both annual and monthly means have been averaged over multiple years as indicated in the section headings. The time periods have been chosen so that a maximum number of instruments can be compared in each case. In addition, vertical and meridional profiles are shown for individual months in order to focus on particular height/latitude regions and to determine if differences between datasets are persistent over the entire year. In addition to the absolute values, differences between individual instruments and the multi-instrument mean (MIM, see *Section 3.3.1* for definition) are presented. Note the MIM is not intended as a "best" climatology; rather its use is motivated by the need for a reference that does not favour a certain instrument. The differences with respect to the MIM reflect not only instrument errors, but also incomplete monthly or latitudinal data coverage, which impact the calculated annual or zonal means to some extent. Note, sampling affects the water vapour annual and monthly averages much less than for ozone, mostly in the region below 100 hPa where dynamical varaiblity is strongest (*Section 3.2.1*; also *Toohey et al.* [2013]). Where not shown in the main evaluations, monthly zonal mean cross sections can be found in *Appendix A4.2*.

LIMS (1978-1979) and SAGE II (1984-1990)

LIMS provides the earliest available H_2O measurements from space. Here, we compare with SAGE II monthly zonal mean fields since these measurements are closest in time, and LIMS does not have enough data to produce an annual mean climatology. The evaluation is done for those months during which LIMS and SAGE II have the most overlap in latitudinal coverage. Note that we do not account for possible trends between the chosen time periods or the influence of the solar cycle on H_2O in the LM [*Nedoluah et al.*, 2009]. *Hurst et al.* [2011] show that trends calculated from balloon-borne H_2O measurements near 20 hPa are small, and that the evolution of H_2O during the 1985-1990 period, when SAGE II is measuring, is relatively stable.

Figure 4.2.1a shows monthly zonal mean H_2O fields for LIMS and SAGE II. The figure reveals the key features of the H_2O distribution in the middle atmosphere, which results from transport by the Brewer-Dobson circulation and a stratospheric source of H_2O. Air entering the stratosphere is dehydrated at the cold tropical tropopause, creating a minimum in H_2O just above the tropopause. As the air ascends to higher altitudes, H_2O concentration is increased through the oxidation of methane [*Bates and Nicolet*, 1950]. Isentropic mixing between the ascending branch of the Brewer-Dobson circulation in the tropics (with low H_2O values) and the descending branch in the extra-tropics (with high H_2O values) produces the typical downward-sloping H_2O isopleths. Dehydration in the cold winter polar vortex can lead to an additional minimum in observed H_2O at high latitudes in the lower stratosphere.

Figure 4.2.1b reveals quantitatively that LIMS and SAGE II show very good to excellent agreement in the tropics (within ±2.5-5% of the MIM, corresponding to inter-instrument differences of 5-10%), and for the most part, agree well in the extra-tropics (within ±5 to ±10% of the MIM, or 10-20% inter-instrument differences), even though the satellite measurements do not overlap in time. Generally, SAGE II is somewhat lower (higher) than LIMS below (above) 10 hPa. LIMS measurements exhibit atypical isopleths that do not slope down strongly enough into the mid-latitudes (**Figure 4.2.1a**). As a consequence, the differences from the MIM increase at higher latitudes, with LIMS showing positive deviations from the MIM. Validation of LIMS H_2O V6.0 with a limited number of available correlative profile measurements at mid-latitudes, confirm that LIMS between 10 and 70 hPa is higher by about 10-15%, although within the stated measurement uncertainties of the respective instruments [*Remsberg*, 2009]. Below 80-100 hPa, the differences from the MIM increase to over ±20% across all latitudes, with SAGE II showing negative and LIMS showing positive deviations.

SAGE II, UARS-MLS, and HALOE (1991-1993)

Figure 4.2.2 shows cross sections of annual zonal mean H_2O for SAGE II, UARS-MLS, and HALOE averaged over

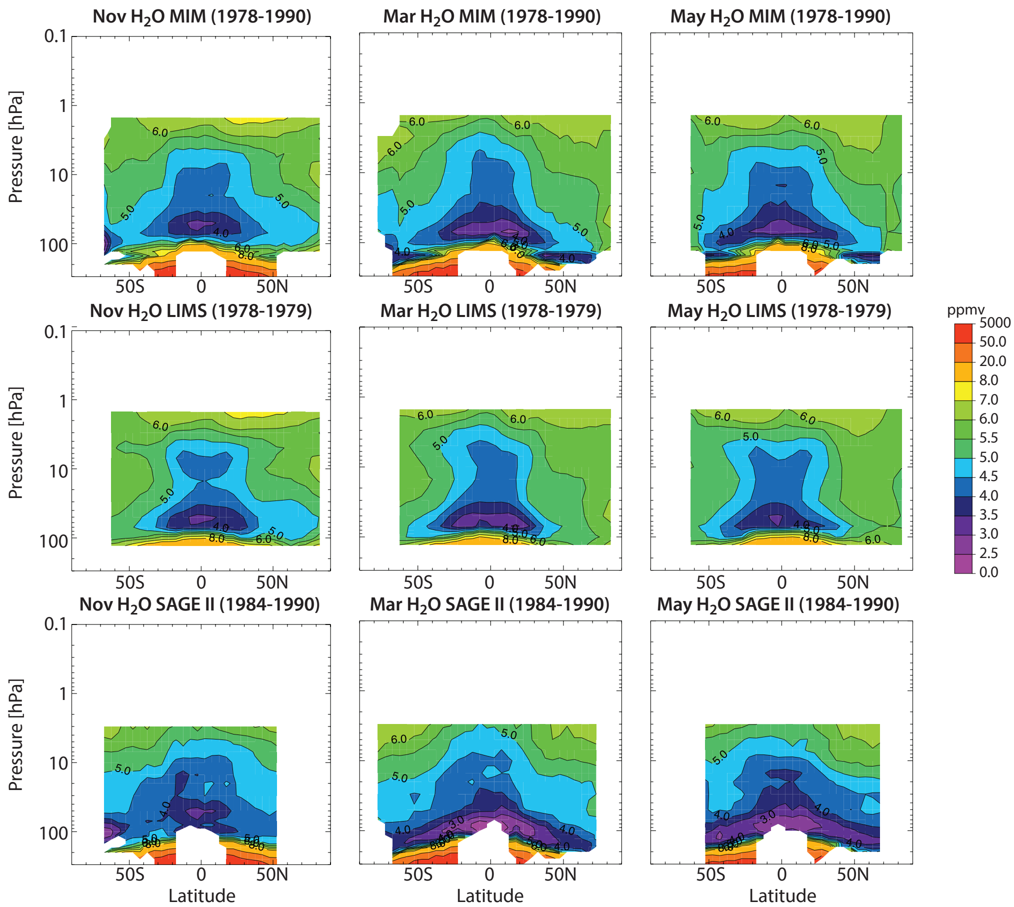


Figure 4.2.1a: ***Cross sections of monthly zonal mean H_2O for LIMS (1978-1979) and SAGE II (1984-1990).*** *November, March and May monthly zonal means (from left to right) are shown for the MIM, LIMS, and SAGE II (from top to bottom).*

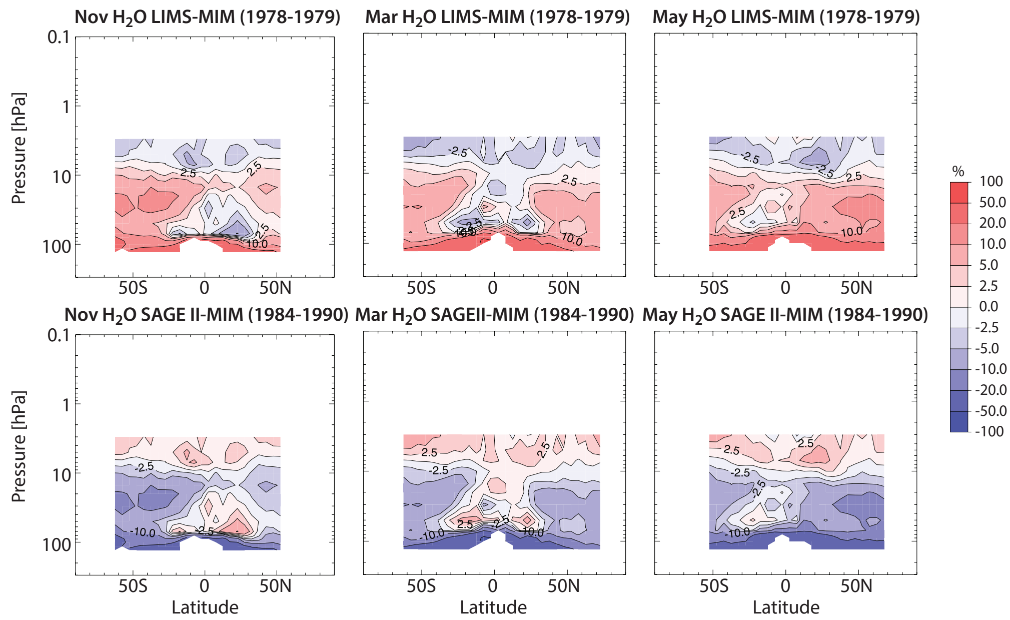


Figure 4.2.1b: ***Cross sections of monthly zonal mean H_2O differences relative to the MIM for LIMS (1978-1979) and SAGE II (1984-1990).*** *Same order as in Figure 4.2.1a.*

the time period 1991-1993, together with their relative differences from the MIM. Note that this time period is not an ideal choice for comparison due to the eruption of Mt. Pinatubo, which brought additional aerosol into the stratosphere adversely affecting the retrievals of solar occultation measurements. Therefore, the inter-instrument differences derived for this time period may not be consistent with differences derived for later time periods. However, it is the only time period that allows direct comparison with measurements from the UARS-MLS instrument.

The relative differences from the MIM are considered small, with values between ±2.5% and ±5% throughout most of the MS, US, and LM indicating excellent to very good agreement between the instruments. HALOE values generally lie between the (lower) UARS-MLS values and the (higher) SAGE II values. *Pumphrey* [1999] showed that the UARS-MLS H_2O data version used here (called prototype version 0104 at that time) yielded uniformly drier values than HALOE (by 0.1 to 0.4 ppmv), and values ~0.6 ppmv drier than the ATMOS measurements obtained from the

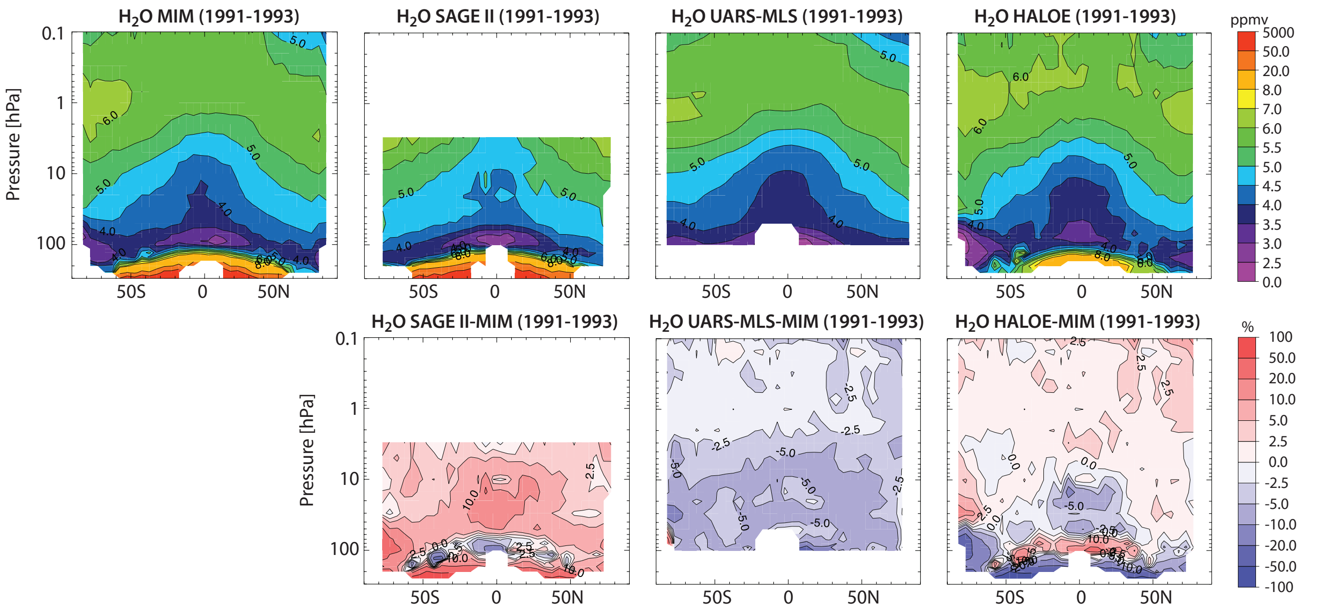


***Figure 4.2.2: Cross sections of annual zonal mean H_2O and differences for 1991-1993.** Shown from left to right are the MIM, SAGE II, UARS-MLS, and HALOE. Upper panels show absolute values, lower panels the differences relative to the MIM.*

Space Shuttle, but compared well to the average of 16 coincident frost point hygrometer profiles. In the UTLS, where UARS-MLS is only available above 100 hPa, SAGE II and HALOE show reasonably good agreement, with increasing differences below 100 hPa especially in the tropics and the SH polar region (around ±20% from the MIM), with HALOE on the low side of the MIM. An interesting feature is the 'sandwiched' layer near the tropical tropopause in the SAGE II and HALOE cross sections, with differences of opposite sign from the values above and below this layer. This indicates that the instruments' measurements do not agree on the mean pressure level of minimum tropical H_2O values in the LS. The effect could be due to the impact of heavy aerosol loading after the eruption of Mt. Pinatubo, the different vertical resolutions of the instruments, or an altitude registration error. More likely it is the result of the temporal sampling of the two instruments: due to the Mt. Pinatubo eruption in June 1991, SAGE II data is limited to the winter months of 1991 (and hence samples smaller H_2O values due to a higher and colder tropopause during these months), while HALOE samples the region during all months of 1993 (*cf.*, **Table 4.2.1**). Indeed, **Figure A4.2.1b** in *Appendix A4* confirms that the feature is not present in the monthly mean evaluations. The sampling issue is also seen to disappear when comparing HALOE and SAGE II in later time periods with better temporal coverage (see next section).

Figure 4.2.3 shows meridional profiles for four different pressure levels for March averaged over 1991-1993. At 1 hPa, UARS-MLS and HALOE show very good agreement, with differences from each other that are smaller than ±5%. At 10, 50, and 80 hPa, UARS-MLS, HALOE, and SAGE II agree well (mostly within ±10%), with UARS-MLS generally on the low side of the other two instruments. The climatological profile of SAGE II is noisier than the other two instruments [*cf.*, *Taha et al.*, 2004], as expressed in the larger SEM values for SAGE II, and shows a mostly positive offset of 10-15% from the MIM.

Figure 4.2.4 shows vertical profiles of H_2O concentration and their differences from the MIM at selected latitudes for April. Focusing on this time period reveals that UARS-MLS and HALOE agree to within 3% above 10 hPa at all latitudes. SAGE II and HALOE also show excellent agreement (within 5%) in the extra-tropical MS and LS above 100 hPa, with UARS-MLS on the low side. However, in the tropics around 20-30 hPa, HALOE exhibits even lower values of H_2O than UARS-MLS, causing the differences between SAGE II and HALOE of up to 30% in this month. In the UTLS, SAGE II and HALOE profiles diverge, with relative differences from the MIM of up to ±40% indicating considerable disagreement, with HALOE on the low side.

SAGE II and HALOE (1996-1998 versus 2002-2004)

Figure 4.2.5a and **b** show cross sections of annual zonal mean H_2O and relative differences to the MIM for SAGE II and HALOE for the years 1996-1998 and 2002-2004. While the two instruments cannot be regarded as totally independent (the correction of the measurement channel shift in SAGE II was based on HALOE measurements), a comparison of the two time periods 1996-1998 and 2002-2004 may indicate any potential drift in one of the instruments.

The comparison reveals that the two instruments show similar overall structures in the H_2O distribution, but with some obvious differences. In particular, HALOE seems to underestimate H_2O mixing ratios in the extra-tropical UTLS below 150 hPa, showing weaker gradients in H_2O across the tropopause than SAGE II. On the other hand, the two instruments agree on a large drop in H_2O at the tropical tropopause (around 100 hPa) between the early and the later period, which is consistent with the findings of *Randel et al.* [2006].

Throughout the MS, the differences relative to the MIM are very similar in both time periods, with values generally

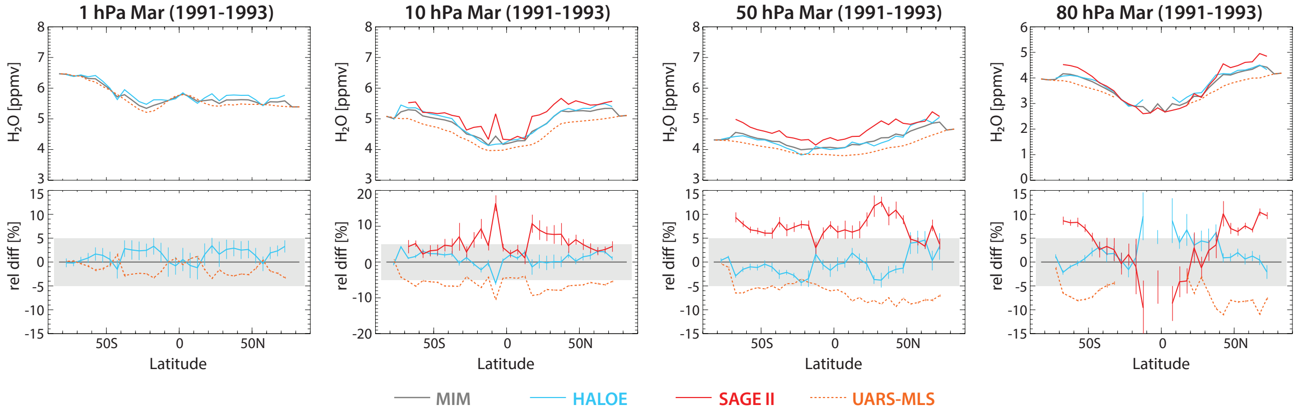


Figure 4.2.3: ***Meridional profiles of monthly zonal mean H_2O for 1991-1993.*** *Shown are meridional profiles for March at 1, 10, 50, and 80 hPa (from left to right). Upper panels show absolute values, lower panels relative differences between the individual instruments (SAGE II, HALOE, and UARS-MLS) and the MIM, respectively. The grey shading indicates where the relative differences are smaller than ±5%. Error bars indicate the uncertainty in the relative differences based on the SEM of each instrument's climatology.*

smaller than ±2.5% (or inter-instrument differences of only 5%) showing excellent agreement between the two instruments. Here, SAGE II (HALOE) is on the low (high) side. The differences increase towards the tropical tropopause to ±5% (equivalent to inter-instrument differences of 10%) and into the extra-tropical UTLS below 100 hPa, where the two instruments show differences of up to ±50% from the MIM. Here, SAGE II (HALOE) is on the high (low) side, findings that are consistent with the study by *Taha et al.* [2004]. Differences are also larger in the SH polar region. As noted in *Section 3.2.1*, sampling biases in the solar occultation measurements may explain more than 10% of the differences between the two instruments in these regions. Temporal sampling biases introduced by

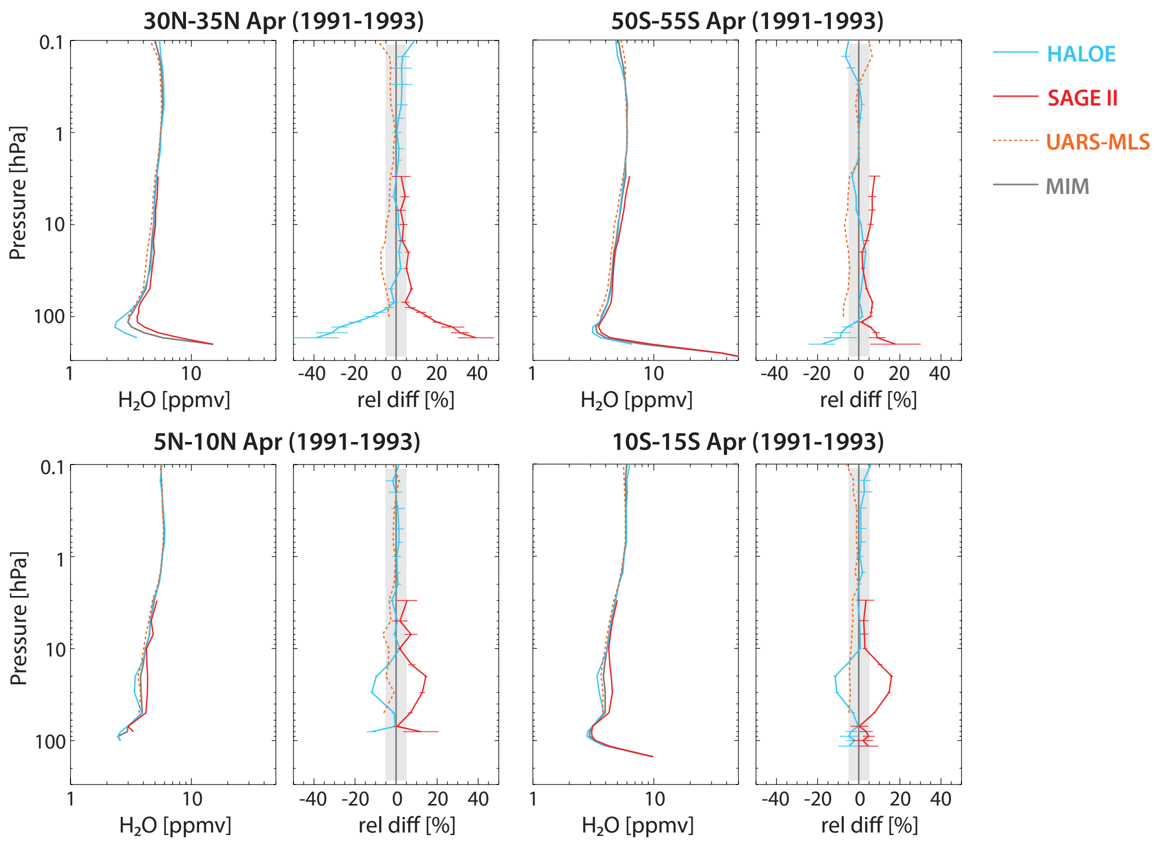


Figure 4.2.4: ***Vertical profiles of monthly zonal mean H_2O for 1991-1993.*** *The H_2O profiles are shown for 30°N-35°N and 50°S-55°S (upper panels), and 5°N-10°N and 10°S-15°S (lower panels) for April. The relative differences between the individual instruments (SAGE II, HALOE, and UARS-MLS) and the MIM are shown on the right of each H_2O profile panel. Error bars indicate the uncertainty in the relative differences based on the SEM of each instrument. The grey shaded area indicates where relative differences are smaller than ±5%.*

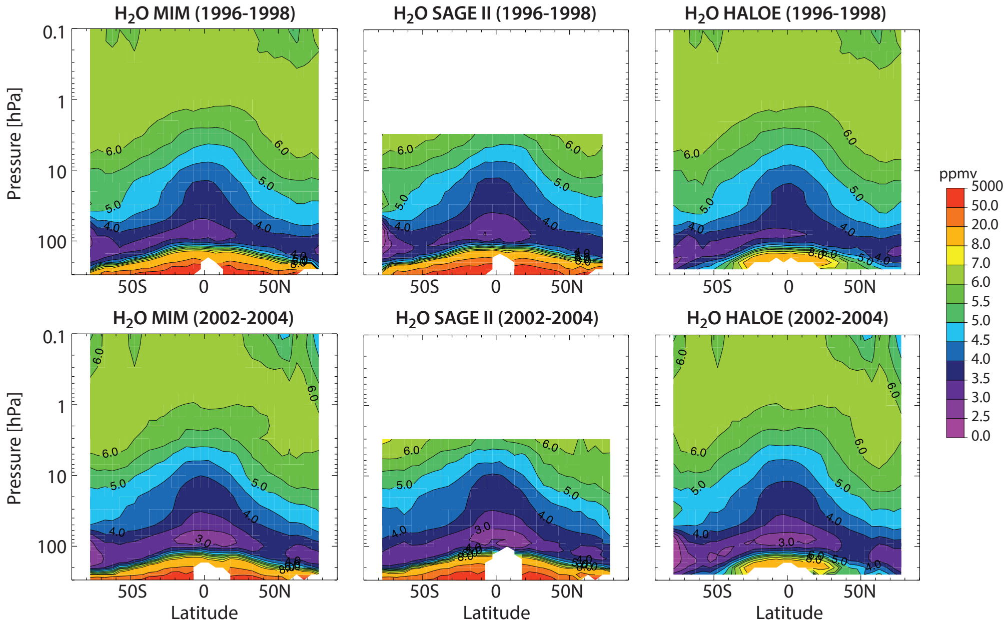


Figure 4.2.5a: ***Cross sections of annual zonal mean H_2O for 1996-1998 and 2002-2004.*** *Shown from left to right are the MIM, SAGE II, and HALOE. The upper (lower) panels show the climatologies for the earlier (later) time period.*

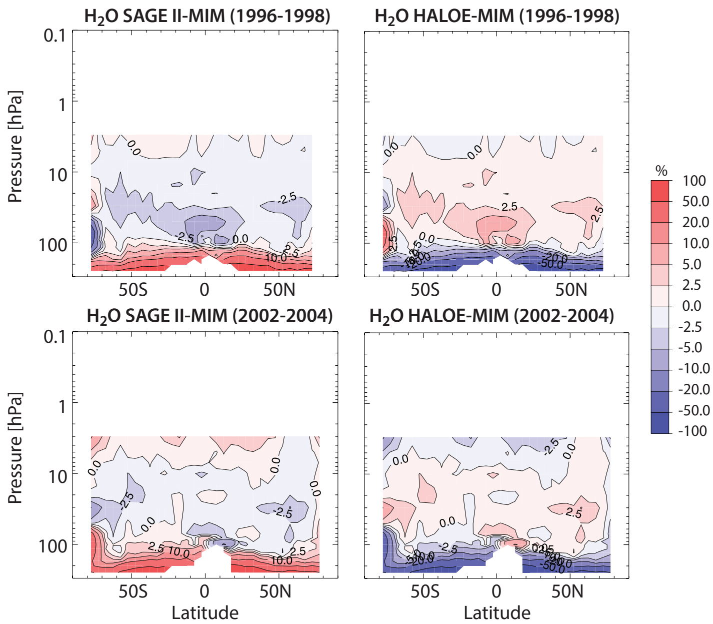


Figure 4.2.5b: Cross sections of annual zonal mean H_2O differences for 1996-1998 and 2002-2004. *Shown are the relative differences for each instrument (SAGE II and HALOE) with respect to the MIM. Same ordering as in Figure 4.2.5a.*

less frequent measurements towards the end of the missions may also be the reason for inter-instrument differences, which increase slightly in the US but decrease in the LS from the earlier to the later time period.

Note that the differences in 1996-1998 and 2002-2004 are of reversed sign in the tropical LS compared to the early 1990s (**Figure 4.2.2**). As discussed earlier, this is most likely the result of enhanced stratospheric aerosol after the Mt. Pinatubo eruption, affecting the retrievals. Also, there is no 'sandwiched' layer as seen in the differences around the tropical tropopause in the early 1990s, supporting the explanation that this issue is attributable to the particular temporal sampling.

Figure 4.2.6 contrasts meridional profiles between the two time periods for different months. The profiles show that the monthly evaluation can sometimes reveal larger discrepancies between the instruments than seen in the annual zonal mean evaluation. For example, at the 5, 10, and 200 hPa pressure levels for January, July, and October, respectively, the differences remain similar between the two time periods. At 80 hPa in the tropical LS in April on the other hand, the differences from the MIM decrease from ±10% to an average of ±2.5% (corresponding to inter-instrument differences of 20% and 5%, respectively). However, evaluation of the 80 hPa level during other months reveals that this decrease is not a consistent feature (not shown).

Figure 4.2.7 shows the vertical profiles in different seasons at subtropical and extra-tropical latitudes, confirming the mostly excellent agreement between the two instruments in the stratosphere in these regions. However, the differences increase strongly below 100 hPa. Only minor changes in the differences are found in between the two time periods at these latitudes.

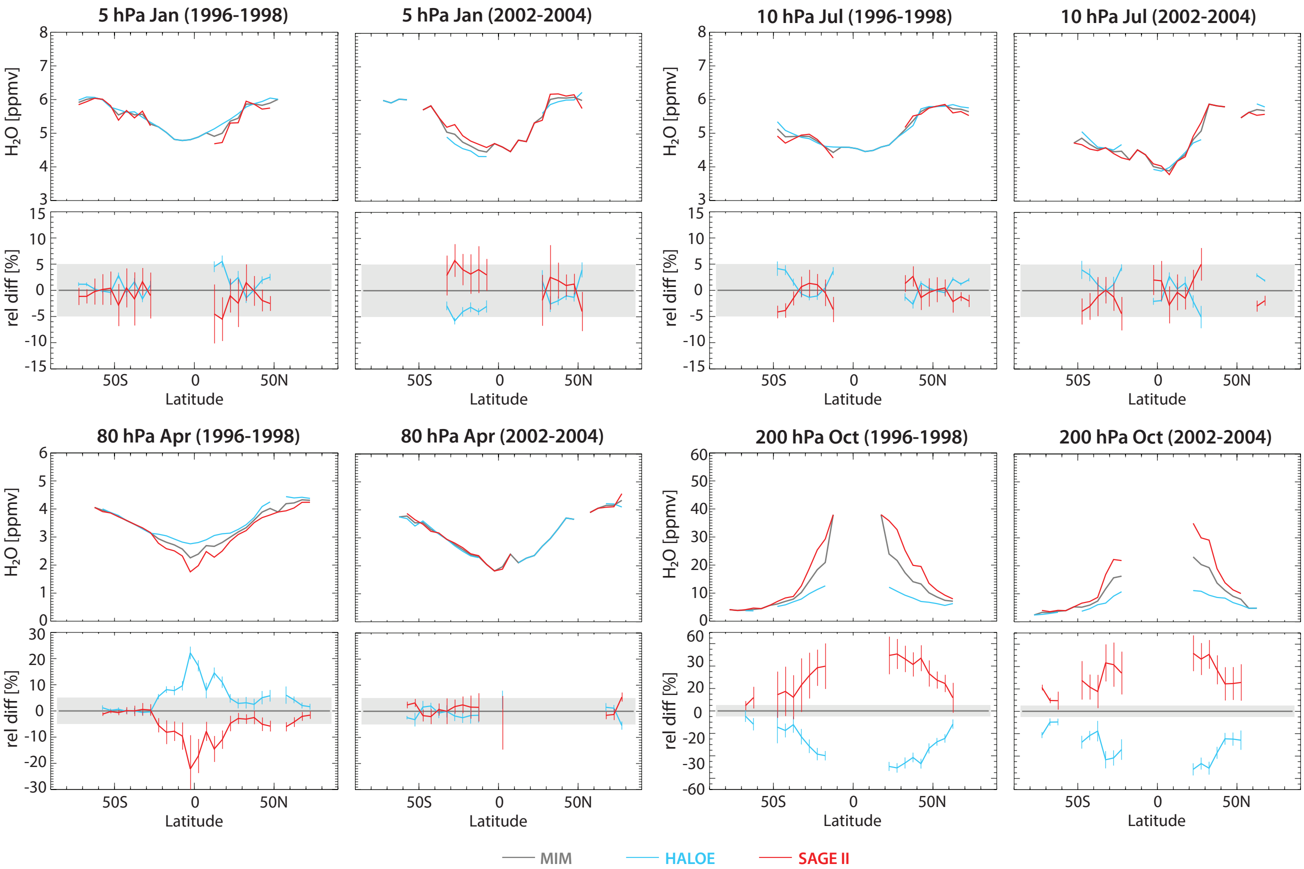


Figure 4.2.6: Meridional profiles of monthly zonal mean H_2O for 1996-1998 versus 2002-2004. *Meridional profiles are shown at 5 hPa for January, 10 hPa for July, 80 hPa for April, and 200 hPa for October (from upper left to lower right) for the two time periods. Upper panels show absolute values, lower panels relative differences between the individual instruments (SAGE II and HALOE) and the MIM, respectively. The grey shading indicates where the relative differences are smaller than ±5% from the MIM. Error bars indicate the uncertainty in the relative differences based on the SEM of each instrument.*

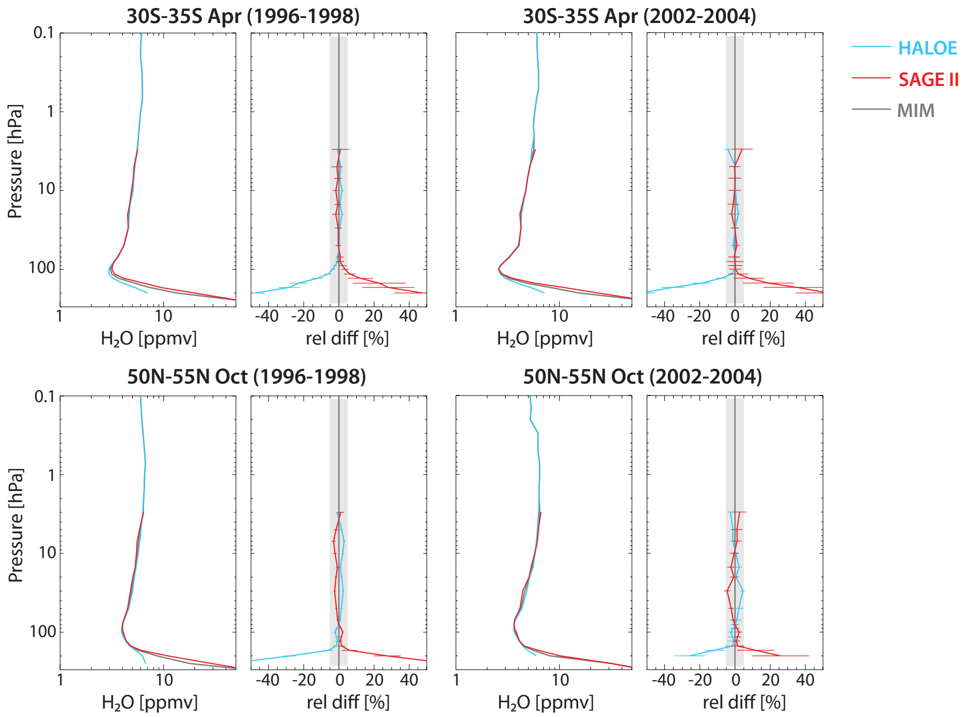


Figure 4.2.7: Vertical profiles of monthly zonal mean H_2O for 1996-1998 versus 2002-2004. *The H_2O profiles and their relative differences from the MIM are shown for April 30°S-35°S and October 50°N-55°N for the two time periods, respectively. Error bars indicate the uncertainty in the relative differences based on the SEM of each instrument. The grey shaded area indicates where relative differences are smaller than ±5%.*

SAGE II, HALOE, POAM III, SMR(1,2), SAGE III, MIPAS(1), and SCIAMACHY (2003)

Figures 4.2.8a and **b** show the annual zonal mean and relative difference cross sections for the year 2003. This period includes seven instruments, with MIPAS(1) measuring in the high spectral resolution mode (see **Tables 4.2.1** and **4.2.2**).

The annual zonal mean MIM shows the key features from Antarctic dehydration, a minimum in mixing ratios above the tropical tropopause, and a maximum in the H_2O values in the USLM. Note that the MIM does not include SMR(2), because of a large bias in the data. The instruments mostly capture the features found in the H_2O distribution, however with rather large inter-instrument differences in the absolute values as detailed below.

SAGE II, HALOE, and SMR(1) are on the low side of the MIM throughout most of the atmosphere (except SMR(1) in the tropical MS). POAM III, MIPAS(1) and SAGE III on the other hand are on the high side of the MIM. SMR(2) shows an unrealistically flat structure of the zonal mean H_2O (mixing ratio) isopleths in the UTLS, with a large positive deviation from the MIM below and a large negative deviation above 100 hPa. A low bias at these altitudes in SMR(2) has also been found by *Urban* [2008] and *Urban et al.* [2012] in comparisons with MIPAS(1), Aura-MLS, and ACE-FTS. SCIAMACHY shows very good agreement with the MIM in the extra-tropical LS, however it shows increasing positive deviations from the MIM of greater than +20% towards the tropopause region.

The differences between SAGE II and POAM III are consistent with the results from the validation exercise using coincident measurements by *Taha et al.* [2004] showing SAGE II with a low bias compared to POAM III, which is somewhat stronger in the SH (around 15%) than in the NH (around 10%). The same study pointed out the differences between HALOE and SAGE II, with SAGE II exhibiting somewhat lower values than HALOE throughout the MS (by about 5%), but reversed behaviour in the UTLS with HALOE showing much lower values than SAGE II. These findings are also consistent with our evaluations of these two instruments in the early 1990s. *Thomason et al.* [2010] also validated SAGE III in comparison with these instruments using coincidences, highlighting the excellent agreement (within 5%) with POAM III, and positive differences of 10-15% compared to HALOE and SAGE II.

Figure 4.2.9 shows the meridional profile comparison for 2003. At 0.5 hPa, only HALOE, SMR(1), MIPAS(1), and POAM III (although very limited) provide data. SMR(1) exhibits the lowest values (with a difference of -20% with respect to the MIM). HALOE is close to the MIM, and MIPAS(1) exhibits the highest values (with a difference of around +10% from the MIM). At 10 hPa, all instruments show very good agreement, within ±5% except for POAM III, which shows a positive deviation from the MIM of about 10%. At 80 hPa, MIPAS(1), SAGE II, SAGE III, HALOE, and SCIAMACHY all agree within about 10-15% in the extra-tropics. SCIAMACHY shows larger positive deviations from the MIM of up to 20% during October than April. SMR(2) shows large negative deviations from the MIM of 20-40% across all latitudes. At 200 hPa, inter-instrument differences increase to up to 100%. MIPAS(1), POAM III, SAGE II and SAGE III agree within about 30-40%, with HALOE being much lower than the other instruments. SCIAMACHY shows a somewhat noisier meridional profile at this level with largest positive deviations from the MIM of up to 30-50%.

Figure 4.2.10 shows the vertical profile comparisons for 2003. Most instruments lie within a range of about ±20% relative difference from the MIM through most of the atmosphere. The instruments agree best in the MS at 10 hPa, with relative differences from the MIM of ±5-8%. An exception is the UTLS, where relative differences from the MIM increase strongly, to up to ±40% and more. SMR(2) shows the largest negative deviations from the MIM above and the largest positive deviations below 100 hPa. HALOE shows large negative deviations from the MIM below 100 hPa. The large discrepancies in the UTLS between the instruments are partly caused by strong dynamical variability and large gradients in this region. As discussed in *Section 3.2.1,* the resulting sampling biases can be larger than ±10%. Another contributing factor may be the different altitude resolutions of the instruments.

In the USLM, SMR(1) exhibits the lowest and MIPAS(1) the highest values, with average differences of ±20%, and HALOE values lie approximately in the middle. A comparison for Southern Hemisphere high latitudes also includes POAM III. This instrument exhibits the highest values throughout the stratosphere, and shows large negative deviations from the MIM in the UTLS. The next section will discuss the issues identified here in greater detail.

SAGE II, HALOE, POAM III, SAGE III, SMR(1,2), MIPAS(1,2), SCIAMACHY, ACE-FTS, Aura-MLS (1998-2008)

Figures 4.2.11a and **b** show the annual zonal mean and relative difference cross sections for climatologies obtained over the years 1998-2008. Despite the fact that the climatologies of the individual instruments span different time periods (as indicated in the figure titles), this approach has been chosen in order to be able to compare a maximum number of instruments, and to limit the influence of reduced sampling by HALOE and SAGE II in the early 2000s. The comparison results for the 1998-2008 time period are consistent with results obtained from single-year evaluations such as the one presented for 2003, or an evaluation performed for instruments covering the years 2006-2009 only (not shown), providing confidence that trends in H_2O over this time period are not large enough to impact the comparison. Note that the evaluation of the 1998-2008 climatologies will be used as the basis for the summary plots in the conclusion *Section 4.2.8*.

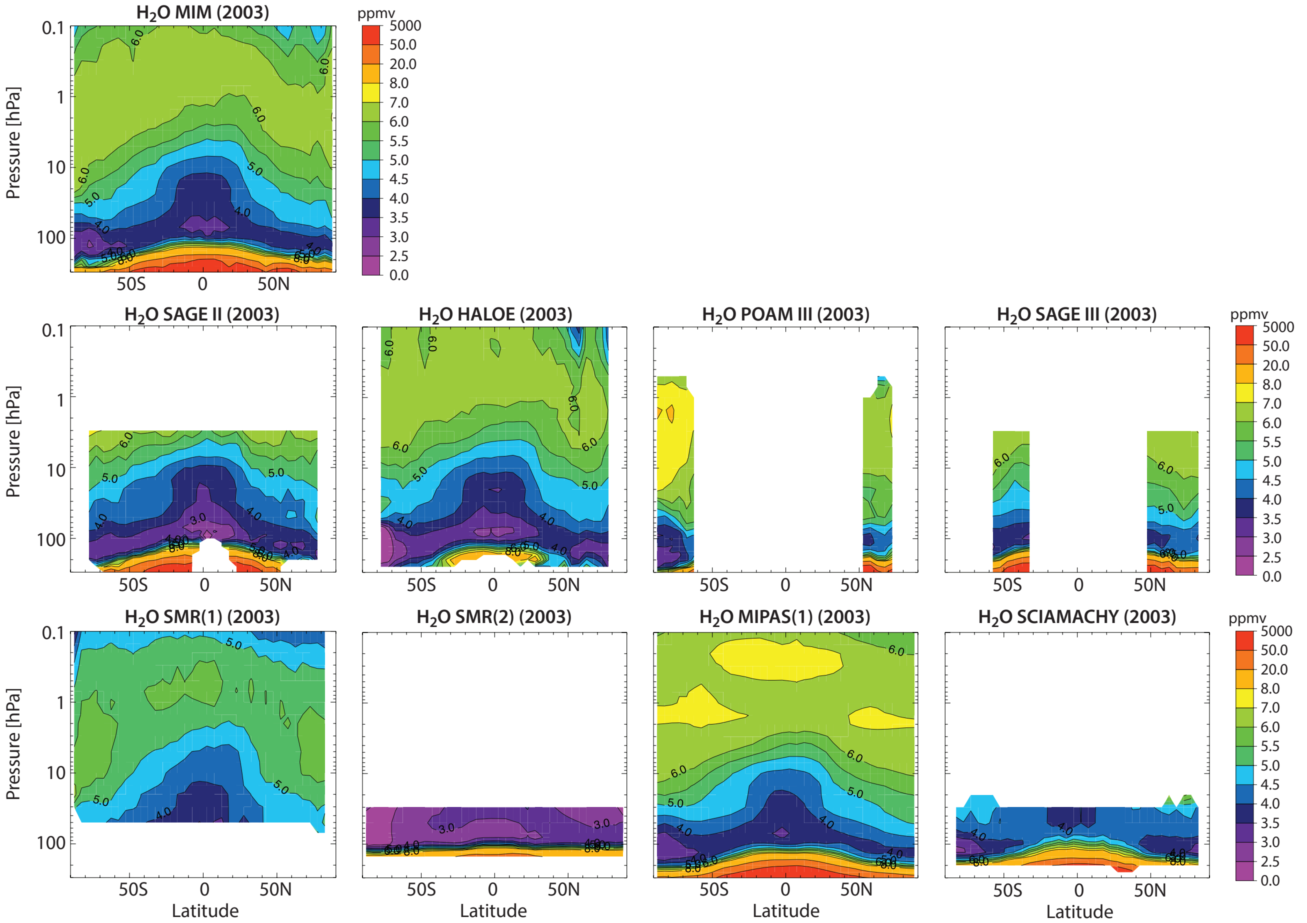


Figure 4.2.8a: Cross sections of annual zonal mean H_2O for 2003. *Shown are the annual mean cross sections for the MIM (upper row), SAGE II, HALOE, POAM III, SAGE III (middle row), and SMR(1), SMR(2), MIPAS(1), and SCIAMACHY (lower row). Note that SMR(2) is not included in the MIM.*

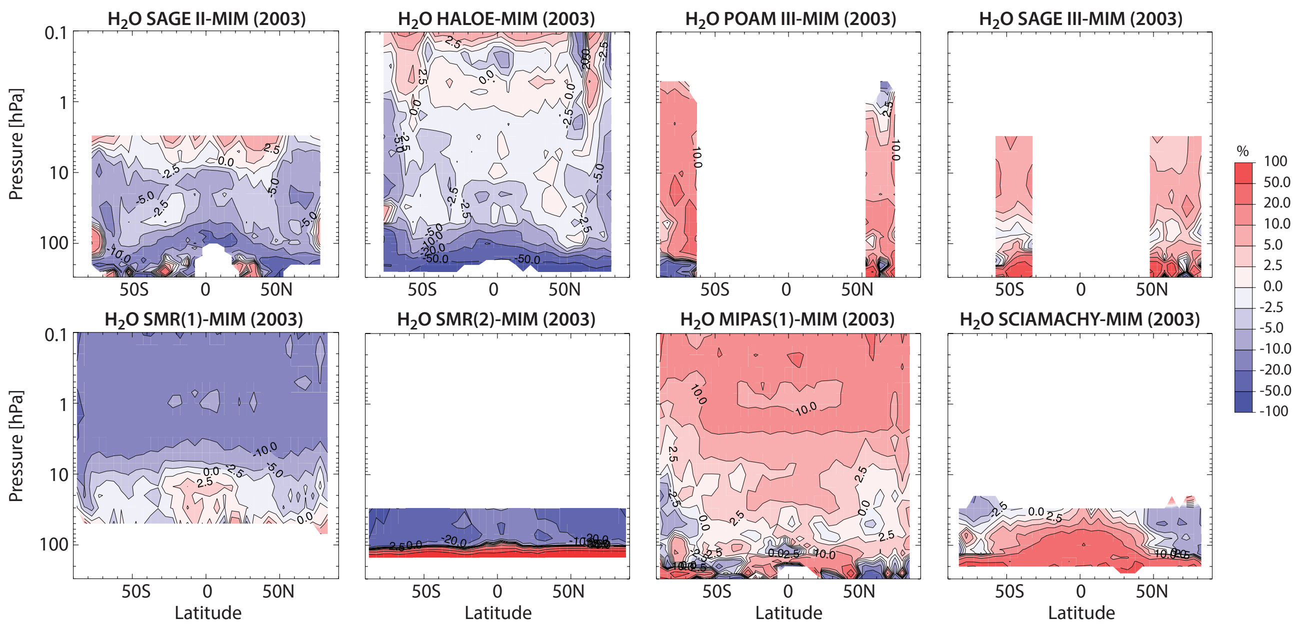


Figure 4.2.8b: Cross sections of annual zonal mean H_2O differences for 2003. *Relative differences with respect to the MIM are shown for the individual instruments shown in Figure 4.2.8a. Note that SMR(2) is not included in the MIM.*

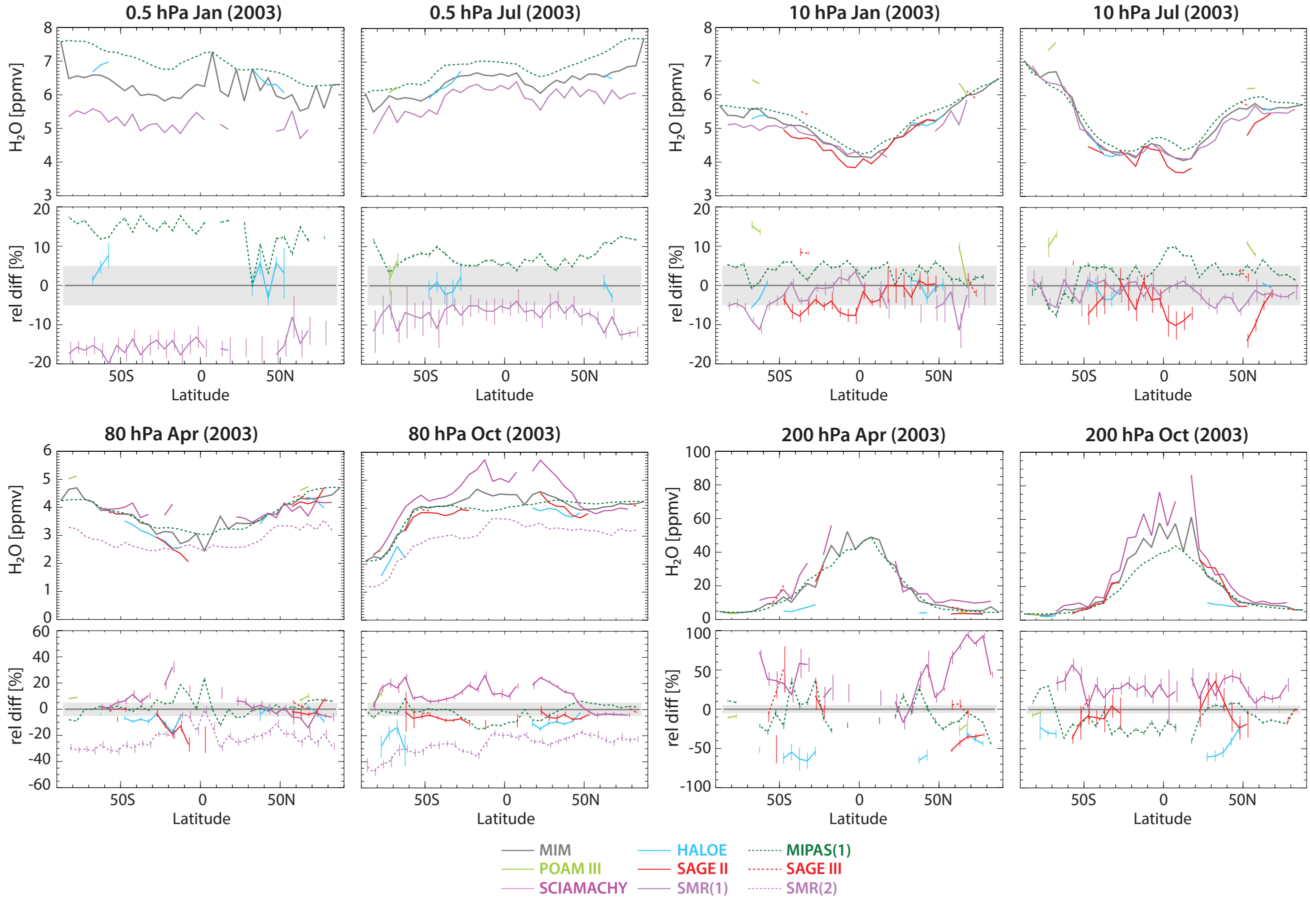


Figure 4.2.9: Meridional profiles of monthly zonal mean H_2O for 2003. *Meridional profiles are shown at 0.5 and 10 hPa for January and July (upper row), and at 80 and 200 hPa for April and October (lower row). Upper panels show absolute values, and lower panels show relative differences between the individual instruments (SAGE II, HALOE, POAM III, SMR(1,2), SAGE III, MIPAS(1), and SCIAMACHY) and the MIM. The grey shading indicates where the relative differences are smaller than ±5%. Error bars indicate the uncertainty in the relative differences based on the SEM of each instrument.*

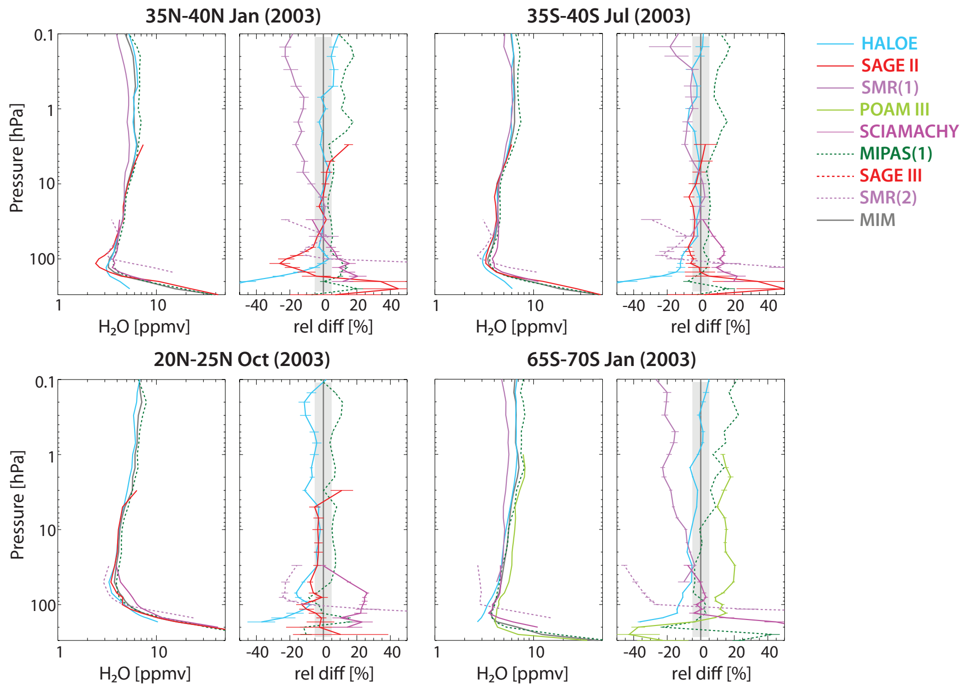


Figure 4.2.10: Vertical profiles of monthly zonal mean H_2O for 2003. *The H_2O profiles and their differences relative to the MIM are shown for January 35°N-40°N, July 35°S-40°S, January 65°S-70°S, and October 20°N-25°N. Error bars indicate the uncertainty in the relative differences based on the SEM of each instrument. The grey shaded area indicates where relative differences are smaller than ±5%.*

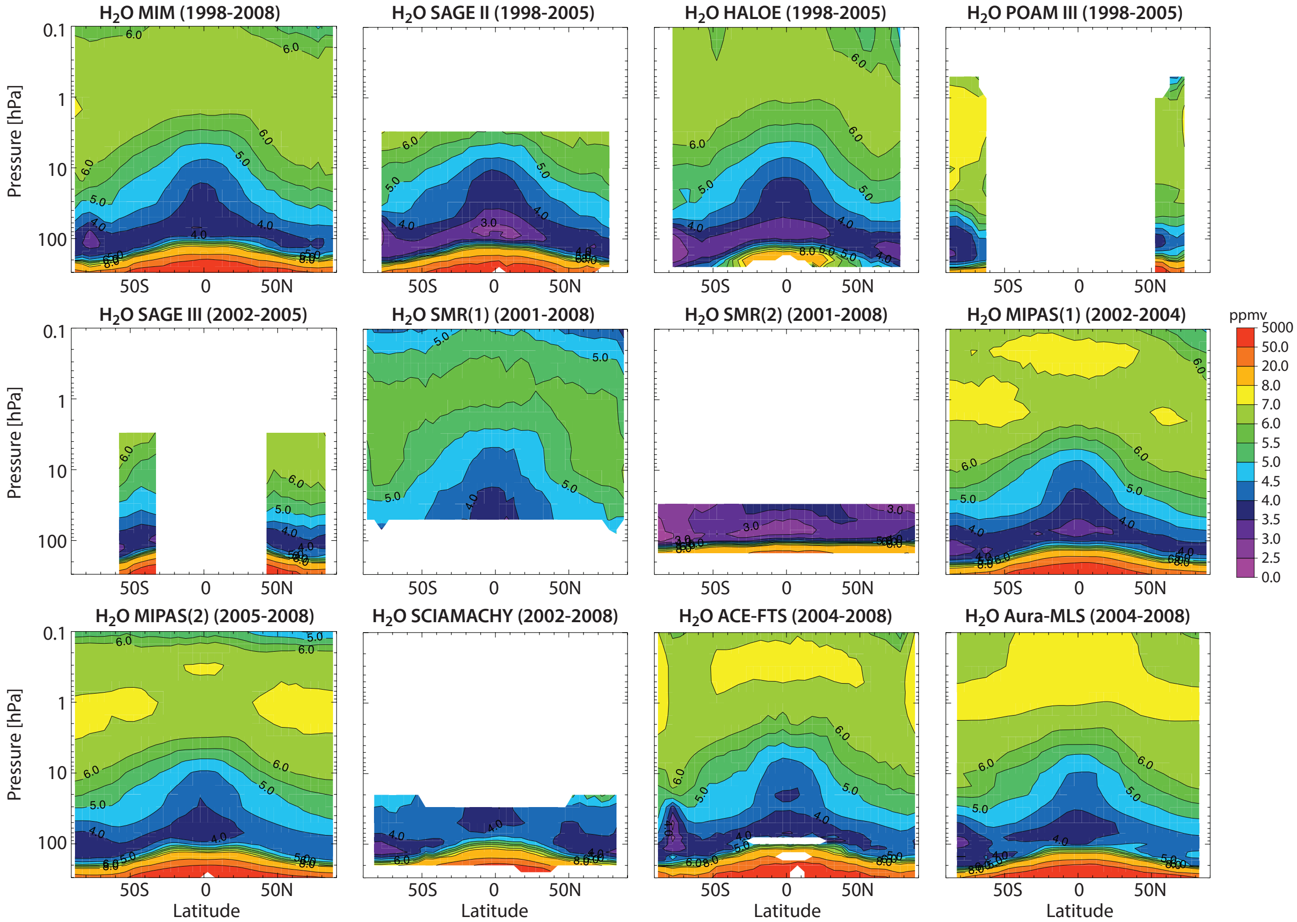


Figure 4.2.11a: Cross sections of annual zonal mean H_2O for 1998-2008. *Shown are the MIM, SAGE II, HALOE, POAM III (upper row), SAGE III, SMR(1), SMR(2), MIPAS(1) (middle row), and MIPAS(2), SCIAMACHY, ACE-FTS, and Aura-MLS (lower row). Note, SMR(2) and MIPAS(1) are not included in the MIM.*

A somewhat intriguing result is that the older set of the instruments (SAGE II and HALOE, together with SMR(1)) show much smaller values than the newer set of instruments (MIPAS(1), MIPAS(2), SAGE III, ACE-FTS, Aura-MLS) throughout most of the stratosphere with differences from the MIM of up to -10% (resulting in inter-instrument differences of up to 20%). POAM III is an exception; it belongs to the older set, but exhibits rather large positive deviations from the MIM. In the USLM, SMR(1) shows the largest negative differences (around -15%) and Aura-MLS the largest positive differences from the MIM (around +10%). MIPAS(2), as in its earlier mode MIPAS(1), reports positive deviations compared to the MIM through most of the stratosphere. However, in contrast to MIPAS(1), MIPAS(2) shows negative differences in the LM and positive differences in the UTLS (except in the tropical UT). Below 100 hPa, SAGE II and HALOE show deviations from the MIM that are larger than -20%. SMR(2) exhibits relative differences of up to +100% below and up to -50% above 100 hPa, respectively. This data product is known to yield less reliable information above 50 hPa [*Urban et al.*, 2012]. Aura-MLS shows a 'sandwich' structure in the UTLS, with a layer of negative deviations in between layers of positive deviations.

Figure 4.2.12 shows meridional profiles for 1998-2008. At 0.5 hPa, MIPAS(1), MIPAS(2), Aura-MLS and ACE-FTS agree within 5-10%, while HALOE and in particular SMR(1) show much lower values. These results are similar to what has been seen for the 2003 evaluations. An independent study by *Nedoluha et al.* [2009] using the Water Vapour Millimeter-wave Spectrometer (WVMS) measurements over Mauna Loa for validating HALOE, ACE-FTS, and Aura-MLS mesospheric H_2O measurements, confirms that Aura-MLS and ACE-FTS are within ±0.5-1.5% of the WVMS measurements, while HALOE is biased low by around 10%. The monthly zonal means of SMR(1) are even lower than HALOE therefore can also be considered to have a low bias.

At 10 hPa, most instruments agree well (within ±5%). Exceptions are SAGE II, which shows much lower values, and POAM III, which shows much higher values than the other instruments (up to 15% deviation from the MIM) in both months. SMR(1) is on the low side of the other measurements. At 80 hPa, the spread in the measurements increases strongly to ±20%, with somewhat smaller discrepancies in the extra-tropics. SMR(2), SAGE II and to a somewhat lesser extent HALOE are all on the low side of the MIM. SCIAMACHY shows a large positive deviation from the MIM of up to 40% in the tropical region during April, however agrees well with the other instruments in the extra-tropics and during October. MIPAS(1), MIPAS(2), ACE-FTS and Aura-MLS agree within 15%. At 200 hPa, the

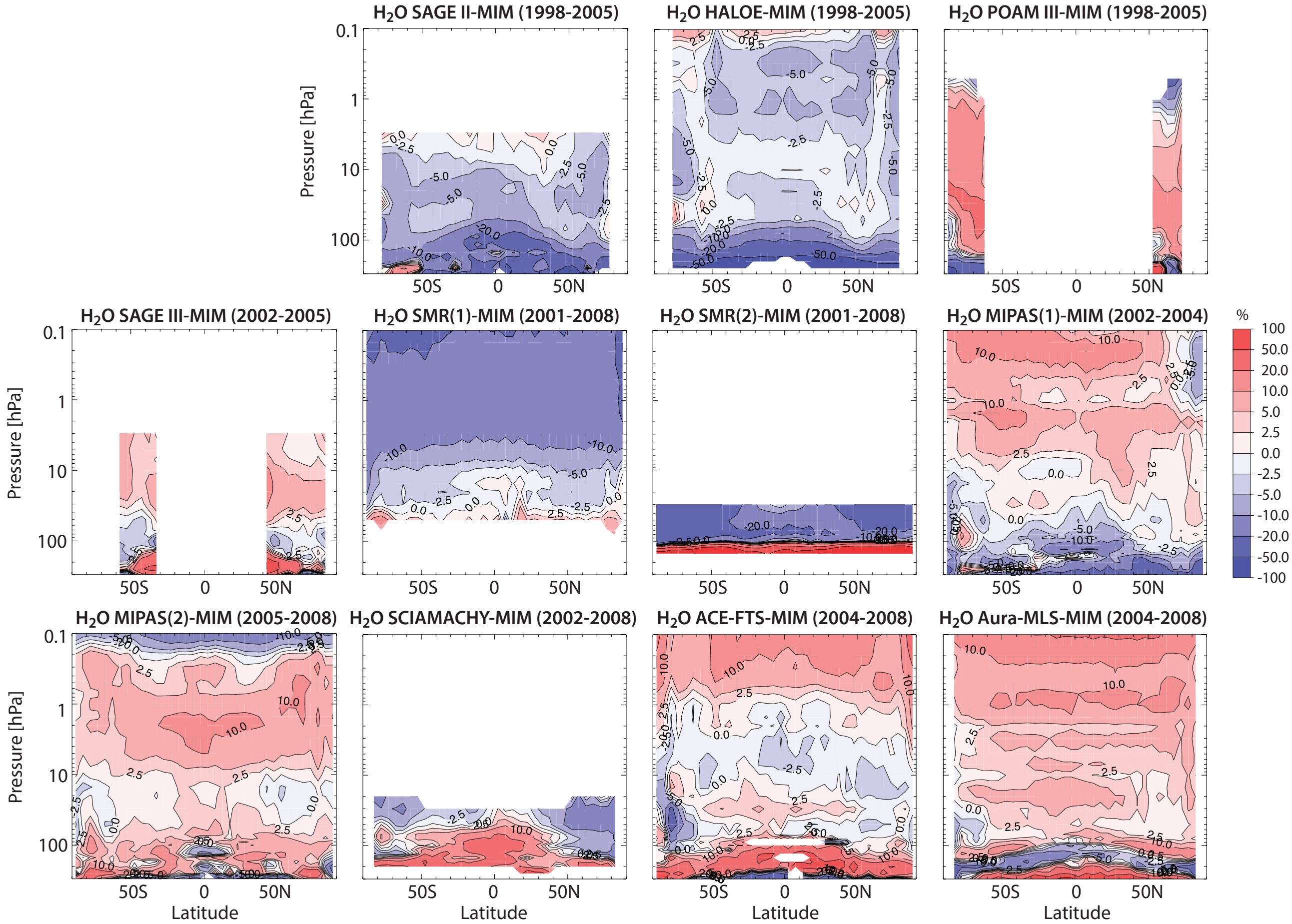


Figure 4.2.11b: ***Cross sections of monthly zonal mean H_2O differences for 1998-2008.*** *Differences relative to the MIM are shown for the individual instruments shown in Figure 4.2.11a. Note that SMR(2) and MIPAS(1) are not included in the MIM.*

instruments agree mostly within ±50% from the MIM, with SCIAMACHY and ACE-FTS showing largest positive and HALOE largest negative deviations.

Figure 4.2.13 shows the vertical profile comparison for the time period 1998-2008, highlighting the vertical structure in the differences of the individual instruments. It also shows that the monthly mean differences are somewhat larger when compared to the annual zonal mean evaluation. The vertical profiles emphasise the good agreement between most instruments in the MS, and identify the instruments that are outliers. Note that the SEM provides a measure of how well the climatologies are defined, and therefore whether the inter-instrument differences are significant or not. These SEM values are generally much smaller for the limb-emission sounders, and are larger in the UTLS than in the MS. The differences between the individual instruments in the UTLS are therefore less well defined.

The validation results based on the comparison of annual and monthly zonal mean climatologies presented here largely confirm validation results obtained for the different satellite instruments using the classical coincidence validation method that compares single profile matches. Other validation activities using ground-based, balloon or aircraft measurements yield further insight into the relative differences between the satellite instruments or help confirm our findings. For example, *Lucke et al.* [1999] found in early comparisons between POAM III and HALOE absolute differences of around 20-25% in the LS and 10-15% in the MS, with POAM III on the high side. These results have been confirmed and extended by *Lumpe et al.* [2006], showing the very good agreement (within 5%) of POAM III with coincident ER-2 and FISH aircraft measurements in the extra-tropical UTLS (between 100-300 hPa). *Thomason et al.* [2010] found mostly consistent results based on profile comparisons between SAGE II, POAM III, Aura-MLS, and HALOE, as did *Carleer et al.* [2008] for comparisons between these instruments and ACE-FTS. The latter found also a very good agreement between ACE-FTS and lidar measurements, with differences of 5% in the MS and US, and increasing differences toward the LM, consistent with the vertical structure seen in the differences of ACE-FTS in **Figures 4.2.11b** and **4.2.13**. Comparisons with aircraft measurements indicate that ACE-FTS exhibits uncertainties of 30% in the UT and 18% in the LS, respectively [*Hegglin et al.*, 2008]. Note that *Hegglin et al.* [2008] were also using a climatological approach to validate H_2O measurements in the UTLS, which accounted for the high geophysical variability in this region and were able to reduce previously reported uncertainties in the UTLS based on the classical validation method using coincident measurements by up to 50%.

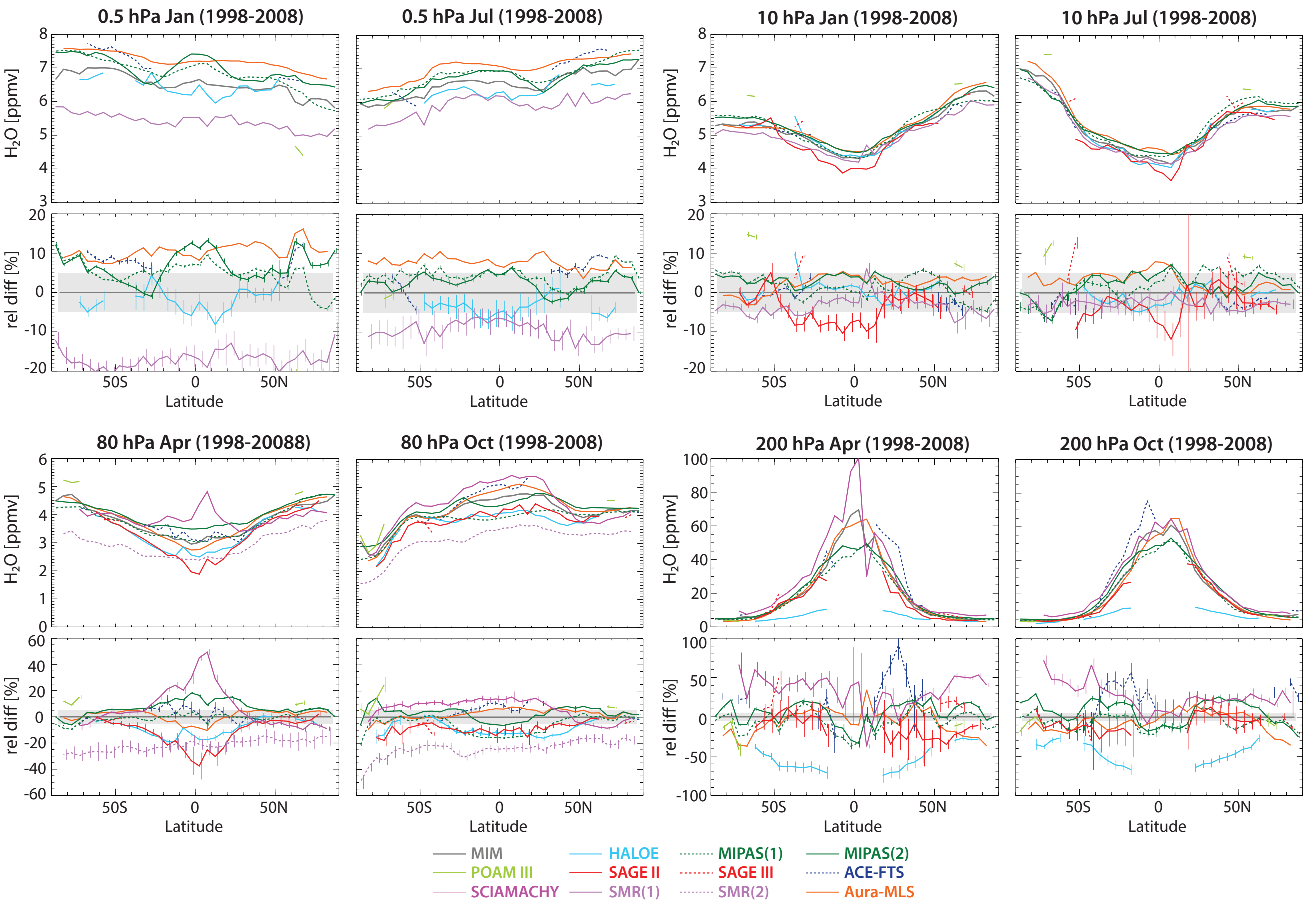


Figure 4.2.12: Meridional profiles of monthly zonal mean H_2O for 1998-2008. *Meridional profiles are shown at 0.5 and 10 hPa for January and July (upper row), and at 80 and 200 hPa for April and October (lower row). Upper panels show absolute values, lower panels relative differences between the individual instruments (SAGE II, HALOE, POAM III, SMR(1,2), SAGE III, MIPAS(1,2), SCIAMACHY, ACE-FTS, and Aura-MLS) and the MIM, respectively. The grey shading indicates where the relative differences are smaller than ±5%. Error bars indicate the uncertainty in the relative differences based on the SEM of each instrument.*

For MIPAS, MIPAS(1) water vapour measurements have been validated by *Milz et al.* [2009]. They have confirmed the MIPAS precision estimates of 5-10%. The MIPAS(2) reduced spectral resolution measurements have been validated by *Stiller et al.* [2012] in the framework of the MOHAVE-2009 campaign [*Leblanc et al.*, 2011]. They found that between 12 km and 45 km, MIPAS(2) water vapour (version V4O H_2O 203) was well within 10% of the data of all correlative instruments. The well-known dry bias of the MIPAS(2) water vapour standard product from nominal observations above 50 km due to neglect of non-LTE effects in the current retrievals has also been confirmed.

Lambert et al. [2007] have shown that Aura-MLS H_2O values compare quite well, overall, with other satellite datasets, in ways that are consistent with the results shown here. Namely, the stratospheric Aura-MLS values tend to be 5-10% wetter than HALOE H_2O, but 5-10% drier than POAM III H_2O. Other studies have shown that HALOE H_2O values tend to typically be lower than other datasets [*e.g.*, *SPARC*, 2000]. Comparisons by *Nedoluha et al.* [2007; 2009; see discussion above] of Aura-MLS and HALOE H_2O with upper stratospheric H_2O from the WVMS results above Lauder and Mauna Loa also show that HALOE H_2O values are smaller than the other two datasets. These authors also conclude that good correlations exist between the observed seasonal and interannual variations from Aura-MLS and WVMS. The Aura-MLS H_2O measurements have also been shown to compare very well with cryogenic frost-point hygrometer (CFH) profiles in the LS and MS; MLS V2.2 values are about 2-3% larger than CFH values [*Read et al.*, 2007; *Voemel et al.*, 2007]. Note that the SPARC Data Initiative climatologies are based on V3.3 data, with inferred deviations from the CFH values of about 5-6%, since they show a slight increase in mixing ratios compared to V2.2 used in these studies. The differences increase with decreasing altitude, and around 216 hPa Aura-MLS exhibits a negative bias of up to 25%.

4.2.3 H_2O evaluations: Seasonal cycles

Water vapour exhibits strong seasonal cycles in both the tropical and extra-tropical UTLS due to its dependence on transport and Lagrangian cold-point temperatures [*Fueglistaler et al.*, 2009; *Hoor et al.*, 2010]. Most attention has focused on the tropics between 80 and 100 hPa, where the stratospheric entry value of water vapour is slaved to the

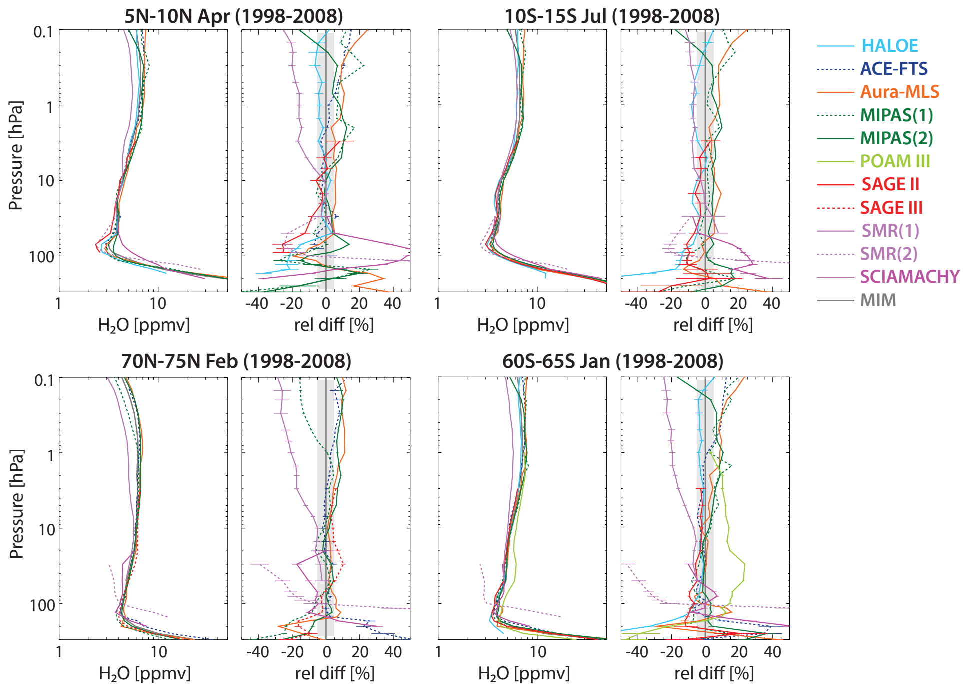


Figure 4.2.13: Vertical profiles of monthly zonal mean H_2O for 1998-2008. *The H_2O profiles and their relative differences to the MIM are shown for April 5°N-10°N, July 10°S-15°S, February 70°N-75°N, and January 60°S-65°S. Error bars indicate the uncertainties in the relative differences based on the SEM of each instrument. The grey shaded area indicates where relative differences are smaller than ±5%.*

seasonally changing cold-point temperatures [*e.g.*, *Fujiwara et al.*, 2010]. However, the seasonal cycle is also of interest in the extra-tropics (especially at the lower levels 200 and 300 hPa) where it reflects the impact of stratosphere-troposphere exchange, and hence is insightful for the evaluation of transport processes in chemistry-climate models [*e.g.*, *Hegglin et al.*, 2010].

Figure 4.2.14 (left and middle panels) shows the seasonal cycles in water vapour at 80 and 100 hPa in the tropical LS averaged over the years 1998-2008 for all available instruments, respectively. The seasonal cycles show a minimum in H_2O during February to April and a maximum during September to October. The seasonal cycle peaks somewhat later at 80 hPa because of the time needed to transport the tape recorder signal upwards into the stratosphere. The absolute values in the seasonal cycle are somewhat better constrained at 80 hPa (with a 1σ uncertainty range of ±15%) than at 100 hPa (±22.5%). HALOE and SAGE II show year-round much lower values than the other instruments. SMR(2) shows lower values than the MIM at 80 hPa, but is in excellent agreement with the MIM at 100 hPa. SCIAMACHY on the other hand shows the highest monthly values throughout the year. The high bias in SCIAMACHY results from the way the climatologies were compiled given the instrument's specific vertical sampling. The sampling altitudes of SCIAMACHY (~70 and 130 hPa) are located relatively far above and below the 80 and 100 hPa levels. Interpolation of the retrieved data onto the SPARC Data Initiative pressure levels therefore leads to a strong smearing of the high tropospheric values into the lower stratosphere. A seasonal cycle taken at 70 hPa shows much better agreement between SCIAMACHY and the other instruments (not shown). We find the best agreement between the mean monthly values of ACE-FTS, Aura-MLS, MIPAS(1) and MIPAS(2). Note that the mean H_2O values are an essential performance metric, although their evaluation is not included in the Taylor diagram. **Figure 4.2.14** (right panel) shows the H_2O seasonal cycle in the UT at 150 hPa. The seasonality at this level is less pronounced and the mean values are less well constrained (±30%, if SMR(2) and ACE-FTS are excluded from the evaluation).

Focusing on the seasonal cycle's amplitude and phase, the Taylor diagrams reveal better agreement between the instruments at 100 than at 80 hPa. At 80 hPa, HALOE and SCIAMACHY agree on amplitude and phase, with MIPAS(1), MIPAS (2), and SMR(2) showing a smaller, and ACE-FTS, SAGE II, and Aura-MLS showing a larger amplitude. The seasonal cycle is not well constrained by ACE-FTS due to the instrument's limited temporal sampling of tropical latitudes. Nevertheless, the available monthly data are distributed such that the amplitude and phase are fairly well captured. At 100 hPa, SMR(2) is the instrument with the best skill score, with monthly mean values that are closest to the MIM. This is especially noteworthy since SMR(2) shows large negative (positive) deviations above (below) this level in the zonal mean cross sections. Aura-MLS shows the best correlation with the MIM at both levels, however with a slightly larger amplitude than the other instruments especially at 80 hPa. SCIAMACHY's amplitude is close to the MIM, despite its aforementioned too high mean values. MIPAS(1) and MIPAS(2) are well correlated with the MIM on both levels, however, both show amplitudes that are slightly too low compared to the MIM. The too low amplitude is explained by state-dependent averaging kernels;

H_2O profiles are better resolved in altitude in a more humid atmosphere, while the averaging kernels are widened in the case of a very dry atmosphere. This means that the sharp signature of the hygropause in the dry phase of the seasonal cycle cannot be properly resolved, leading to a reduced amplitude of the seasonal cycle. Application of MIPAS averaging kernels within comparisons would hence remove the problem. Note that the sampling of HALOE and SAGE II in the tropics is more limited towards the end of the missions, so that seasonal cycles calculated for 2003-2005 do not capture the amplitude and phase properly (not shown). At 150 hPa, SAGE II, HALOE and MIPAS(2) agree well on phase (correlation of 0.7) and amplitude. SCIAMACHY agrees with the correlation, however shows a larger amplitude. SMR(2) and ACE-FTS do not reproduce the seasonal cycle. Here, SMR(2) is below the recommended altitude range, and ACE-FTS suffers from inadequate sampling.

Seasonal cycles in water vapour for the Southern and Northern Hemisphere mid- to high latitudes at different pressure levels (100, 200, and 300 hPa) are displayed in **Figure 4.2.15**. The maxima in the seasonal cycle at 300 and 200 hPa are seen during summer, while the maximum at 100 hPa is found during winter, reflecting that the 100 hPa level is slaved to the tropics with a time lag of about 3 to 4 months, while the lower levels are affected by transport processes across the extra-tropical tropopause on a shorter time scale and the tropopause height itself [*Hegglin et al.*, 2010]. The seasonal cycle mean values are better constrained at 100 hPa than at the lower levels with an associated 1σ-uncertainty range that is about ±15% at 100 hPa year-around, but up to ±25-50% during summer peak values at 200 and 300 hPa.

The seasonal cycle at 100 hPa in the Southern Hemisphere is influenced by both dehydration at the tropical tropopause during Northern Hemisphere winter and dehydration within the polar vortex during Southern Hemisphere winter. Instead of the expected maximum during winter (compare to Northern Hemisphere) this leads to a semi-annual cycle with one minimum occurring during February/March and another minimum occurring during August/September (also compare to **Figure 4.2.19**). ACE-FTS shows the best agreement with the MIM, reflected in a high skill score and also in terms of monthly mean values. The same is true for MIPAS(2), although its mean values are somewhat larger than those of the MIM. At 100 hPa in the Northern Hemisphere, Aura-MLS, MIPAS(2), SCIAMACHY and SAGE II agree very well in terms of correlation and phase. However, Aura-MLS shows much higher and SAGE II and SCIAMACHY show much lower monthly mean values than the other instruments. SMR(2) and HALOE are also on the low side of the MIM. Best agreement with the monthly mean values is seen for ACE-FTS, MIPAS(1), SMR(2) and SAGE III. Note that the seasonal cycle in this region is very weak and signals are therefore hard to interpret given the sampling limitations of the individual instruments.

The instruments show the largest spread in skill at 200 hPa in both hemispheres. In the Southern Hemisphere, the spread is mainly due to a disagreement in the amplitude, while in the Northern Hemisphere the spread is also due to a disagreement in the phase. HALOE exhibits no discernible seasonal cycle at and below this altitude (at pressure levels smaller than 200 hPa) in both hemispheres. Note that HALOE performs much better at higher altitudes, although still with monthly mean values that are smaller than the MIM (not shown). In the Southern Hemisphere, SAGE III exhibits a much stronger amplitude than the MIM. SCIAMACHY, despite showing excellent agreement in the phase, shows a slightly too high amplitude

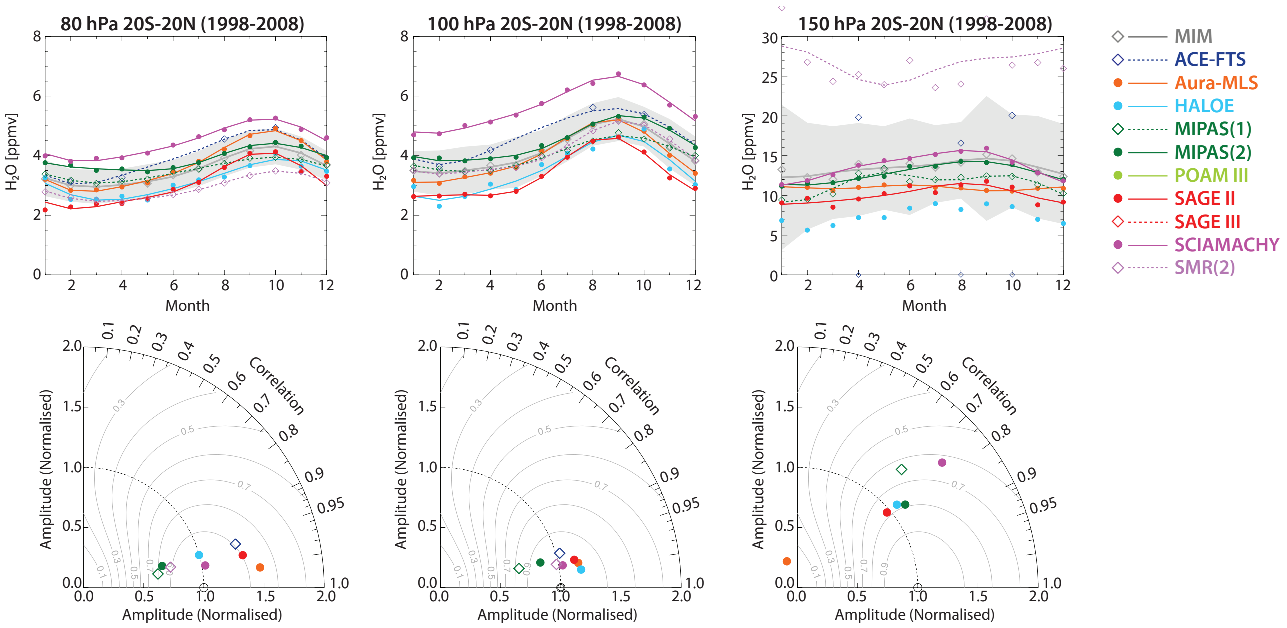


__Figure 4.2.14: Seasonal cycles of H_2O in the tropics for 1998-2008.__ Seasonal cycles and corresponding Taylor diagrams of monthly zonal mean H_2O averaged over 20°S-20°N are shown at 80 (left column), 100 hPa (middle column) and 150 hPa (right column). Coloured lines represent fits including an annual and a semi-annual component to the available monthly data points. The grey line indicates the multi-instrument mean (MIM) and the grey shading ±1σ.

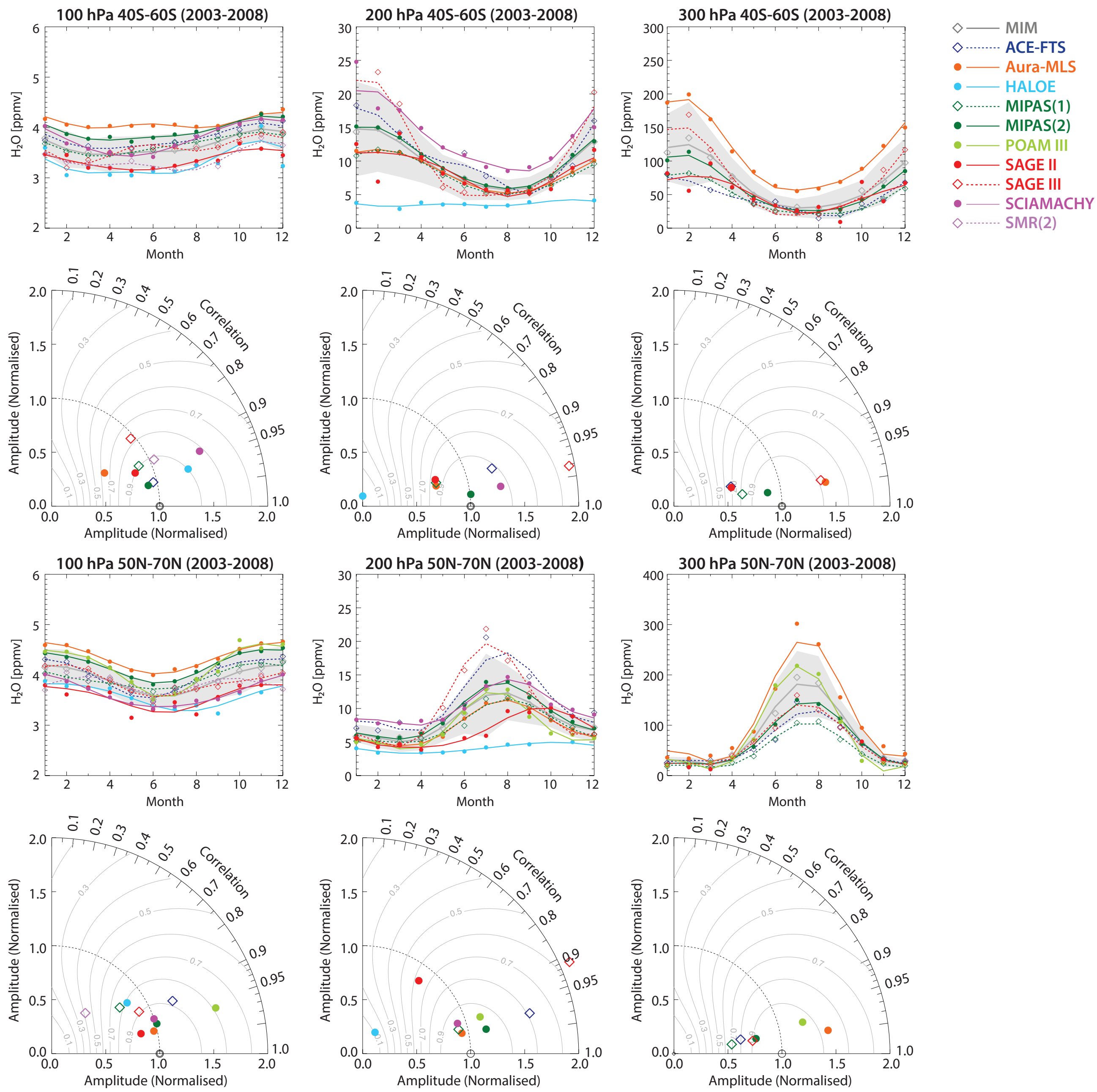


Figure 4.2.15: Seasonal cycles of H_2O in the SH and NH mid-latitudes for 1998-2008. *Seasonal cycles and corresponding Taylor diagrams of monthly zonal mean H_2O averaged over 40°S-60°S (upper two rows) and 50°N-70°N (lower two rows) are shown at 100, 200, and 300 hPa (from left to right). Coloured lines represent fits including an annual and a semi-annual component to the available monthly data points. The grey line indicates the multi-instrument mean (MIM) and the grey shading ±1σ.*

and is an outlier regarding its much larger mean values when compared to the other instruments throughout the year. Reasonably good agreement and hence constraint on the seasonal cycle is achieved by MIPAS(1) and (2), SAGE II and Aura-MLS. In the Northern Hemisphere the agreement between the instruments is somewhat better. Here SCIAMACHY, MIPAS(1) and (2), Aura-MLS, and POAM III agree very well in correlation, amplitude, and mean value, while ACE-FTS and SAGE III exhibit too large amplitudes, and SAGE II a wrong phase in the seasonal cycle peaking two to three months later than the other instruments.

At 300 hPa, we find better agreement, with two clusters of instruments in both the Southern and Northern Hemisphere that show high correlations (>0.95), but large differences in their amplitudes. In the Southern Hemisphere, the cluster of instruments consists of ACE-FTS, SAGE II, and MIPAS(1) showing much smaller amplitudes than Aura-MLS and SAGE III. MIPAS(2) shows the best agreement with the MIM, in terms of amplitude, phase, and mean values. In the Northern Hemisphere, it is again Aura-MLS, together with POAM III, which shows a much larger amplitude than the other instruments. MIPAS(1), MIPAS(2), ACE-FTS, and SAGE III agree with each other, but are on the low side of the MIM.

The difficulties of reproducing the annual cycle in water vapour at different levels in the UTLS are related to the strong vertical gradients in water vapour found across the tropopause and the narrow vertical region over which the annual cycle extends, both requiring high vertical resolution measurements and/or high vertical sampling to be adequately resolved. Also, instrument limitations resulting from cloud interference and high extinction exist in this altitude region. Clearly, instruments with less frequent sampling show less robust results, *e.g.*, ACE-FTS agrees well with other instruments at 300 hPa in the Northern Hemisphere, but seems to overestimate the amplitude at 200 hPa. UTLS-specific evaluations using tropopause co-ordinates or equivalent latitude may help improving the comparisons in the future and define better constraints for model-measurement comparisons.

4.2.4 H_2O evaluations: Tape recorder

The atmospheric tape recorder [*Mote et al.*, 1996] is one of the most pronounced spatio-temporal patterns in equatorial water vapour, showing the slow upward propagation of a minimum in H_2O from the tropical tropopause region up to altitudes of around 30 km. The signal is produced by seasonal variations in tropical tropopause temperatures that determine the H_2O saturation mixing ratios in air masses entering the tropical stratosphere. A realistic characterisation of the tape recorder is a key aspect of the physical consistency of the different datasets, provided that the sampling is adequate.

Figure 4.2.16a shows the tape recorders of the individual instruments for which tropical data were available for a latitude band between 15°S and 15°N and the time period 2000-2010. No tape recorder could be produced for SAGE III and POAM III, which have no tropical coverage. Most of the satellite instruments do capture the upward propagation of low water vapour mixing values. Although a tape recorder is also visible for SCIAMACHY, the minimum in H_2O just above the tropical tropopause is much weaker, and the higher mixing ratios reach further into the stratosphere as seen for the MIM. As discussed earlier, this is due to the coarse sampling of SCIAMACHY in the tropopause region that leads to strong smearing of the values across the tropopause. SMR(2) shows much lower mixing ratios than the other instruments throughout the tape recorder signal. Due to limited temporal coverage in the tropics, the ACE-FTS had to be interpolated in time and altitude to obtain a tape recorder, but captures the main features of the tape recorder well.

We find that the tape recorders of the individual instruments show much stronger signals (*i.e.*, lower minimum mixing ratios) for 2000-2005 than for 2005-2010. Also, we see large relative differences throughout the stratosphere for the overlapping time period 2002-2004 (see **Figure 4.2.16b**), which indicates that the early and later data records cannot simply be concatenated for use in trend analyses.

Figure 4.2.16b shows the differences in the tape recorders with respect to the MIM. It reveals that for the period 2000-2005, SAGE II and HALOE seem to agree well, with differences that have a rather noisy structure, which implies that the two instruments have no systematic biases and that the structure (tape recorder signal) they reproduce is physically consistent. Both these instruments show lower values than the new generation of instruments (SMR(1) and MIPAS(1)) that contribute to the MIM at the beginning of 2002. Since SMR(1) yields the most negative deviations from the MIM after 2004 when more instruments are available, it must follow that HALOE (and SAGE II for this matter) would be on the low side of these as well. In the later period, MIPAS(2), SCIAMACHY, and Aura-MLS exhibit structures in the differences in the LS that resemble the tape recorder itself, implying a systematic difference, which may be due to the effects of different vertical resolutions (see **Table 4.2.2**). Resolution issues would affect the derived amplitude of the tape recorder, which is often used as a diagnostic in model-measurement comparisons. MIPAS(2) and Aura-MLS have higher values in the MS when compared to the ACE-FTS and SMR(1). The interpolated ACE-FTS data show – aside from the effects discussed above – differences relative to the MIM within the range of the other instrument differences. SMR(2) shows negative deviations of > 20% from the MIM in the 50-100 hPa range. However, the noise in the relative deviations indicates that it captures the seasonal cycle reasonably well compared the other instruments.

A tape recorder has also been derived for the LIMS instrument (see **Figure 4.2.17**). While the tape recorder shows a distinct minimum in H_2O above the tropical tropopause, there seems a lack of propagation of the signal into the middle stratosphere. Note that the data are very limited in time. Not shown is the tape recorder for UARS-MLS, which however captures the tape recorder signal in the LS and MS as demonstrated before by *Pumphrey* [1999].

4.2.5 H_2O evaluations: Horizontal tape recorder

Seasonal variations in the imprint of the cold point tropopause temperatures on H_2O saturation mixing ratios not only propagate upwards into the stratosphere, but they also spread poleward on shorter time scales due to strong horizontal transport and mixing [*SPARC*, 2000] as is depicted in **Figures 4.2.18** and **4.2.19**. A minimum in H_2O is observed between February through May near 10°N-20°N, which consequently is mixed into higher latitudes, but also into the Southern Hemisphere. During August to October a strong maximum in H_2O is observed with two peaks centered at 30°N and 10°S for most of the instruments. These maxima are due to higher tropopause temperatures during Northern Hemisphere summer and may also be partly influenced by transport of moister air into the stratosphere within the summer monsoons. These higher values slowly spread to higher latitudes, also in the winter hemisphere. Note that during the later period (2005-2010, **Figure 4.2.19**), the air entering the stratosphere is moister than during the earlier period (1998-2005, **Figure 4.2.18**) as seen from the comparison of individual instruments available in both periods.

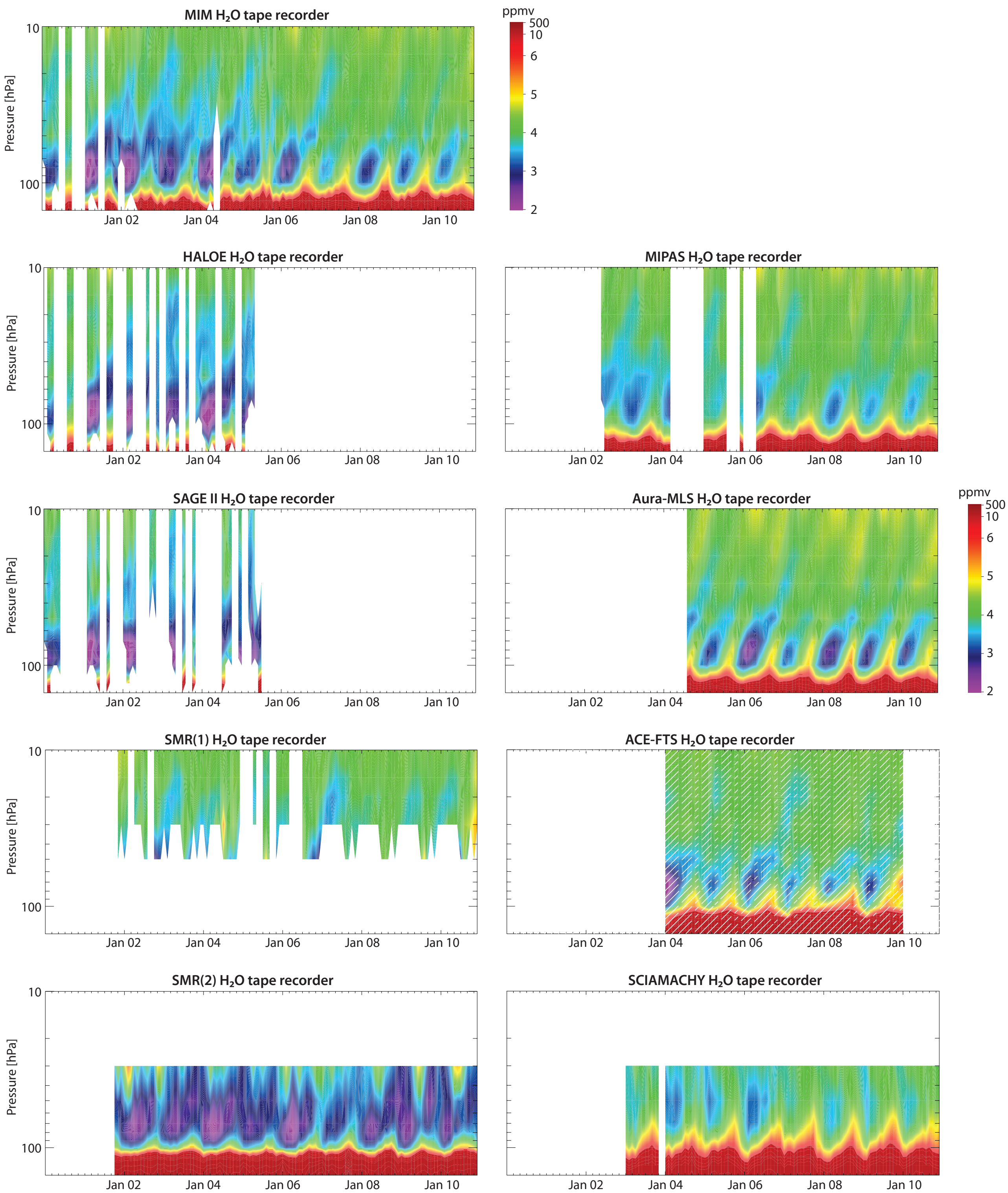


***Figure 4.2.16a:** **H_2O tape recorder.** Shown is the altitude-time evolution of H_2O averaged over 15°S-15°N for the time period 2000-2010. The very limited tropical ACE-FTS data were interpolated in time and altitude; white hatching indicates regions that do not contain data. Note that the SMR(2) and SCIAMACHY products are not included in the MIM.*

The individual instruments show different degrees of skill in reproducing the horizontal tape recorder. The horizontal gradients are relatively small and hence pose a challenge to the instruments. Aura-MLS shows slightly higher H_2O mixing ratios in the extra-tropics than the other instruments and the minimum during Northern Hemisphere winter to be centered at the equator, similar to SAGE II and HALOE in **Figure 4.2.18**. SMR(2) reproduces the main features of the MIM although shows a somewhat noisier field and without the split in the maxima during August through November. SCIAMACHY suffers from the earlier mentioned fact that the SPARC Data Initiative 100 hPa level shown here lies in between the two native retrieval levels leading to smearing across the tropopause. This issue leads to too high H_2O mixing ratios in the tropics year-around. In the extra-tropics the effect of the smearing is smaller and the

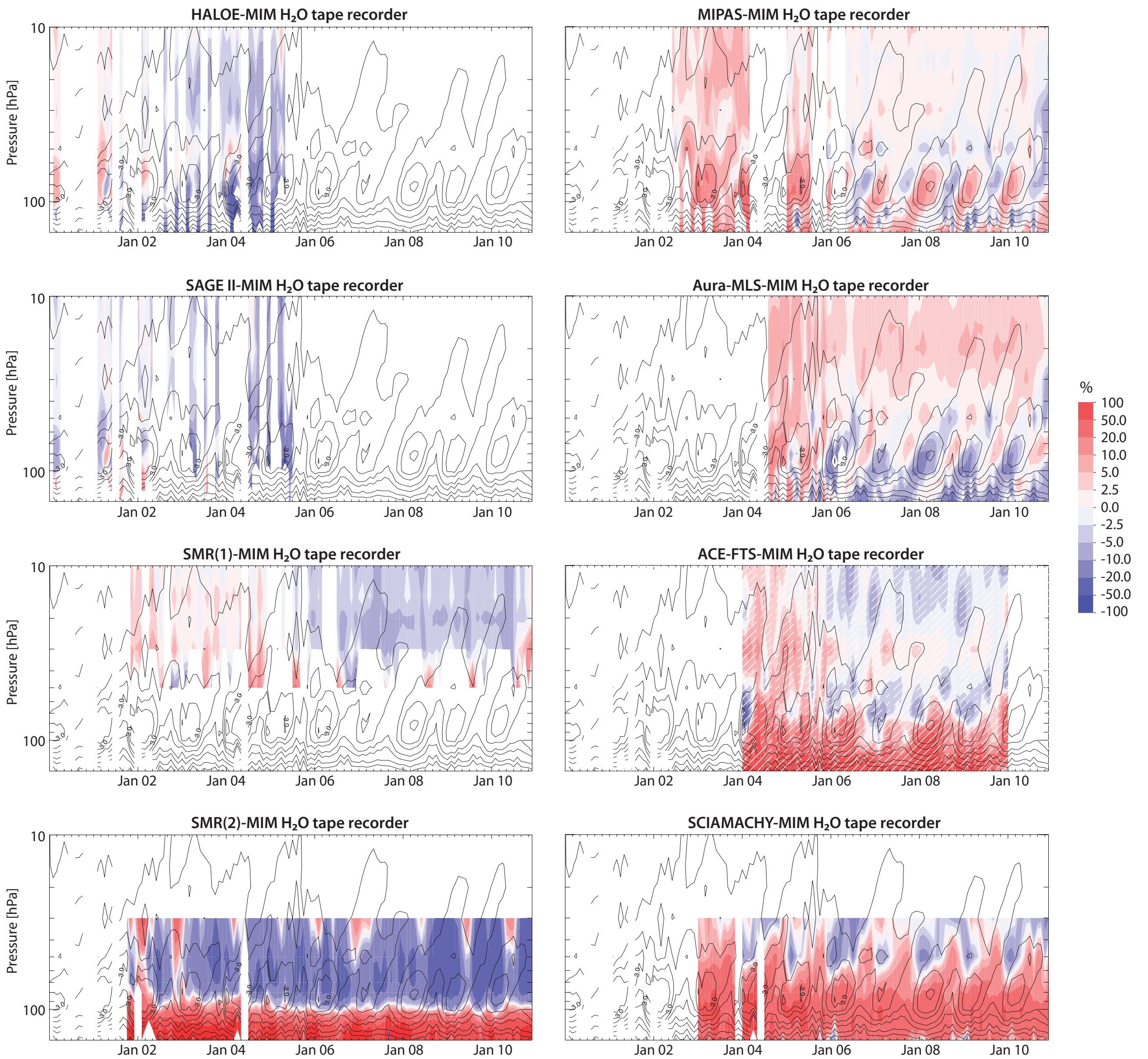
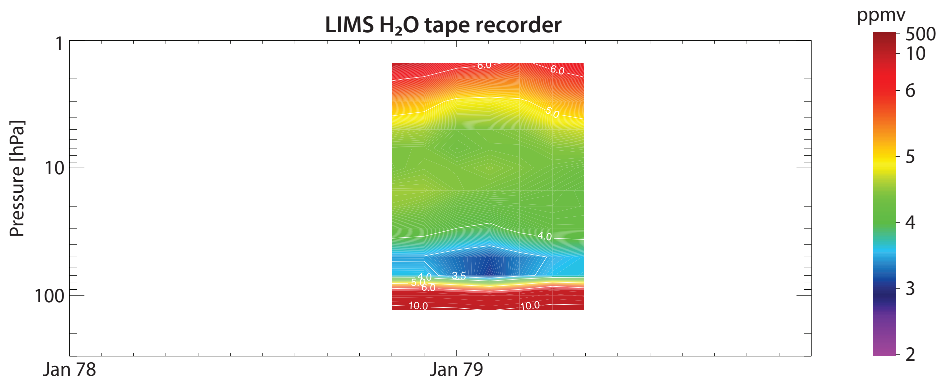


Figure 4.2.16b: Differences for H_2O tape recorder. *Altitude-time evolutions of H_2O differences relative to the MIM are shown for the time period 2000-2010 and each individual instrument (same ordering as in Figure 4.2.16a). Contour levels (2.5, 3, 3.5, 4, 5, 6, 8, 10, 50, 100 ppmv, with the 3-ppmv isopleths labelled) reproduce the MIM from Figure 4.2.16a.*

Figure 4.2.17: Tape recorder from LIMS H_2O measurements. *Note, the available time period does not span a full annual cycle.*



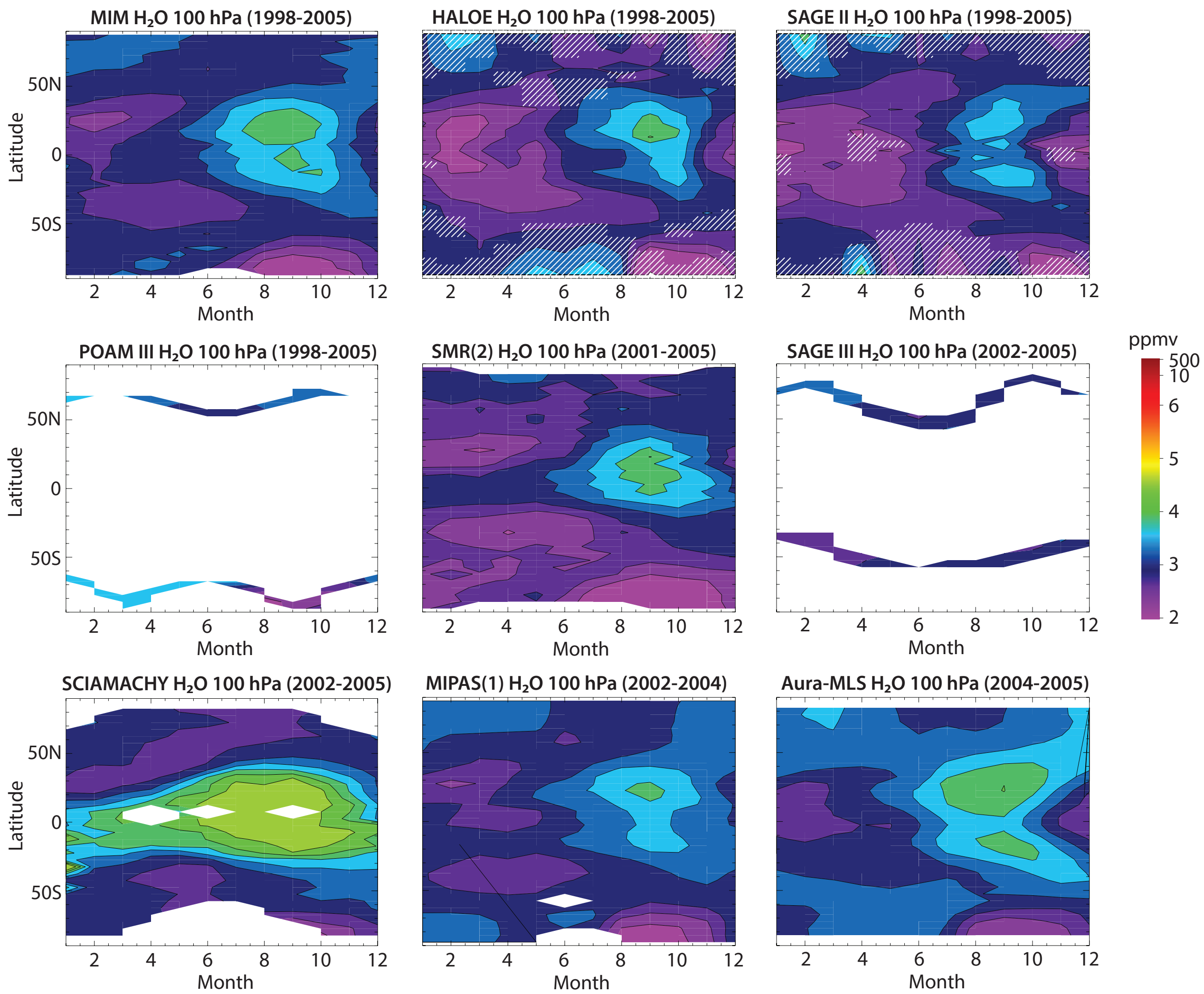


Figure 4.2.18: The horizontal tape recorder during 1998-2005. *Shown is a latitude-time evolution of H_2O at 100 hPa averaged over this period (or periods within this timeframe as indicated in the panel headers). HALOE and SAGE II show interpolated data; white hatching indicates the areas where no data was available.*

structures in H_2O better when compared to the MIM. Note that the feature derived from the solar occultation instruments would show better coverage when shown in equivalent latitude, however they are still useful to judge differences in absolute values between the individual instruments. POAM III measurements show slightly higher values than the other instruments particularly in the Southern Hemisphere, while SAGE III seems to agree better with MIPAS(1) than Aura-MLS. Most instruments with sufficient latitude coverage capture the Antarctic polar vortex dehydration between July and December although to a different extent.

4.2.6 H_2O evaluations: Polar vortex dehydration

Another spatio-temporal pattern that is seen in H_2O is the descent of aged and H_2O-enriched air masses and subsequent dehydration in the polar vortex of the Southern Hemisphere. Since this phenomenon predominantly happen in winter/early spring, occultation instruments will obviously not capture its full extent. However, for satellite instruments, which are measuring in darkness, the evaluation provides a stringent test of whether the retrieval in this region is being hampered by the presence of ice particles. The time period 2002-2009 has been chosen, since it encompasses most of the satellite instruments used in this study and allows for the evaluation of interannual variability in this region.

The only additional instrument to be tested is UARS-MLS, which is depicted in **Figure 4.2.20**. Since the simultaneous measurements from SAGE II and HALOE were strongly impacted by the Pinatubo aerosol, no ideal comparison can be made. However, it can be stated that UARS-MLS measures polar vortex H_2O in a physically plausible way. MS values seem rather on the low side compared to later years (see **Figure 4.2.21a**), which is consistent with our results from the annual zonal mean cross sections showing a general low bias in this instrument at these altitudes.

Figures 4.2.21a and **b** show the absolute values within the South polar vortex region averaged over 60°S-90°S and their differences to the MIM, respectively, between 2002 and 2010. Air masses containing more H_2O descend in branches from the upper stratosphere starting in autumn (March), and undergo dehydration during the winter months at lower

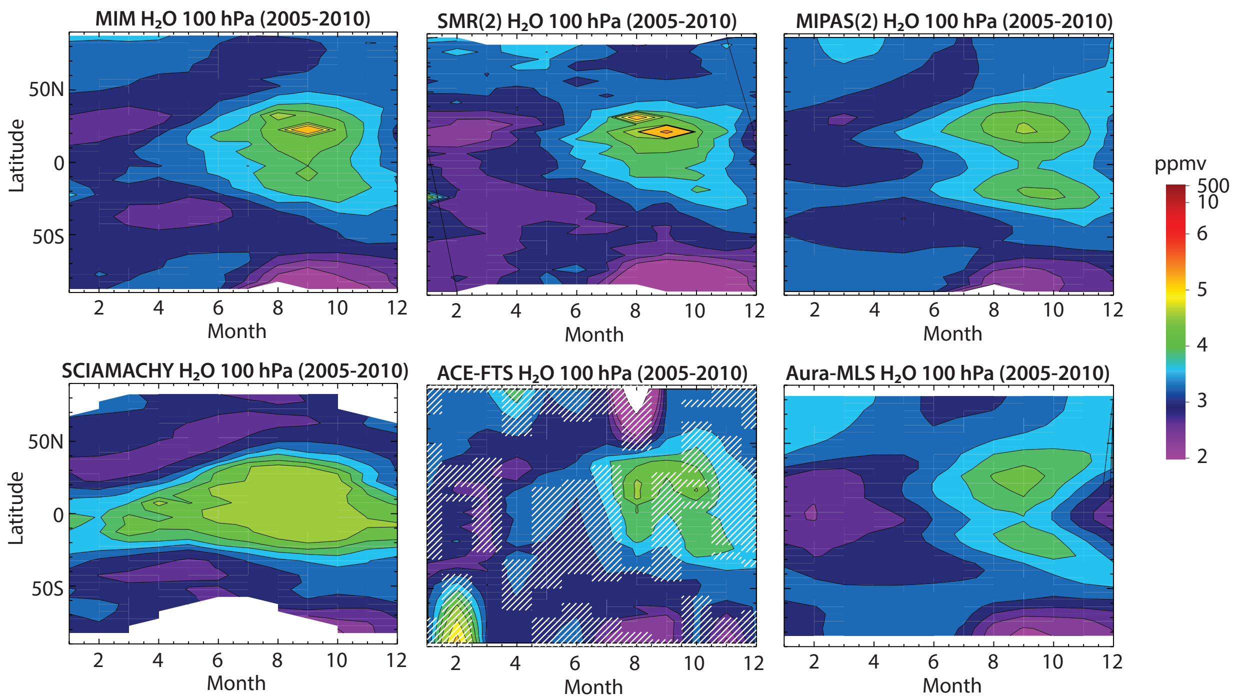


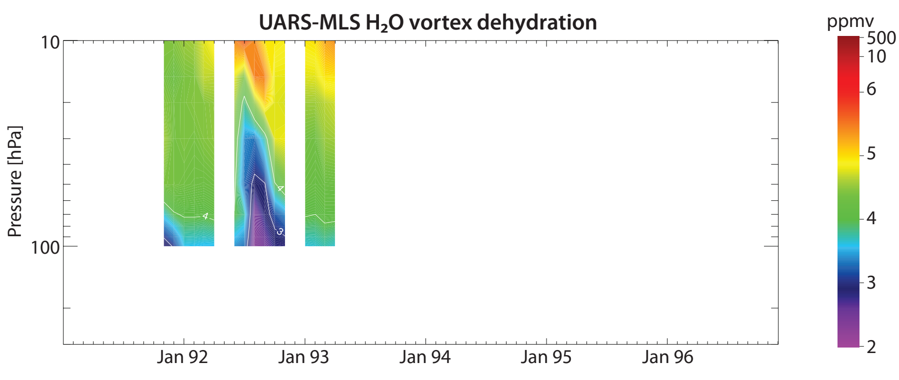
***Figure 4.2.19: The horizontal tape recorder during 2005-2010.** Shown is a latitude-time evolution of H_2O at 100 hPa averaged over this period. ACE-FTS shows interpolated data; white hatching indicates the areas where no data was available. Note the differences in Aura-MLS, SMR(2), and SCIAMACHY when compared to the earlier time period (Figure 4.2.18).*

altitudes (July-September). The most comprehensive results are obtained from Aura-MLS and MIPAS, two emission sounders, which are able to measure H_2O also during polar night. Note in this evaluation MIPAS(1) and MIPAS(2) are shown in the same panel. Many of the solar occultation results are showing the right physical structure, however, the less frequent sampling limits the overall picture. Nevertheless, SAGE II, HALOE and the ACE-FTS show mostly good agreement with the other instruments. Note that POAM III exhibits a better sampling of the polar region (see **Figure 2.7**). Nevertheless, POAM III shows larger deviations from the MIM than the previously mentioned solar occultation instruments. SMR(2) shows much too low values and too prominent dehydration structures that extent into the January-April period (compare also **Figures 4.2.18** and **4.2.19**). SMR(1) on the other hand, performs well for the higher altitudes, although it exhibits a little lower mixing ratios as MIPAS and Aura-MLS. SCIAMACHY shows consistent features, but does not capture the strength of the events. This is most probably due to the fact that only measurements at SZAs smaller than 85° were used to construct the SCIAMACHY H_2O climatologies, limiting its sampling to the outer parts of the polar vortex. Aura-MLS shows relatively strong negative deviations from the MIM around 200 hPa, but agrees well with POAM III.

4.2.7 H_2O evaluations: Interannual variability

In addition to the seasonal cycle in water vapour, which is driven by the solar forcing and discussed in *Section 4.2.3*, water vapour is characterised by non-seasonal variations related to ENSO and the QBO [*e.g.*, *Niwano et al.*, 2003; *Randel et al.*, 2004], and to a smaller extent by interannual variability in tropical convection or polar vortex temperatures. Long-term variability involves changes in methane, a source for water vapour in the stratosphere, and decadal variability. The evaluation of interannual variability using deseasonalised anomalies yields insight into whether an instrument's record produces physically consistent time series in comparison to other datasets. While the longer-term evolution of the anomalies is expected to be consistent

***Figure 4.2.20: Polar vortex dehydration.** Southern Hemisphere polar vortex descent and dehydration as observed in UARS-MLS in the Antarctic polar vortex 60°S-90°S between 1991 and 1993.*



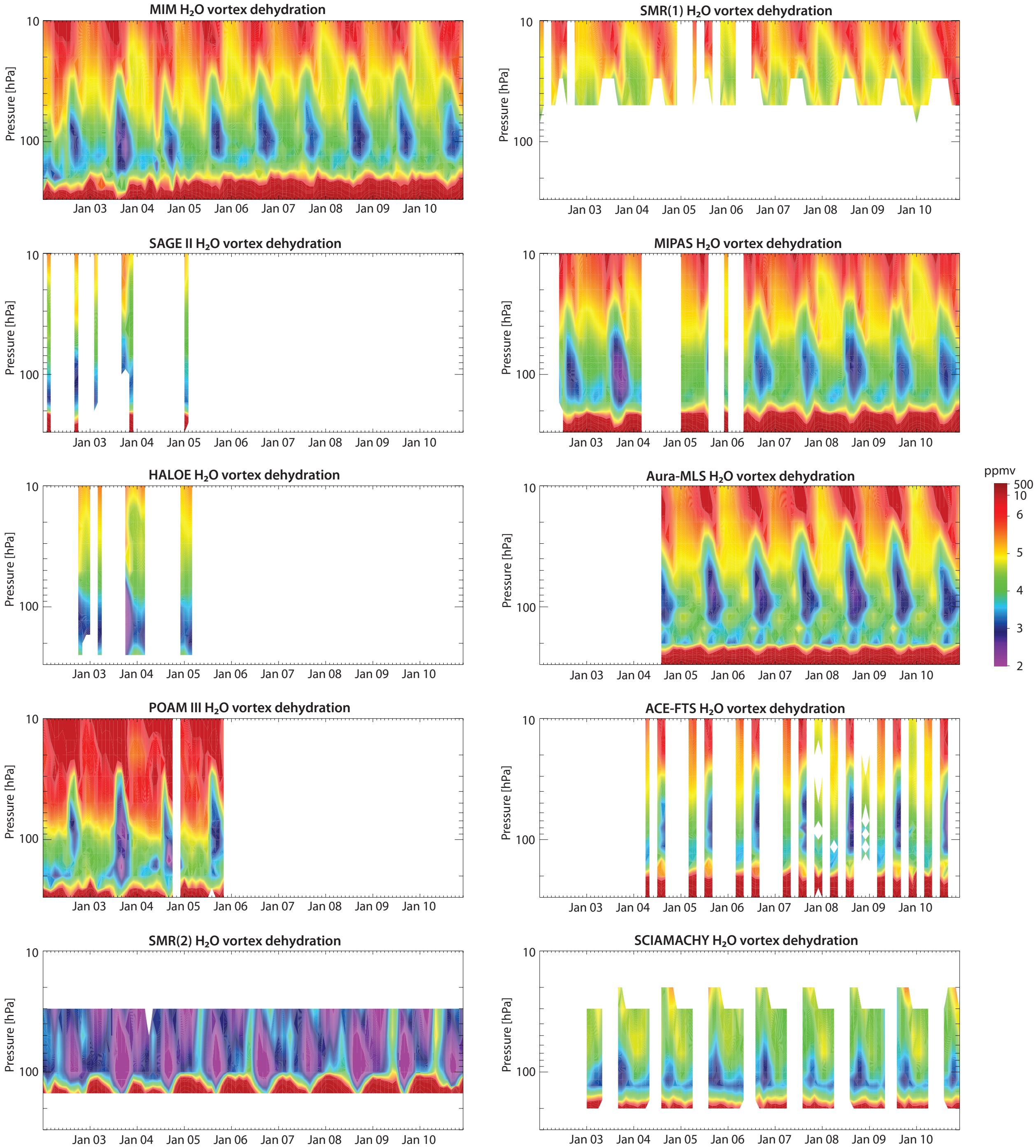


Figure 4.2.21a: Polar vortex dehydration. *The altitude-time evolution of Antarctic polar vortex descent and dehydration between 2002 and 2010 is shown for individual instruments and the MIM (uppermost left panel) using H_2O averaged over 60°S to 90°S. Note that SMR(2) is not included in the MIM.*

between the instruments, monthly differences are likely to be introduced by noise or sampling issues.

Figure 4.2.22 shows time series of deseasonalised H_2O anomalies at 80 hPa in the tropics, and at 100 and 10 hPa in the Northern extra-tropics between 1997 and 2010. See *Section 3.3.4* for the method used to calculate the anomaly time series. We start the evaluation in 1997, beyond Pinatubo's effect on the HALOE and SAGE II time series. The different instruments show very good agreement with generally consistent long-term tendencies and the QBO leaving the most pronounced signature in the anomalies. Note that while the QBO is a tropical phenomenon, it has also a distinct influence on extra-tropical water vapour, although with a somewhat attenuated signal due to mixing processes, which also shows a delay compared to the tropical signal related to stratospheric transport time scales. It is noteworthy that the instruments also agree on the breakdown pattern of the QBO signal on the tropical 80 hPa and the extra-tropical 100 hPa levels after 2008, as in the early 2000's.

In the tropics at 80 hPa, the evaluation reveals that compared to SAGE II, HALOE exhibits somewhat higher anomalies in

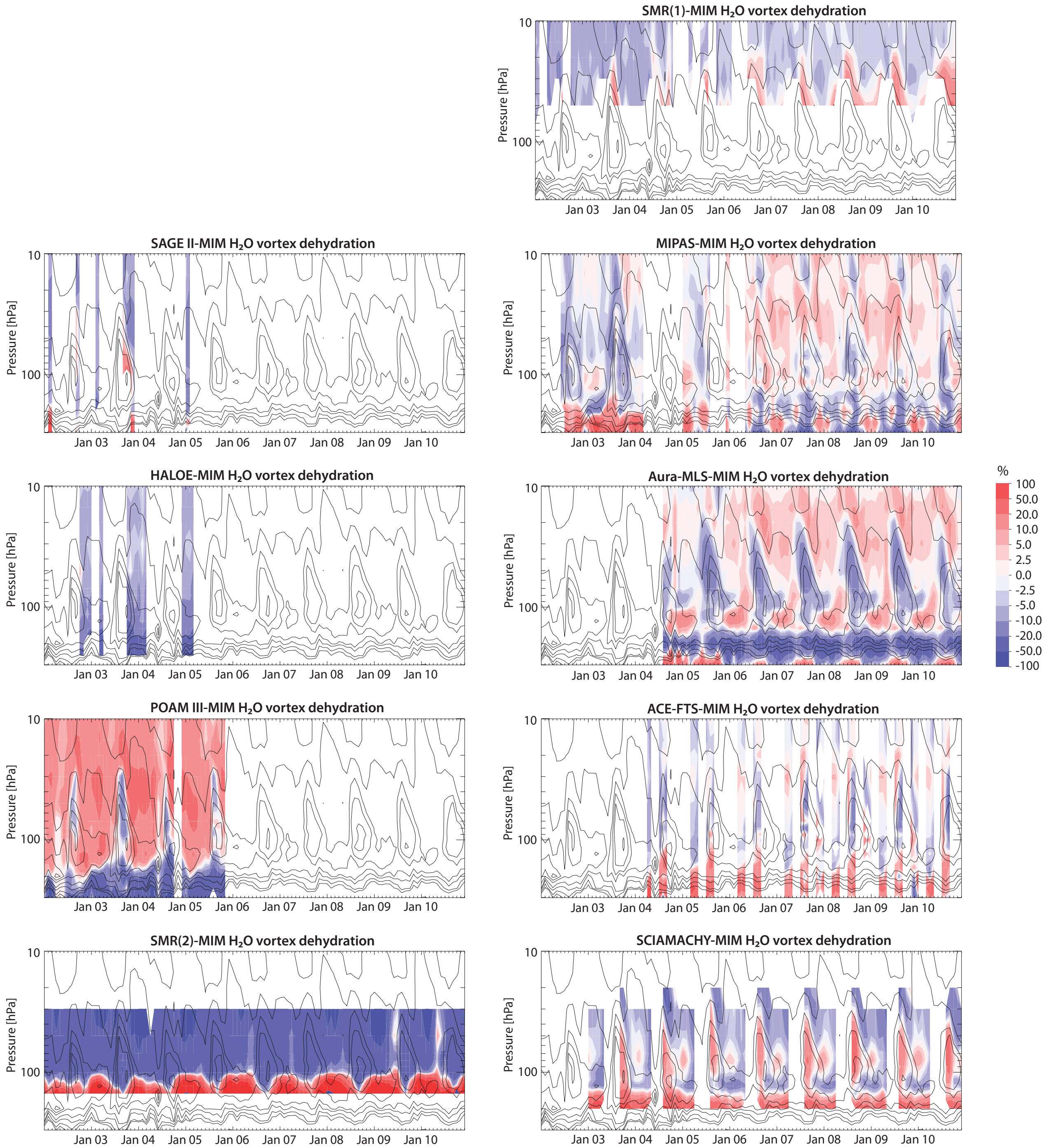


Figure 4.2.21b: Differences for polar vortex dehydration. *The time-altitude evolution of H_2O differences relative to the MIM between 2002 and 2010 is shown for each individual instrument (same ordering as in Figure 4.2.21a). Contour levels (2.5, 3, 3.5, 4, 5, 6, 8, 10, 50, 100 ppmv, with the 3-ppmv isopleths labelled) reproduce the MIM from Figure 4.2.21a.*

the early part (1997-1999), but somewhat lower anomalies in the later part of the record (2003-2005). As mentioned earlier, this relative drift may be caused by a more limited sampling of HALOE (or SAGE II for that matter) towards the end of the instrument's time series. The SMR(2) time series is characterised by some spike-like structures, which are not found in the other instruments after 2007. SCIAMACHY and Aura-MLS on the other hand agree very well in the amplitude of the QBO signal and also the month-to-month fluctuations, while MIPAS(1) and (2) show a somewhat smaller QBO signal with similar month-to-month variations. This issue is consistent with the evaluation of tropical seasonal cycles and is explained in more detail in *Section 4.2.3*. The ACE-FTS agrees fairly well with MIPAS and Aura-MLS, although its very infrequent tropical sampling does not allow definitive conclusions and produces some outliers, which most likely are attributable to sampling.

In the extra-tropics, HALOE and SAGE II agree very well on the anomalies, with POAM III confirming the magnitude of the variability at both 10 and 100 hPa. SMR(2) exhibits even more noise at 100 hPa in the extra-tropics (despite its good performance in the mean seasonal cycle at this level) and is hence not shown. SAGE III follows the mean behaviour well, however starts slightly at too positive

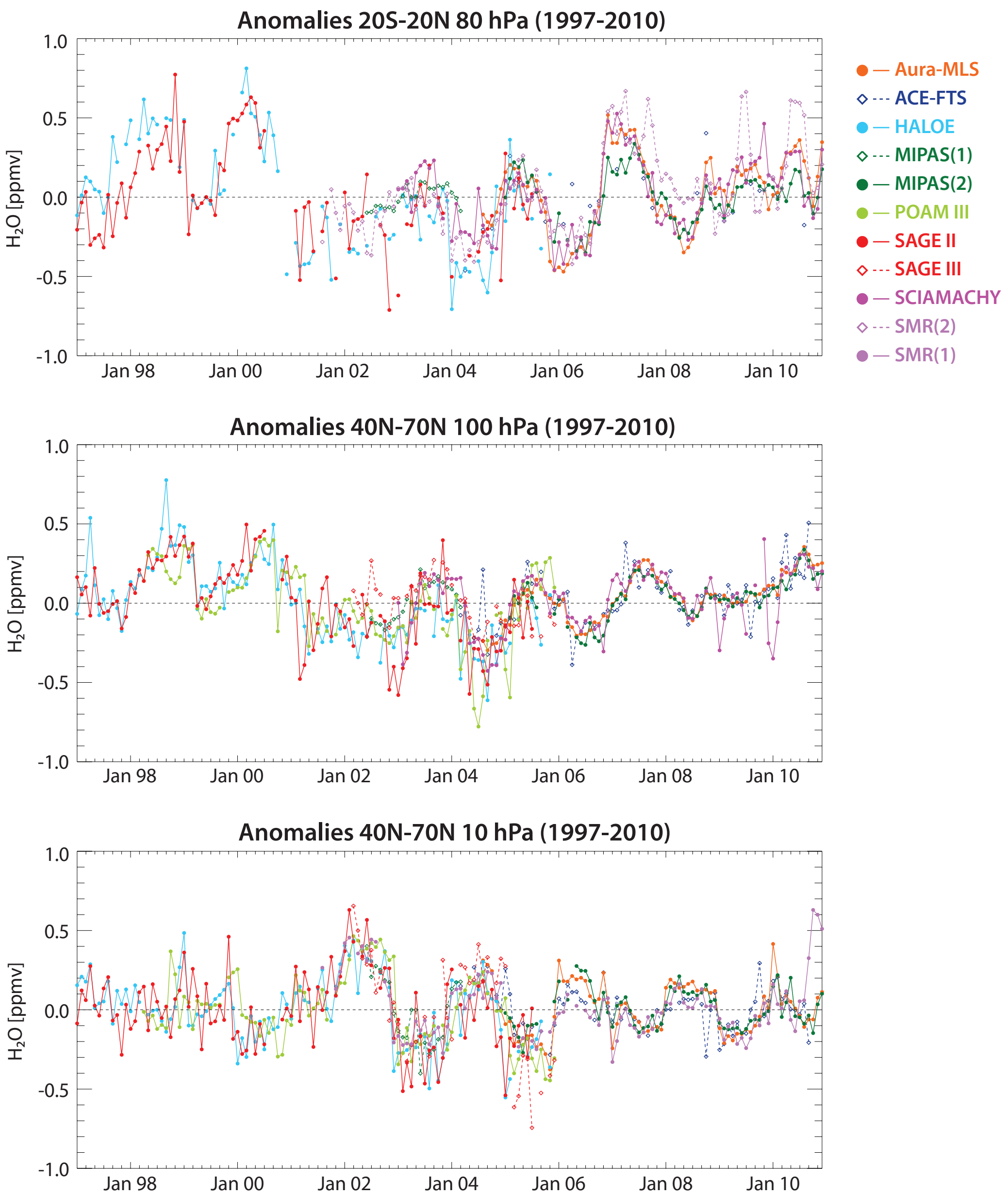


***Figure 4.2.22:** **Time series of deseasonalised anomalies of H_2O for 1997-2010.** Time series of deseasonalised anomalies in H_2O at 80 hPa between 20°S and 20°N (upper panel), and at 100 hPa (middle panel) and 10 hPa (lower panel) between 40°N and 70°N, respectively.*

anomalies at 100 hPa or ends at too negative anomalies at 10 hPa in the extra-tropics indicating a potential sampling issue (or drift) in the instrument. SCIAMACHY shows a somewhat noisier field or month-to-month fluctuations after 2008. ACE-FTS has a better sampling coverage in the extra-tropics, and the anomalies show here very similar behaviour to Aura-MLS and MIPAS(2), although with a somewhat smaller amplitude at 10 hPa. SMR(1) at 10 hPa shows also good agreement with these latter instruments (except during 2010), but similar to ACE-FTS exhibits a somewhat too low amplitude in the anomalies.

4.2.8 Summary and conclusions: H_2O

In this report, we assessed the quality of 13 water vapour products from 11 different limb-viewing satellite instruments (LIMS, SAGE II, UARS-MLS, HALOE, POAM III, SMR, SAGE III, MIPAS, SCIAMACHY, ACE-FTS, and Aura-MLS) which provide measurements over the time period from 1978 to 2010 (see **Table 4.2.1**). Overall findings on the water vapour annual mean state and important characteristics of the individual datasets are discussed below. Two summary plots are provided. The first (**Figure 4.2.23**) aims to provide information on our current estimate of the water vapour annual mean state and its overall uncertainty as derived from the spread between the different datasets as a function of latitude and altitude. The second figure (**Figure 4.2.24**) aims to summarise the specific inter-instrument differences, which are expressed through the median (or mean) deviation from the MIM of each instrument averaged over a particular region, together with the spatial homogeneity (or smoothness) of that deviation, expressed as the MAD (or standard deviation). Note that both pieces of information (average deviation from the MIM and spatial variability of that deviation) are important for a meaningful assessment of inter-instrument differences. See *Section 3.3.5* for more detailed information on the summary plots.

The comprehensive comparison of H_2O climatologies from the different available limb-viewing satellite instruments results in the following summary and conclusions on the atmospheric mean state, performance by region, and performance of individual instruments.

Atmospheric mean state

- Our knowledge of the atmospheric mean state in H_2O derived from the full set of instruments available between 1998 and 2008 (excluding SMR(2) and MIPAS(1)) is best in the lower and middle stratosphere tropics and mid-latitudes, with a relative uncertainty of ±2-6% (1σ) (**Figure 4.2.23**).
- The relative uncertainty (1σ) in the atmospheric mean state in H_2O (1998-2008) increases toward the polar latitudes (±10% and 15% for NH and SH, respectively), the lower mesosphere (±15%) and the troposphere (±30-50%). Note that the uncertainty in H_2O is largest in the subtropical jet region (30-50°N/S), partly due to a large dynamical variability in tropopause height, which affects the climatologies due to sampling issues (**Figure 4.2.23**).
- The minimum in the annual zonal mean of H_2O found just above the tropical tropopause shows values ranging from approximately 2.5 to 4.5 ppmv when including all instruments, with a mean of 3.5±0.5 ppmv (or ±14%, 1σ-uncertainty) (**Figure 4.2.23**). The 1σ uncertainty is somewhat larger (15-20%) when looking at individual months (see seasonal cycle evaluation **Figure 4.2.14**).
- The maximum found in the annual zonal mean of H_2O in the lower mesosphere shows an absolute range of approximately 5.5-7.5 ppmv, with a mean of 6.5 ± 0.7 ppmv (or ±9%, 1σ-uncertainty) (**Figure 4.2.23**).

Performance by region

Lower Mesosphere (0.1-1 hPa)

In the tropical and extra-tropical LM, the instruments agree well, within approximately ±10% of the MIM (corresponding to inter-instrument differences of up to 20%). The newer set of instruments (ACE-FTS, Aura-MLS, and MIPAS(1) and (2)) even show excellent agreement, within 5% of each other. A clear exception to this is SMR(1), which shows deviations from the MIM of up to 18%. Together with the older instruments HALOE and UARS-MLS, SMR(1) is on the low side of the MIM. Earlier results from validation studies using coincident measurements from other independent instruments support these findings: UARS-MLS was found to have a low bias of 5% when compared to the ATMOS instrument (and HALOE) [*Pumphrey et al.*, 1999]. Note that the spatial variability of the deviations within one region is relatively small for most instruments, indicated by small MADs (around ±3%), POAM III shows a larger range, indicated by a larger MAD (±6%) (**Figure 4.2.24**).

Upper Stratosphere (1-5 hPa)

In the tropical and extra-tropical US, the instruments show a good agreement, within ±10% of the MIM, and very small MADs (±1.5%) for most instruments indicating a narrow distribution of deviations from the MIM within these regions. This means that while individual instruments may disagree with each other, their differences are well defined. Most instruments agree even very well, within ±5%. Exceptions in the tropical region are UARS-MLS and SMR(1), which show larger negative deviations, and MIPAS (2), which shows a larger positive deviation from the MIM than the other instruments. Exceptions in the extra-tropical regions are LIMS, SMR(1), and UARS-MLS. POAM III data in the extra-tropics show the highest values, although close to those from MIPAS (1) and (2) (**Figure 4.2.24**).

Middle Stratosphere (5-30 hPa)

In both the tropical and extra-tropical MS, most instruments agree very well to within ±5% of the MIM. Notable is the excellent agreement (within ±2.5%) between ACE-FTS, Aura-MLS, HALOE, LIMS, MIPAS (1) and MIPAS(2) in the extra-tropics. Small MADs (mostly ±3 to ±4%) indicate small variability in the deviations and hence that the instrument differences are well defined. Exceptions are ACE-FTS, LIMS, and SCIAMACHY in the tropics, and POAM III and SCIAMACHY in the extra-tropics (**Figure 4.2.24**).

Lower Stratosphere (30-100 hPa)

In the tropical LS, the instruments show only reasonably good agreement, mostly within ±20% of the MIM. The agreement is much better in the extra-tropical LS with, deviations of only ±5% of the MIM. Exceptions are LIMS, POAM III and UARS-MLS with deviations of ±10% of the MIM, and SMR(2) with a deviation of -22% from the MIM. Very good agreement is found for the ACE-FTS, Aura-MLS, HALOE, MIPAS(1), MIPAS(2), SAGE II, SAGE III, SCIAMACHY, and SMR(1). The instruments' MADs indicate better defined deviations in the extra-tropics than in the tropics (**Figure 4.2.24**).

Upper Troposphere/Lower Stratosphere (100-300 hPa)

Considerable disagreement between the instruments is found for the lowest levels between 100 and 300 hPa of both the tropical and extra-tropical UTLS, with differences from the MIM of ±40% in the tropics and 30% in the extra-tropics. Nevertheless, very good agreement within ±5% of the MIM is found for Aura-MLS, MIPAS(1), MIPAS(2), POAM III, and SAGE III in the extra-tropics. Large MADs (±10% or more) indicate spatial inhomogeneity of the deviations in the two regions and hence not well defined instrument behaviour. Note SMR(2) shows deviations from the

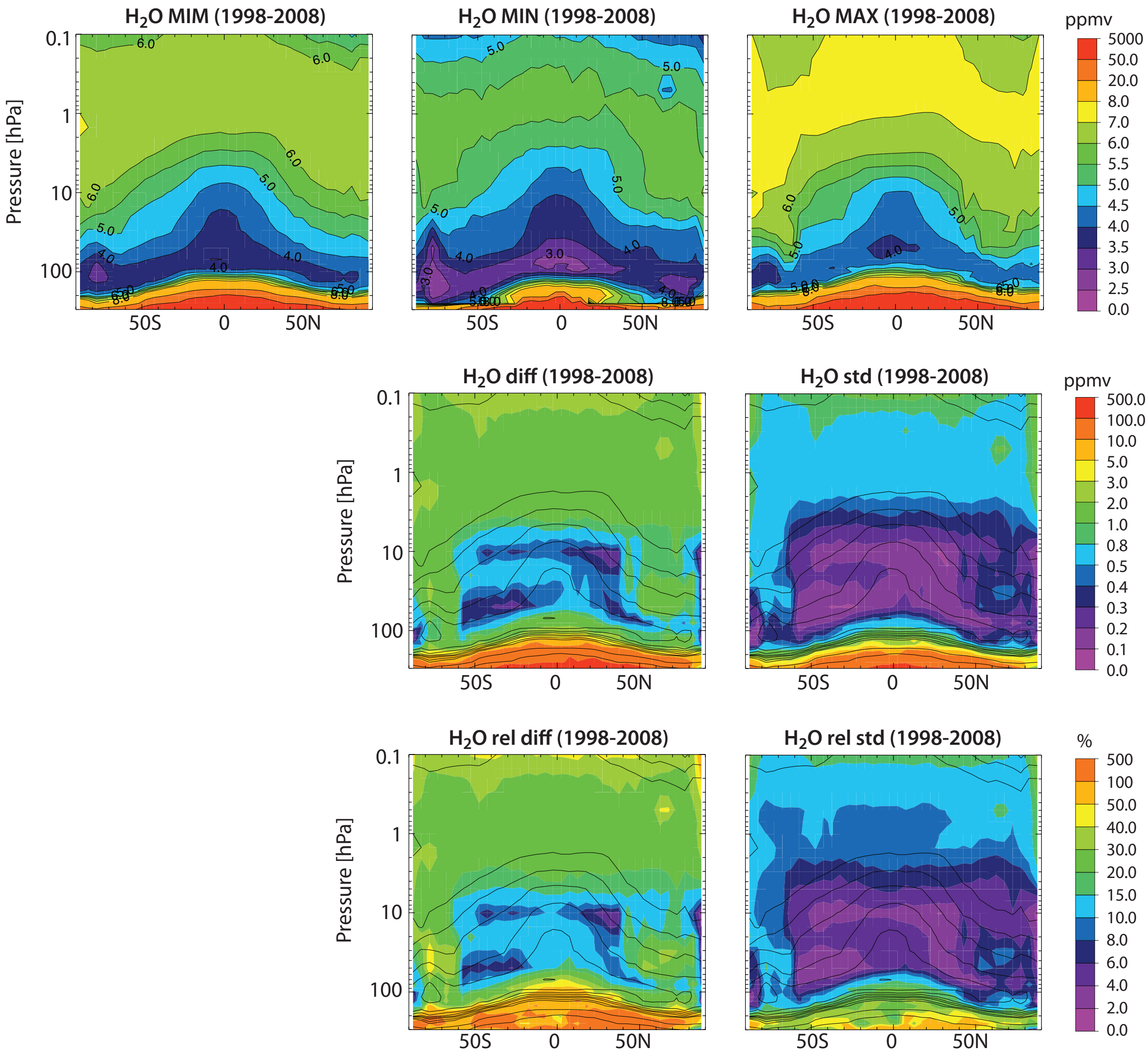


__Figure 4.2.23: Summary of H_2O annual zonal mean state for 1998-2008.__ Shown are the annual zonal mean cross sections of the MIM, minimum (MIN) and maximum (MAX) H_2O values (upper row), the absolute differences (MAX-MIN) and absolute standard deviations (middle row), and relative differences and relative standard deviations with respect to the MIM (lower row). Black contour lines in the lower panels repeat the MIM distribution. Instruments considered are SAGE II, SAGE III, HALOE, POAM III, ACE-FTS, Aura-MLS, MIPAS, SAGE III, SMR(1), and SCIAMACHY.

MIM of more than +50%, and its use is not recommended below 100 hPa. The poor agreement in the UTLS may partly be explained by sampling issues and partly by the difficulties the instruments encounter to measure accurately in the UTLS. Large dynamical variability and steep gradients across the tropopause limit especially instruments with low temporal (occultation sounders) or vertical resolution (emission sounders). Also, cloud interference and saturation of the measured radiances pose challenges to the instruments depending on the measurement mode applied.

Instrument-specific conclusions

LIMS (V6.0) provides the earliest H_2O observations available to the SPARC Data Initiative. The LIMS record extends over only a few months. Using SAGE II as transfer, LIMS shows very good agreement, within ±5% of the MIM, in the MS and the tropical US, however large negative deviation from the MIM of around -12% in the extra-tropical US, and large positive deviations from the MIM of +15% in the LS and +30 to +40% in the UTLS (between 100 and 300 hPa), respectively.

SAGE II (V6.2) provides the longest H_2O record. Evaluations of the data indicate a low bias when compared to the newer generation of instruments. This fact may be explained by the chosen retrieval channel, which was switched from 935 nm to 945 nm, to better agree with HALOE data. The shift was necessary since the first channel experienced a drift [*Thomason et. al.*, 2004], although the exact nature of the shift and when it happened could not be established. However, in this study SAGE II V6.2 is shown to perform very well in interannual variability evaluations, and may therefore be useful for data merging activities. Above 3 hPa, SAGE II exhibits a known bias, and so the data above this level are not included in the SPARC Data Initiative monthly zonal mean climatologies. Note that a newer version of SAGE II (V7.0) has become available, which improves on the main issues identified in V6.2 [*Damadeo et al.*, 2013], and is beneficial for data merging [*Hegglin et al.*, 2014].

HALOE (V19) is the most used H_2O dataset. Our evaluations indicate that the instrument's H_2O has a slight low bias throughout the atmosphere. Deviations from the MIM are found to be around -5% through most of the stratosphere and LM consistent with results from SPARC [2000]. HALOE's low bias strongly increases in the UTLS (between 100 and 300 hPa) to values larger than -20%, and the instrument fails at reproducing the seasonal cycles at the 200 hPa level and at lower altitudes in both the tropics and the extra-tropics. However, note that HALOE resolves the seasonal cycle and interannual variability well down to levels above 200 hPa after bias-elimination.

UARS-MLS (V6) offers H_2O measurements over a limited time period in the early 1990's. The measurements are seen to be about 5% lower than HALOE through most of the atmosphere, a result confirmed by validation with *in situ* measurements.

SAGE III (V4.0) is limited to the extra-tropics, however shows excellent agreement with the MIM throughout the atmosphere and even in the UTLS (between 100-300 hPa). While its limited availability restricts its use to a small number of evaluations, it may be considered for use in merging activities.

POAM III (V4.0) is another instrument with a somewhat limited temporal and spatial coverage. The biases derived in our evaluations are consistent with earlier validation studies. POAM III is biased high throughout the stratosphere with somewhat larger deviations from the MIM in the SH (>20%) than in the NH (>10%). However, it performs very well (within 5% from the MIM) at the lowest levels (100-300 hPa). Despite the positive biases, the instrument performs well in evaluations of interannual variability, and compares well to SAGE II and HALOE, making it a potentially useful instrument to study climate variability or to merge HALOE and SAGE II with the newer instruments.

The **SMR(2)** (V2.0) H_2O product (derived using the 544 GHz-band) does not exhibit a correct tropopause-following structure of the trace gas isopleths and the values are too high below and too low above 100 hPa, respectively. Nevertheless, once the bias is removed, SMR(2) exhibits a reasonably good interannual variability in the tropics and also shows a tropical seasonal cycle that agrees well with the MIM. However, the data are less consistent in the extra-tropics. The data product needs further improvement and the recommendation is to restrict its use to between 50 and 100 hPa. SMR(1) (V2.1) provides reasonably good data in the MS (also showing physically consistent interannual variability), while strong negative deviations from the other instruments are found in the USLM. This issue is known and has been related to an imperfect sideband correction of the 488.9 GHz water band.

MIPAS(1) (V13) and **MIPAS(2)** (V220) compare very well to the MIM with deviations from the MIM mostly within ±5% throughout the atmosphere. An exception is the tropical UTLS (100-300 hPa), where deviations for MIPAS(1) and MIPAS(2) increase to -25% and -10%, respectively. The seasonal cycle and interannual variability in the tropical tropopause region exhibit a too low amplitude, which can be explained by a state-dependent averaging kernel. The two data versions agree with each other mostly within a few percent. Exceptions are the UTLS (100-300 hPa), and the tropical LS and US, where MIPAS(1) is about 10% lower than MIPAS(2).

Aura-MLS (V3.3) shows very good to excellent agreement with the MIM throughout most of the atmosphere (with deviations from the MIM between +2.5 and +5%). Exceptions are found in the LM, where the deviations increase to +10%. Good spatial and temporal coverage (also long-term) allow generally a robust assessment of the Aura-MLS deviations from the MIM (except in the UTLS), which makes the data exceptionally useful for data merging.

ACE-FTS (V2.2) performs exceptionally well compared to the MIM in both the tropical and extra-tropical stratosphere, and to a somewhat lesser extent in the LM, despite its disadvantage of being an occultation sounder with small temporal and spatial sampling. The deviations from the MIM are mostly consistent with validation results using coincident measurements. In the UTLS between 100 and 300 hPa, the deviations from the MIM increase to +10% in the extra-tropics and +35% in the tropics, respectively, some of which is likely attributable to limited sampling.

SCIAMACHY (V3.0) H_2O (a relatively new retrieval product) provides promising results, however suffers from a relatively coarse vertical resolution in the UTLS, which leads to smearing of the strong gradients found across the tropopause when interpolating the data onto the SPARC Data Initiative pressure grid. The smearing affects mainly the H_2O mean values, however does not compromise evaluations of interannual variability or amplitudes in H_2O seasonal cycles in this region.

4.2.9 Recommendations: H_2O

- Our evaluations show that most instruments exhibit very good agreement regarding the magnitude and structure of interannual variability in the different regions of the atmosphere (once the instruments' biases are removed), therefore fulfilling a necessary prerequisite that the use of the data for studies of climate variability can be recommended.
- Our findings indicate that our knowledge on the H_2O atmospheric mean state is still unsatisfactory, especially in the tropical UT and LS (300-30 hPa), emphasising the need for limb-sounders with higher quality and vertical

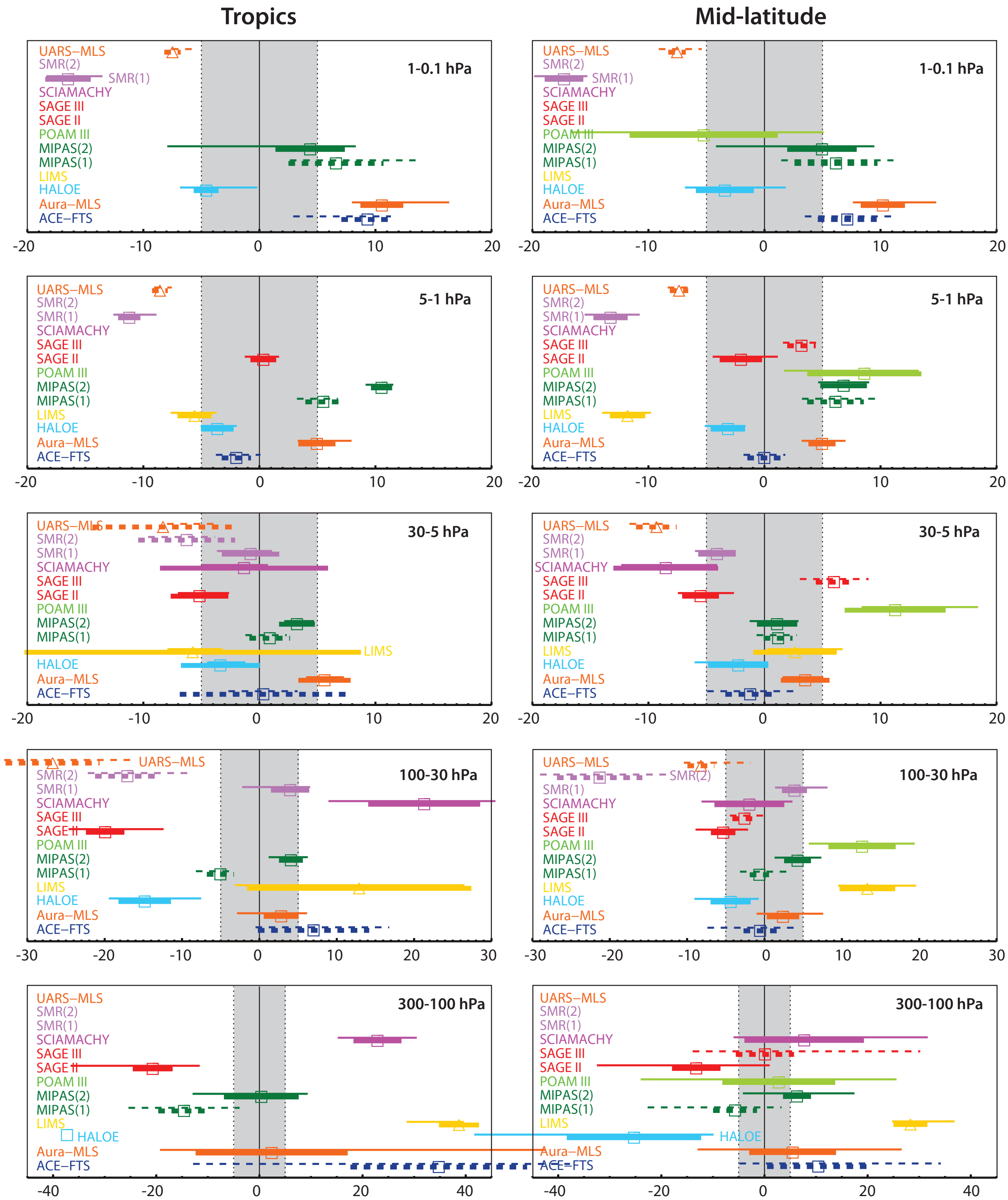


Figure 4.2.24: Summary of inter-instrument differences in H_2O for 1998-2008. *Results are calculated for the tropics 20°S-20°N (left) and extra-tropics 40°S-80°S and 40°N-80°N (right) and for 5 different altitude regions from the UT up to the LM between 300 and 0.1 hPa as defined in Table 0.1. Shown are the median (squares), median absolute deviations (MAD, thick lines), and the mean ±1σ ranges (thin lines) of the relative differences between each individual instrument and the MIM averaged over a given latitude and altitude region. The period of reference is 1998-2008 and the results are directly comparable to the evaluations in Section 4.2.2. Triangles indicate medians of instruments that are obtained outside of the reference period, here LIMS and UARS-MLS, shown with respect to the instrument means of SAGE II and HALOE based on comparisons for 1978-1990 and 1991-1993, respectively.*

resolution, but also for *in-situ* correlative measurements that help validate them.

- The excellent agreement that is typically observed between Aura-MLS, MIPAS(1), MIPAS(2) and ACE-FTS indicates their potential for use in extending the HALOE time series in merging activities. Note that the merging of MIPAS(1) and MIPAS(2) needs to address potential biases between these two datasets in the tropical UTLS (300-100 hPa), LS and US.
- HALOE has been the most frequently used H_2O record up to date. Based on our evaluations, HALOE data show a consistent, but small negative deviation from the MIM of around -2.5 to -5% throughout the atmosphere, for which the user should account for in merging activities and trend studies. This negative deviation increases in the tropical LS to -15%. HALOE data should furthermore be used with care at altitudes below 100 hPa, where the negative biases strongly increase (to values up to -50%). However, the seasonal cycles and interannual variability are nevertheless well resolved at all altitudes above the 200 hPa level.
- In the extra-tropical UTLS, between 100-300 hPa, Aura-MLS, MIPAS(1), MIPAS(2), POAM III and SAGE III are producing consistent results. Both POAM III and SAGE III may be used as transfers between the earlier and the newer sets of satellite instruments.
- The H_2O datasets evaluated here show great potential for improving past model-measurement comparisons. However, careful choices have to be made when choosing instruments to be included in a metric depending on the region of the atmosphere:

  i. Seasonal cycles in H_2O in the UTLS are often used for classic model-measurement comparisons [*Gettelman et al.*, 2010; *Hegglin et al.*, 2010]. While there are still considerable uncertainties in the monthly mean values, which may partly be addressed by accounting for sampling issues, the combined measurements will yield better constraints on amplitude and phase of the seasonal cycles in both the tropics and extra-tropics.
  ii. The derivation of the tape recorder's amplitude and phase, another classical model diagnostic (see *SPARC* [2010]), can be affected by the differences in the instruments' vertical resolutions. The effect of vertical resolution on these metrics should be explored in more detail before conclusions can be drawn on model behaviour.
  iii. We suggest using polar vortex dehydration (time-altitude cross sections) and the horizontal tape recorder (time-latitude cross sections) around 100 hPa as new (or improved) model diagnostics in future model-measurement comparisons.

4.3 Methane – CH_4

Methane (CH_4) is the most abundant hydrocarbon in the atmosphere. It is a very effective greenhouse gas and the second-largest contributor to anthropogenic radiative forcing since preindustrial times after CO_2. CH_4 affects stratospheric ozone chemistry and in the troposphere acts to reduce the atmosphere's oxidizing capacity. CH_4 is emitted by ruminants, from rice fields, waste management, fossil fuel production, and biomass burning, but also has natural sources that amount to about 30% of total emissions [*IPCC*, 2007]. CH_4 has a relatively short atmospheric lifetime of about 10 years and in the troposphere exhibits a strong seasonal cycle as well as a distinct gradient across the equator, similar to CO_2. CH_4 has been widely used to study stratospheric circulation and transport [*Jones and Pyle*, 1984; *Choi and Holton*, 1988; *Russell*, 1993; *Randel et al.*, 1998], and the available long-term measurements now are also used to deduce changes in the stratospheric circulation [*Remsberg*, 2015].

4.3.1 Availability of CH_4 measurements

The first vertically resolved satellite measurements of CH_4 available to the SPARC Data Initiative were made by HALOE in 1991. MIPAS started measuring CH_4 in 2002 providing nearly four years of overlap (although with a major gap in 2004). From 2004 onwards there are also ACE-FTS measurements available for comparison. Not available for the SPARC Data Initiative format and hence not included in the evaluations are CH_4 measurements from SAMS on Nimbus-7 (1979-1981; *Taylor* [1987]), ATMOS (since the mid-1980s; *Gunson et al.*, [1996]), ISAMS on UARS [*Taylor et al.*, 1993], and CLAES on UARS [*Roche et al.*, 1993].

Tables 4.3.1 and **4.3.2** compile information on the availability of CH_4 measurements, including data version, time period, vertical range and resolution, and references relevant for the data product used in this report.

4.3.2 CH_4 evaluations: Zonal mean cross sections, vertical and meridional profiles

Annual zonal mean cross sections for the years 2003-2006 are analysed to investigate mean biases between the various datasets. Additionally, vertical and meridional profiles are presented. We here use the average over the years 2003-2006 for comparison, since there was basically no trend in tropospheric CH_4 between 1998 and 2008 and averaging over 4 years of data will help smear out effects of the QBO. We avoid comparisons over single years, which suffer from other shortcomings. For example, HALOE is not measuring during all months of the year in 2005, which introduces a sampling bias.

HALOE, MIPAS, and ACE-FTS (2003-2006)

Annual zonal mean cross sections for CH_4 are shown in **Figure 4.3.1** along with the relative differences between the individual instruments and the MIM.

CH_4 concentrations decrease with increasing altitude in the atmosphere due to oxidative reaction of CH_4 with hydroxyl

Table 4.3.1: Available CH_4 measurement records from limb-sounding satellite instruments between 1978 and 2010. The red filling of the grid boxes indicates the temporal (January to December) and vertical coverage (300 to 0.1 hPa) of the respective instrument in a given year.

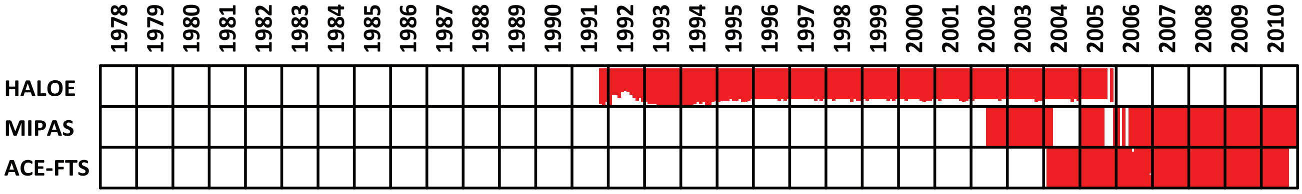


Table 4.3.2: Time period, vertical range, vertical resolution, references and other comments for CH_4 measurements.

Instrument	Time period	Vertical range	Vertical resolution	References	Additional comments
HALOE V19	Oct 91 – Nov 05	up to 80 km	3.5 km	*Grooß and Russell*, 2005 *Park et al.*, 1996	
ACE-FTS V2.2	Mar 04 –	5 km – 62 km	3 - 4 km	*De Mazière et al.*, 2008	
MIPAS MIPAS(1) V11 MIPAS(2) V220	Mar 02 – Mar 04 Jan 05 – Apr 12	Cloud top – 70 km	4 – 5 km 2 – 3.7 km	*Glatthor et al.*, 2005 *von Clarmann et al.*, 2009a	measurement mode switched in 2005 from high spectral to high vertical resolution

radicals (OH), which leads ultimately to the formation of H_2O and CO_2. The stratospheric CH_4 distribution nicely reflects the effects of the Brewer-Dobson circulation on tracers with a tropospheric source and a stratospheric sink, with upwelling of higher values in the tropical region and downwelling of lower values in the extra-tropics. As a result, CH_4 isopleths slope downward toward higher latitudes and follow the shape of the tropopause. The instruments agree fairly well on the overall distribution of CH_4.

HALOE and the ACE-FTS agree better in the UTLS, while the two MIPAS data versions show positive deviations from the MIM, with MIPAS(2) showing largest deviations of up to +10% around 100 hPa. The high bias in the lower atmosphere is a known feature in the MIPAS CH_4 data [*von Clarmann et al.*, 2009a]. Particularly the very high values above 2 ppmv in the tropical UT as seen in MIPAS are unrealistic, given that global tropospheric CH_4 (approximately equal to tropical concentrations) did not exceed 1.8 ppmv in the mid 2000s (source NOAA; see also *Isaksen et al.* [2009]).

Good agreement is found between all instruments in the tropical/subtropical MS with deviations from the MIM of up to ±10%. However, the values diverge towards the USLM, with HALOE largely on the low side and the ACE-FTS on the high side of the MIM. This finding is in agreement with the results from *De Mazière et al.* [2008] who used coincident profiles from HALOE to validate the ACE-FTS. MIPAS(1) in the LM seems closer to HALOE and MIPAS(2) closer to ACE-FTS. Also, towards higher latitudes, where natural variability becomes larger, the deviations from the MIM increase for all instruments. The monthly mean plots as presented in *Appendix A4.3* reveal somewhat less agreement between the instruments with deviations from the MIM reaching up to ±20% in certain regions. Nevertheless, the structures found in the instrument differences are similar for the monthly and annual means.

Monthly mean vertical CH_4 profiles in the polar regions and the Southern Hemisphere tropics are shown in **Figure 4.3.2** together with their differences from the MIM. The months shown have been chosen in order to include the maximum number of instruments possible, which depends on their latitudinal sampling. Also, summer and autumn months show the least variability in the stratosphere, which is important when trying to isolate the uncertainty of the measurement from that introduced by natural variability. The data are averaged over a few years in order to improve the spatial and temporal sampling between the instruments as mentioned above. The profiles indicate that HALOE and ACE-FTS mostly show very good agreement in the LS and MS, but that their values diverge during some months in the USLM. There are, however, other months/latitudes for which HALOE and ACE-FTS show larger disagreement than with MIPAS, so the result is not robust. It is noteworthy that both MIPAS data versions show some rather strong (but opposite) oscillations in these climatological profiles that may arise from the limited vertical resolution of the measurements.

Figure 4.3.3 shows the meridional zonal mean profiles of CH_4 and differences from the MIM. At 100 and 50 hPa, the meridional profiles indicate that MIPAS(2) is higher than the MIM and MIPAS(1) is approximately agreeing with the MIM, while both HALOE and ACE-FTS exhibit lower values, with very good agreement between each other. Note that the MIM at 100 hPa shows spikes, which are an artefact of the MIM consisting of different instruments at different latitudes and not due to one of the instruments showing such spikes.

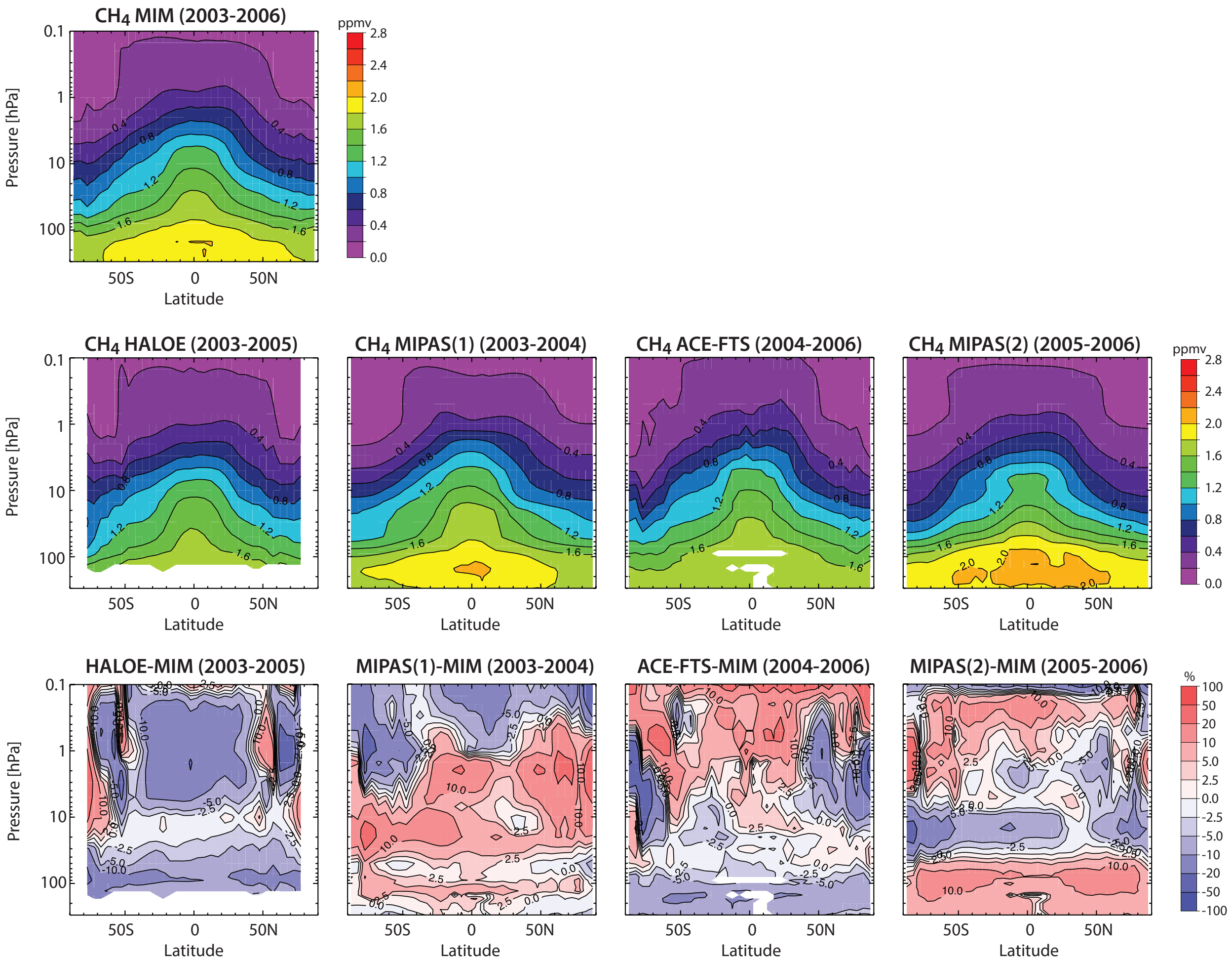


Figure 4.3.1: Cross sections of annual zonal mean CH_4 for 2003-2006. *Upper panel shows the CH_4 cross section for the MIM, middle panels show cross sections for the different instruments (HALOE, MIPAS(1), ACE-FTS, and MIPAS(2)), and lower panels show the relative differences between each instrument and the MIM.*

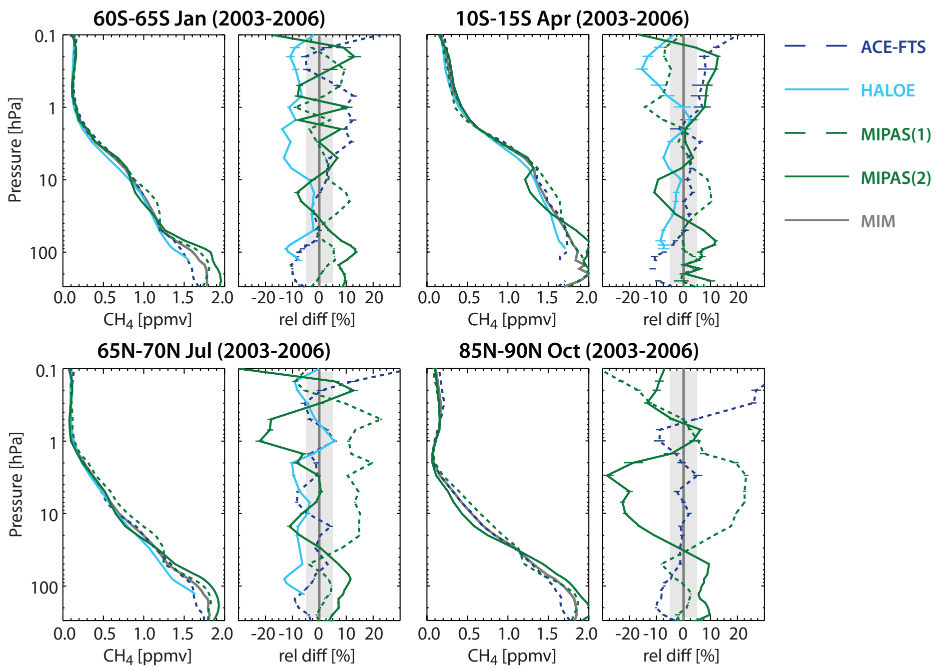


Figure 4.3.2: Vertical profiles of monthly zonal mean CH_4 for 2003-2006. *Vertical CH_4 profiles for 60°S-65°S January and 10°S-15°S April (upper panels) and for 65°N-70°N July and 85°N-90°N October (lower panels) are shown together with their differences from the MIM. HALOE, MIPAS(1), ACE-FTS, and MIPAS(2) data are averaged over the years 2003-2005, 2003-2004, 2004-2006, and 2005-2006, respectively, according to their availability within this time period.*

Otherwise the instruments agree well and lie approximately within the ±10% difference range, except at the highest altitude (the 5 hPa level) and winter high latitudes (here also at 10 hPa), where deviations are as large as ±30%.

4.3.3 CH_4 evaluations: Latitude-time evolution

The latitude-time evolution of CH_4 can be used to test the physical consistency of a particular dataset. **Figure 4.3.4a** shows multi-year climatologies of the latitude-time evolution of CH_4 for the different instruments at 2 and 10 hPa, where distinct features have been found according to previous studies. At 10 hPa, the maximum in CH_4 is centred year-around at the Equator, while at 2 hPa, there are local maxima located in the subtropics of the respective summer hemisphere [*e.g.*, *Jones and Pyle*, 1984; *Ruth et al.*, 1997]. The feature at 2 hPa had been attributed to the equatorial semiannual oscillation [*Choi and Holton*, 1991]; the maxima found in the CH_4 distributions in the tropics coincide with the maxima in upwelling. The CH_4 at 2 hPa at the equator thus should show a semi-annual cycle. Furthermore, the 2 hPa and 10 hPa levels are distinct in the CH_4 variability seen in the polar region. At 10 hPa, the minima in polar regions during autumn and winter coincide with the maxima in downwelling within the Brewer-Dobson circulation [*Randel et al.*, 1998]. Note, CH_4 exhibits a more pronounced minimum in the Southern Hemisphere, since the polar vortex here is stronger and allows less CH_4-rich air to be mixed in from mid-latitudes than in the Northern Hemisphere. At 2 hPa, however, the minima show up in summer/autumn. These minima are the result of photochemistry, with CH_4 lifetimes decreasing to 4 months at these altitudes [*Randel et al.*, 1998; *Solomon*, 1986].

HALOE captures the tropical features well at both 2 and 10 hPa, and also includes both the downwelling at higher latitudes at 10 hPa and the enhanced chemistry during summer months at 2 hPa. MIPAS shows very similar features, but extends further into the polar regions, revealing the full extent and timing of these features. The maxima in both MIPAS(1) and MIPAS(2) are stronger than in HALOE. The ACE-FTS exhibits a noisier field attributable to its more limited sampling. This creates sharp maxima and edges especially in the tropics, where the instrument scans through the lower latitudes only once a season. The use of equivalent latitude would help to reduce the noise introduced by the limited sampling. However, climatologies in equivalent latitudes are not as practical for model-measurement comparison, so knowledge of the quality of ACE-FTS climatologies in geographical latitude as provided here is also valuable. **Figure 4.3.4b** shows the differences in the latitude-time evolution of the different instruments with respect to the MIM. Consistent with the annual zonal mean evaluation at 10 hPa, MIPAS(2) and HALOE agree mostly within 5% (both lying on the low side of the MIM). MIPAS(1) on the other hand shows deviations from these two instruments of up to 15%. At 2 hPa, the difference field

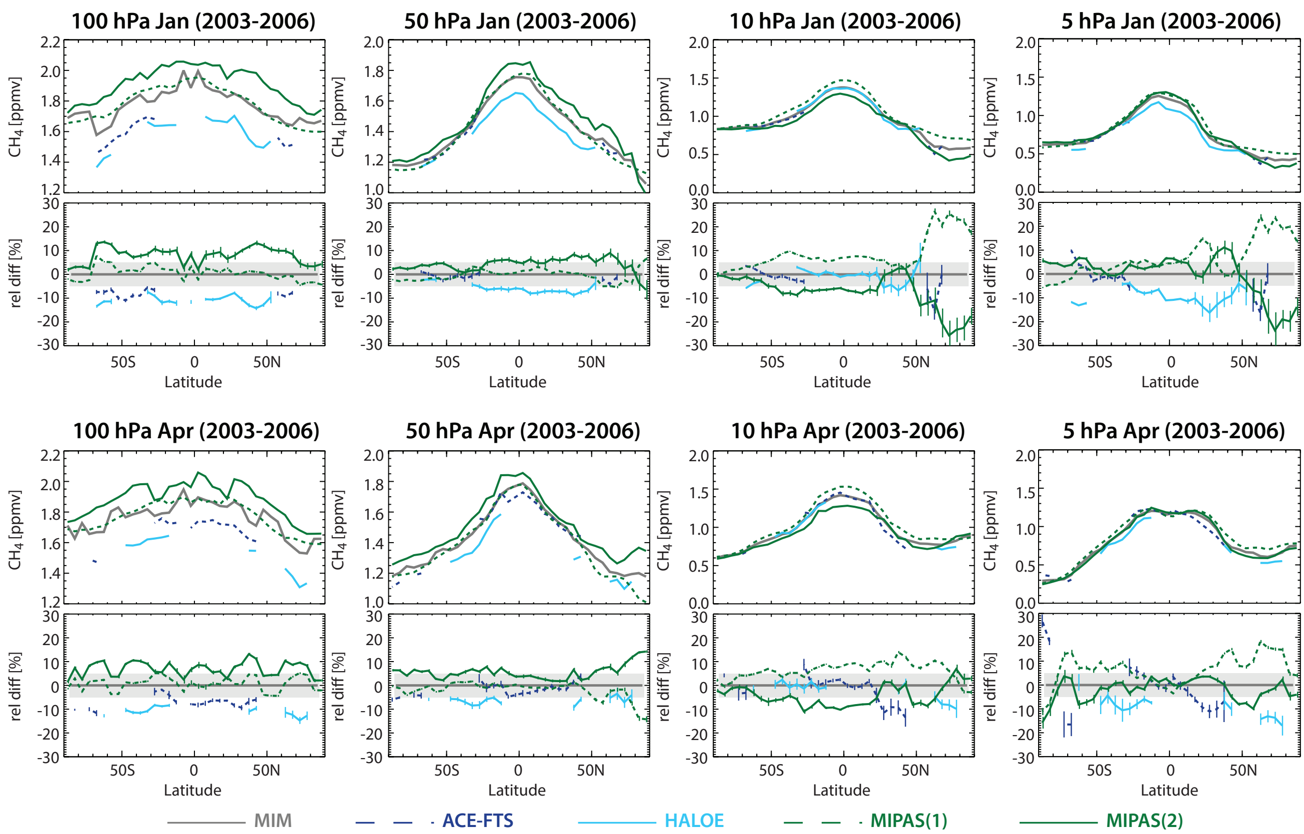


Figure 4.3.3: ***Meridional profiles of monthly zonal mean CH_4 for 2003-2006.*** *Meridional zonal mean CH_4 profiles for HALOE, MIPAS(1), ACE-FTS, and MIPAS (2) are shown at 100, 50, 10, and 5 hPa for January (upper row) and April (lower row), respectively. Differences between the individual instruments and the MIM are shown in the lower panels of each row.*

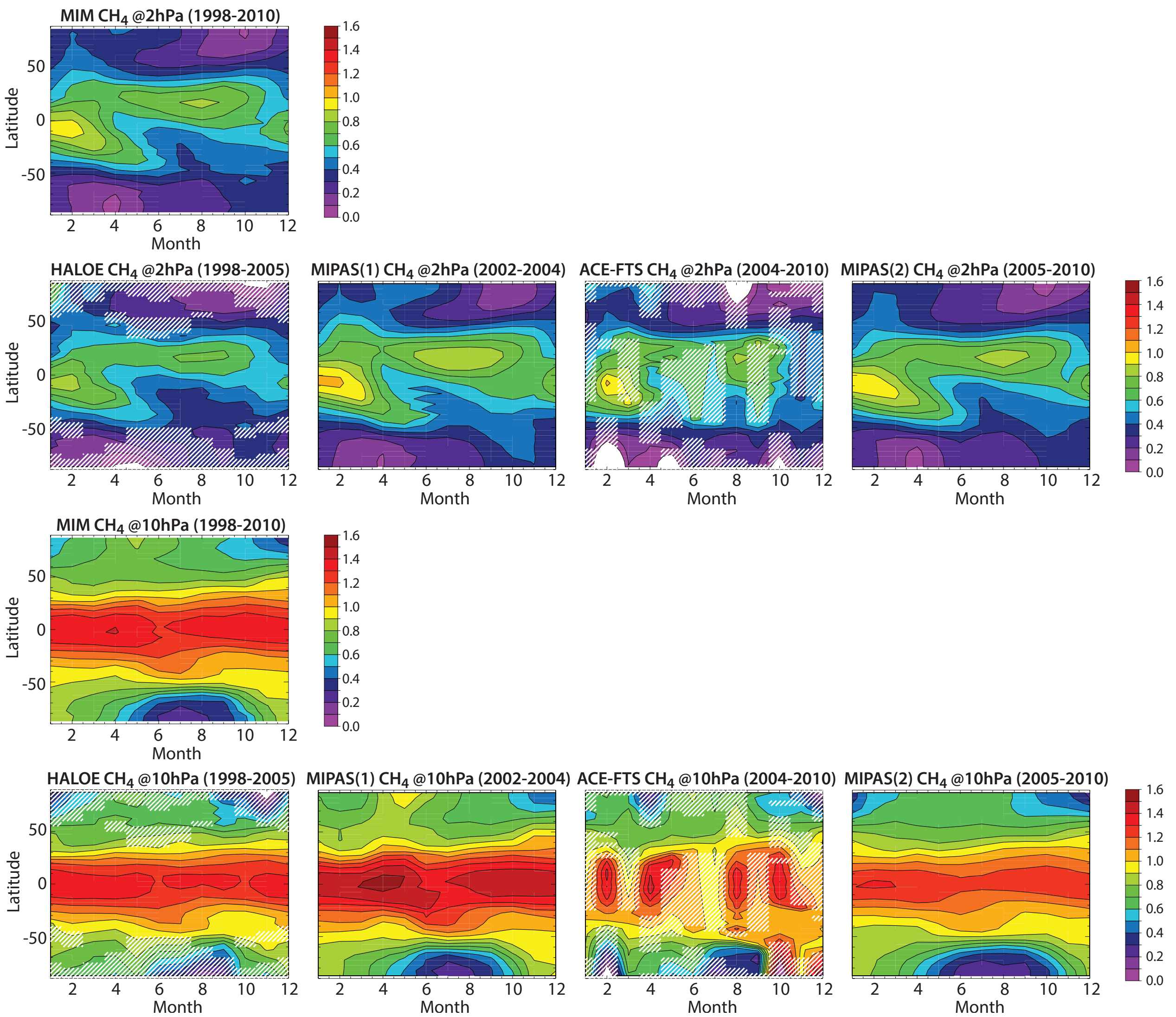


Figure 4.3.4a: Latitude–time evolution of CH_4. *The latitude-time evolution of montly zonal mean CH_4 at 2 hPa (top two rows), and 10 hPa (bottom two rows) are shown for the MIM (1998-2010), HALOE (1998-2005), MIPAS(1) (2002-2004), ACE-FTS (2004-2010), and MIPAS(2) (2005-2010) averaged over the time period given in brackets. HALOE and the ACE-FTS show interpolated fields, with hatched regions indicating where no measurements are available.*

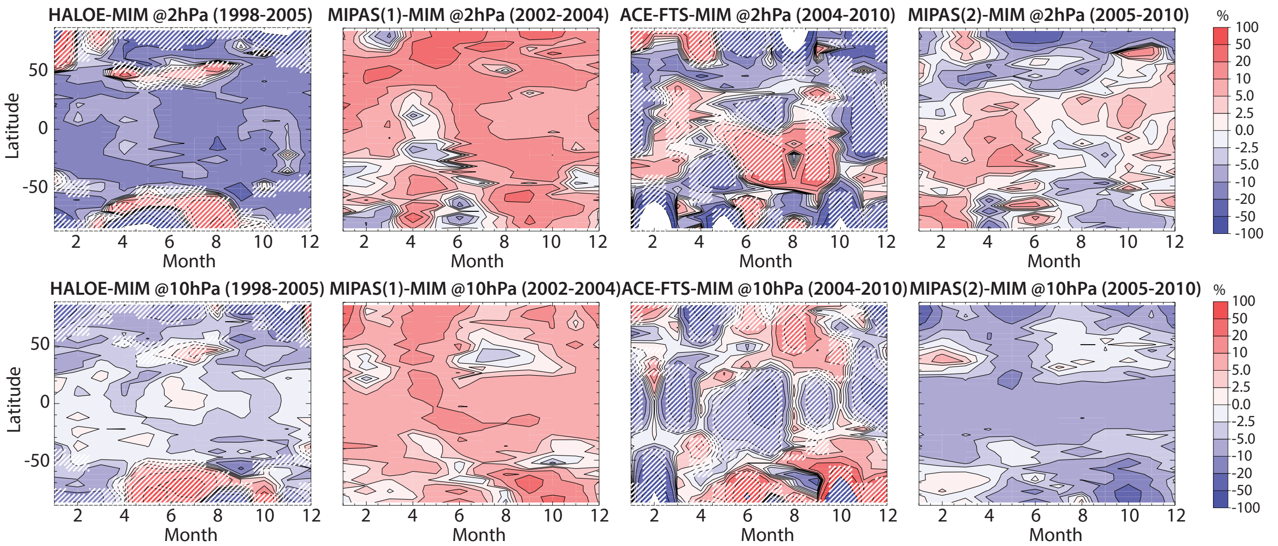


Figure 4.3.4b: Latitude–time evolution of differences in CH_4. *CH_4 differences with respect to the MIM at 2 hPa (top), and 10 hPa (bottom) are shown for HALOE, MIPAS(1), ACE-FTS, and MIPAS(2) over the time period as indicated in Figure 4.3.4a. For HALOE and ACE-FTS, hatched regions indicate where no measurements are available.*

is quite noisy for MIPAS(2) (and also ACE-FTS), but shows differences between MIPAS(1) and HALOE of up to 40%. Note that as a first approximation we assume the CH_4 trend between 1998 and 2010 to be negligible. A comparison of this evaluation limited to the year 2005 did increase and not decrease the differences.

4.3.4 CH_4 evaluations: Interannual variability

Figure 4.3.5 shows deseasonalised anomalies at different pressure levels in the tropics and the Northern Hemisphere mid- and high latitudes. In the tropics at 2 hPa, the interannual variability shows an approximately 2-year long fluctuation linked to the QBO [*Randel et al.*, 1998], with anomalies from the mean of around ±18%. The tropical QBO signal in methane is prominent between about 10 and 1 hPa (35-45 km), and fades away at altitudes below 10 hPa due to too small vertical gradients (not shown). At the tropopause height (around 100 hPa), methane interannual variability is very small and dominated by the long-term tropospheric trend. Although the QBO is a tropical phenomenon, it affects also the extra-tropics, as seen for 10 hPa in Northern mid-latitudes. However, here the QBO signal is somewhat weaker showing an anomaly of ±10% from the mean only. The peak negative anomaly is seen about nine months later than the peak negative anomaly at 2 hPa in the tropics, which reflects the different transport time scales in different regions of the atmosphere. At 50 hPa in the Northern polar region, the QBO signal has basically vanished and the interannual variability is instead driven by the varying strength of the polar vortex during winter months.

The comparisons reveal a very good agreement between the different instruments in terms of the magnitude of and structure in interannual variability. Even ACE-FTS with its limited sampling follows the fluctuations approximately. Note that the same evaluation, however treating MIPAS(1) and MIPAS(2) as continuous time series, reveals some inconsistency between the two datasets, which can be explained by the high bias of MIPAS(1) at 10 hPa and MIPAS(2) at 50 hPa, respectively (see **Figure 4.3.1** and **Figure A4.3.1b** in *Appendix A4*). The comparison also confirms a known high bias of the high spectral resolution CH_4 MIPAS(1) data in the MS [*c.f., Glatthor et al.*, 2005], which has been largely removed in the low spectral resolution data [*von Clarmann et al.*, 2009a].

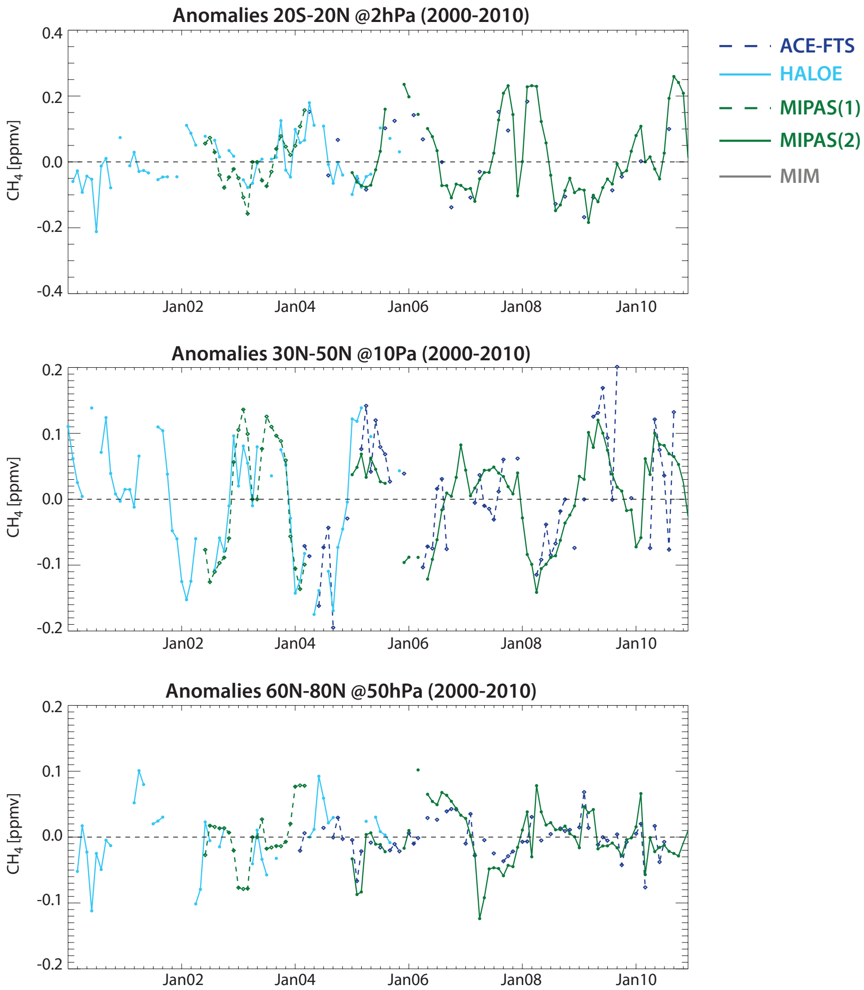


Figure 4.3.5: Time series of deseasonalised CH_4 anomalies between 2000 and 2010. *Deseasonalised CH_4 anomalies are shown for 2 hPa in the tropics (20°S-20°N; upper panel), 10 hPa at Northern mid-latitudes (30°N-50°N; middle panel), and 50 hPa at Northern high latitudes (60°N-80°N; lower panel).*

4.3.5 Summary and conclusions: CH_4

A comparison of three CH_4 climatologies (HALOE, MIPAS, and ACE-FTS) has been carried out. MIPAS data before/after 2005 have been evaluated separately (using MIPAS(1) and MIPAS(2)). Overall findings on the systematic uncertainty in our knowledge of the CH_4 mean state and important characteristics of the individual datasets are presented in the following summary including two synopsis plots. The first summary plot (**Figure 4.3.6**) provides information on the mean state and its uncertainty derived from the spread between the datasets. The second summary plot (**Figure 4.3.7**) shows specific inter-instrument differences in form of the deviations of the instrument climatologies from the MIM climatology. For each instrument and selected region, the deviation to the MIM is given in form of the median (mean) difference over all grid points in this region. Additionally for each instrument the spread of the differences over all grid points in this region is presented. Note that both pieces of information (average deviation and spread) are important for a meaningful assessment of inter-instrument differences. A detailed description of the summary plots can be found in *Section 3.3.5*.

Atmospheric mean state

The uncertainty in our knowledge of the annual mean state of atmospheric CH_4 as derived from the three satellite instruments is smallest in the LS and tropical/NH subtropical MS with a 1σ multi-instrument spread of less than ±6% (see **Figure 4.3.6**). The uncertainty is larger in the UT and

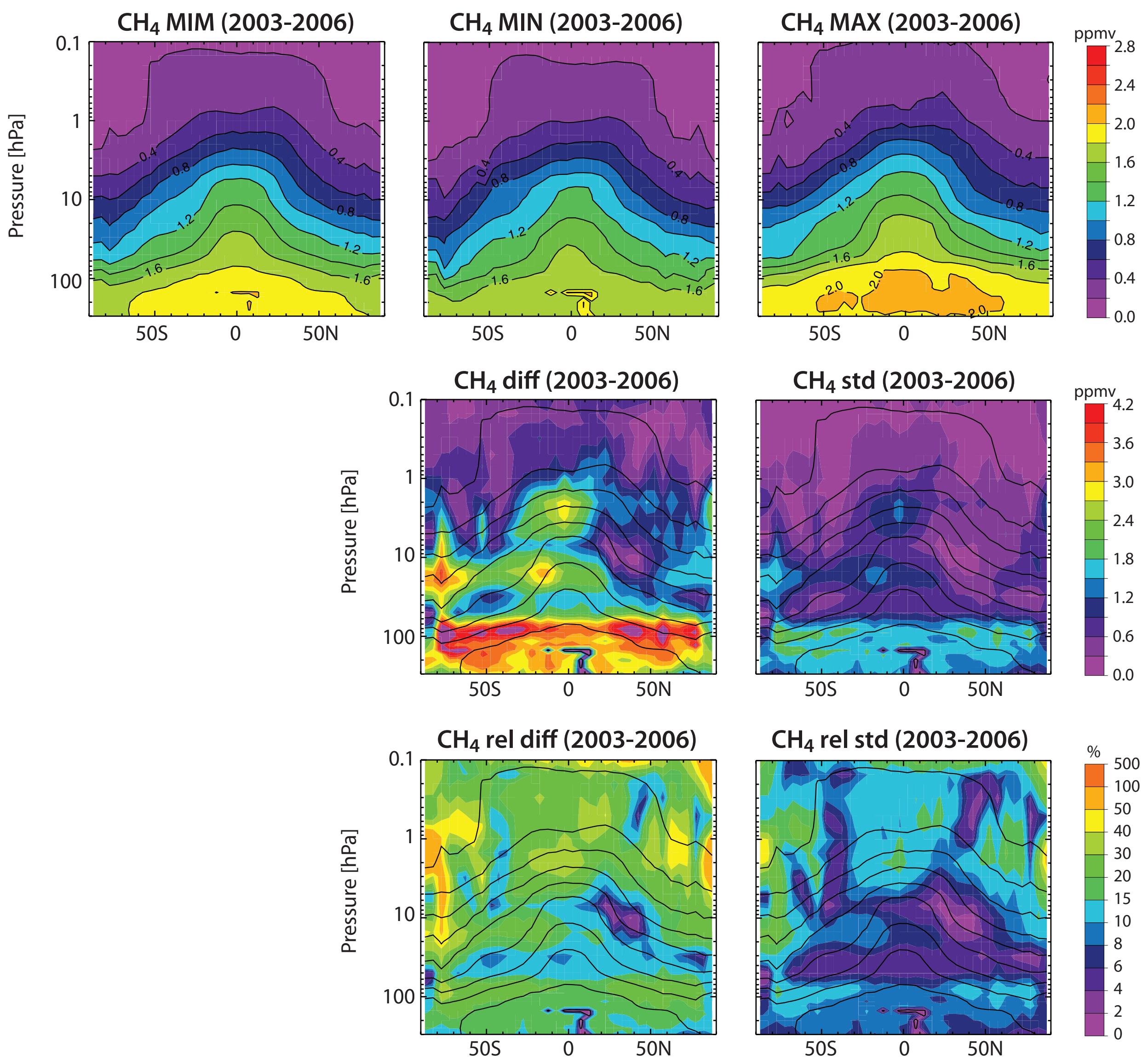


__Figure 4.3.6: Summary of CH_4 annual zonal mean state for 2003-2006.__ Annual zonal mean cross sections for 2003-2006 of the MIM, minimum (MIN), and maximum (MAX) CH_4 values are shown in the upper row. The maximum differences over all instruments (MAX-MIN) and the standard deviation over all instruments are shown in the middle row. The relative differences and relative standard deviations with respect to the MIM are shown in the lower row. Black contours in lower panels repeat the MIM distribution. Instruments considered are HALOE, MIPAS(1), ACE-FTS, and MIPAS(2).

lowermost stratosphere with a 1σ multi-instrument spread of around 10%. The uncertainty increases also towards higher altitudes and latitudes, where 1σ values reach up to ±20% and more. The higher uncertainty in the USLM is explained by CH_4 concentrations close to the detection limit of the instruments.

Performance by region

In the USLM (0.1-5 hPa), all instruments agree within ±15% but show large MAD values of the same magnitude, indicating that the deviations from the MIM are not well defined within the region. The MAD values are somewhat larger in the extra-tropics than in the tropics, most likely due to the larger natural variability in this region. HALOE is consistently lower than the MIM.

In the MS (5-30 hPa), the MADs are much smaller than in the USLM in the tropics, but less so in the extra-tropics. HALOE and ACE-FTS are very close to the MIM in both the tropics and extra-tropics, while MIPAS(1) and MIPAS(2) show the most positive and negative deviations from the MIM, respectively.

In the UTLS (30-300 hPa), ACE-FTS and HALOE are on the low side and MIPAS(1) and MIPAS(2) both on the high side of the MIM. All of the instruments exhibit relatively small MADs indicating that the mean differences from the MIM are well defined. Given that MIPAS has a known high bias in this lower part of the atmosphere [*von Clarmann et al.*, 2009a], ACE-FTS and HALOE reflect more accurately the range of uncertainty in the absolute values of this region.

Instrument-specific conclusions

HALOE provides the longest time series, but exhibits consistently lower values than the other instruments through most of the atmosphere. Previous validation with correlative measurements has indicated agreement of typically better than 15% [*Park et al.*, 1996]; our study shows better agreement through most of the LS and MS, at least with respect to the ACE-FTS.

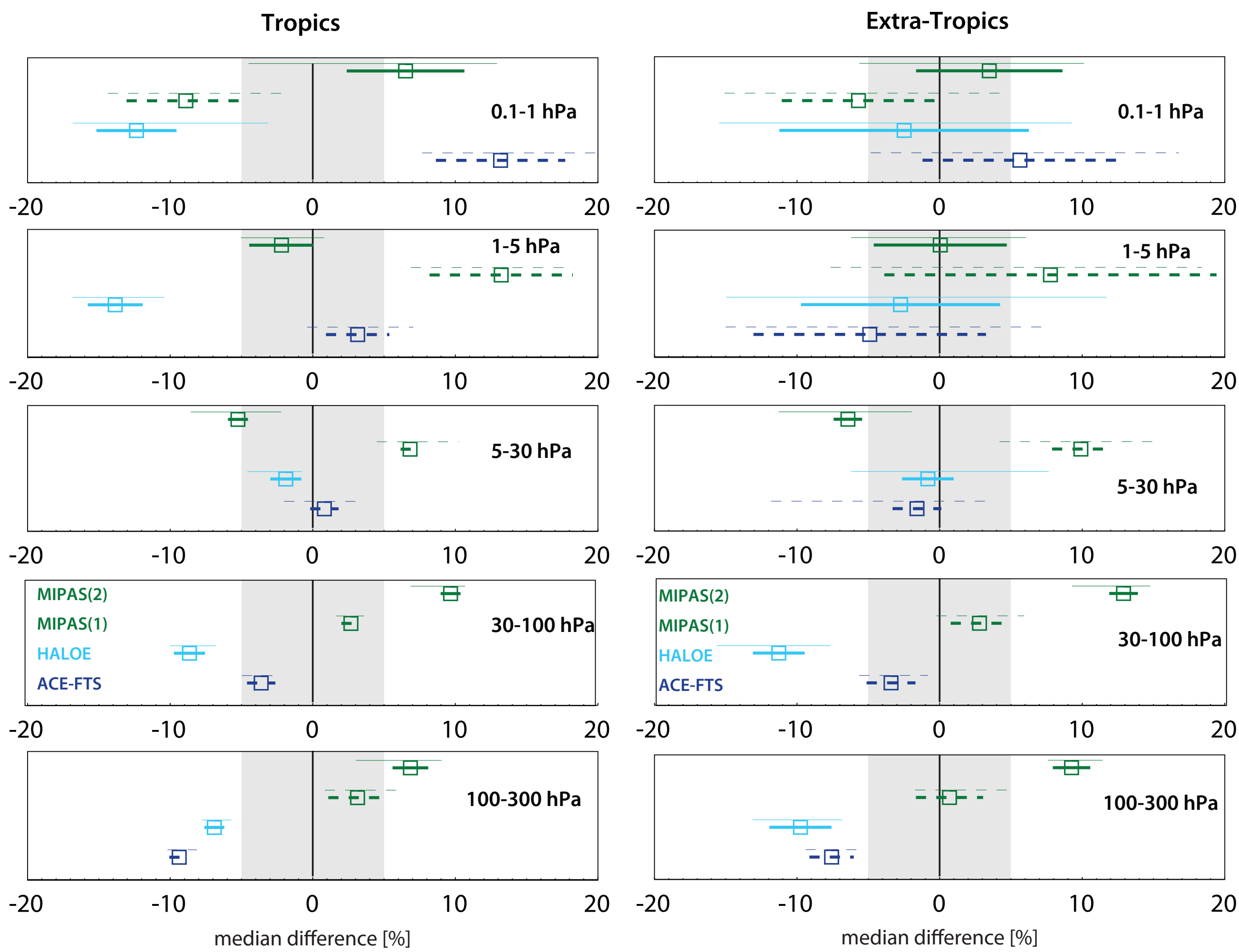


Figure 4.3.7: Summary plot of CH_4 inter-instrument differences for 2003-2006. *Over a given latitude and altitude region the median (squares), median absolute deviation (MAD, thick lines), and the standard deviation (thin lines) of the monthly mean relative differences between an individual instrument-climatology and the MIM are shown. Results are shown for the tropics (20°S-20°N) and extra-tropics (40°S-80°S and 40°N-80°N) and for 4 different altitude regions from the UT to the US between 300 and 1 hPa for the reference period 2003-2006.*

Despite its limited sampling, **ACE-FTS** shows mostly coherent interannual variability and exhibits deviations from the MIM that are mostly within ±5%, except in the UT and LM. These results are broadly consistent with the validation study of *De Mazière et al.* [2008] where the ACE-FTS results were found to reproduce the variability of the atmosphere well. However, our evaluations indicate somewhat smaller inter-instrument differences in the LS and MS than found in *De Mazière et al.* [2008], which may be the result of using a climatological evaluation approach that helps to limit the impact of natural variability on instrument comparisons.

MIPAS(1) and **(2)** both have a known bias in upper tropospheric CH_4 [*von Clarmann et al.*, 2009a], which are above the global mean values derived from tropospheric *in-situ* measurements, and relatively large vertical fluctuations in the deviations from the MIM. Limb emission measurements are less sensitive to CH_4 mixing ratios in the TTL than those above these levels, which can lead to increased retrieval errors or may be reflected in oscillating profiles. MIPAS(1) and MIPAS(2) are generally to be treated as independent datasets. Thus the discontinuity in extra-tropical MIPAS CH_4 is not unexpected and serves as another example that trend analysis of MIPAS data requires a special data merging approach [*von Clarmann et al.*, 2010].

4.3.6 Recommendations: CH_4

For trend studies it will be important to include CLAES, SAMS, and iSAMS observations as well, since these instruments would yield data from further in the past before trends in tropospheric CH_4 flattened.

The CH_4 latitude-time evolution at 2 and 10 hPa may be a useful diagnostic for testing the location and seasonal behaviour of the Brewer-Dobson circulation *versus* chemistry effects in chemistry-climate models. At 2 hPa, CH_4 clearly reveals the upwelling branch in the Brewer-Dobson circulation, which shifts off the equator into the summer hemisphere subtropics [*Randel et al.*, 1998]. The lowest values in CH_4 at this level are found in the polar regions during the summer/autumn months due to photochemical methane destruction. At 10 hPa, strong gradients and very low CH_4 reveal the strong downwelling of older stratospheric air within the polar vortices starting in late autumn and persisting through to early spring.

4.4 Nitrous oxide – N_2O

Nitrous Oxide (N_2O), despite its relatively low atmospheric concentrations, is another important greenhouse gas (approximately 300 times more powerful than CO_2 on a per molecule basis). This is due to its long atmospheric lifetime (about 120 years) and large infrared absorption capacity (per molecule). N_2O is inert in the troposphere, but is destroyed in the stratosphere through photolysis (about 90% of total loss) and reaction with $O(^1D)$ (about 10% of total loss) [*Seinfeld and Pandis*, 1998]. The latter loss reaction leads to the production of NO (see *Section 4.10*), which is involved in the chemical destruction of O_3 in the stratosphere. N_2O is predicted to constitute the single-most important contribution to future emissions of ozone-depleting substances in the 21st century [*Ravishankara et al.*, 2009], although its ozone-depletion potential (and hence effect on the ozone layer) will be strongly dependent on its lifetime, which is set to change under climate change due to changes in the stratospheric circulation [*Plummer et al.*, 2010].

4.4.1 Availability of N_2O measurements

Satellite measurements of N_2O available to the SPARC Data Initiative include those from ACE-FTS, Aura-MLS, MIPAS, and SMR, with the first time series (by SMR) starting in 2001. Earlier N_2O measurements, which are not included in the SPARC Data Initiative, can be obtained from SAMS [*Drummond et al.*, 1980], ISAMS [*Taylor et al.*, 1993], ATMOS [*Gunson et al.*, 1996], CLAES [*Roche et al.*, 1993], CRISTA [*Riese et al.*, 1999], ILAS [*Kanzawa et al.*, 2003], and ILAS-II [*Ejiri et al.*, 2006]. The instruments participating in the SPARC Data Initiative cover the full altitude range considered in this report, except Aura-MLS, which provides measurements for a slightly smaller range between 100 and 0.46 hPa.

Tables 4.4.1 and **4.4.2** compile information on the availability of N_2O measurements, including data version, time period, height range, vertical resolution, and references relevant for the data product used in this report.

Table 4.4.1: Available N_2O measurement records from limb-sounding satellite instruments between 1978 and 2010. *The red filling in each grid box indicates the temporal and vertical coverage (within the pressure range 300-0.1 hPa) of the respective instrument.*

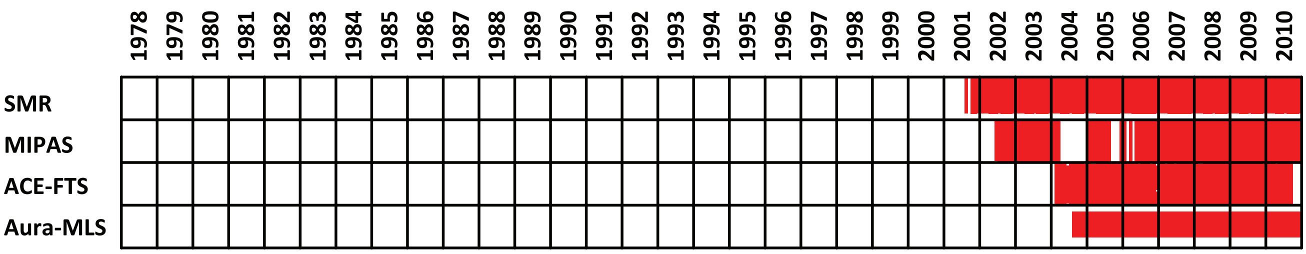


Table 4.4.2: Time period, vertical range, vertical resolution, references and other comments for N_2O measurements.

Instrument	Time period	Vertical range	Vertical resolution	References	Additional comments
Aura-MLS V3.3	Aug 04 –	100 – 0.46 hPa	4 – 6 km for p > 1hPa	*Lambert et al.*, 2007 *Livesey et al.*, 2011	
ACE-FTS V2.2	Mar 04 –	5 km – 60 km	3 – 4 km	*Strong et al.*, 2008	
SMR V2.1	Jul 01 –	12 – 60 km	~1.5 – 3 km (LS)	*Urban et al.*, 2005a,b *Urban et al.*, 2006	
MIPAS MIPAS(1) V11 MIPAS(2) V220	 Mar 02 – Mar 04 Jan 05 – Apr 12	 Cloud top – 70 km	 4 – 5 km 2.5 – 5.8 km	*Glatthor et al.*, 2005 *Funke et al.*, 2008 *von Clarmann et al.*, 2009a	measurement mode switched in 2005 from high spectral to high vertical resolution

4.4.2 N_2O evaluations: Zonal mean cross sections, vertical and meridional profiles

Annual zonal mean cross sections for the time period 2006-2009 are analysed to investigate mean biases between the various datasets. Note, we do not use the years 2005 and 2010 to minimise the effect of data gaps in MIPAS and ACE-FTS. Additionally, vertical and meridional profiles are evaluated in order to focus on specific months, altitude and latitude regions.

Aura-MLS, MIPAS, ACE-FTS, and SMR (2006-2009)

Figure 4.4.1a shows annual zonal mean cross sections averaged over the years 2006-2009 for the multi-instrument mean (MIM) and the four different instruments. Note that we consider the high- and low-spectral resolution versions of MIPAS (MIPAS(1) and MIPAS(2) respectively) separately in order to investigate potential changes in the performance of the instrument. Due to its long lifetime, N_2O is generally well-mixed in the troposphere but decreases exponentially with height in the stratosphere due to photolysis and reaction with $O(^1D)$. The isopleths are shaped similarly to those of CH_4, sloping downwards towards higher latitudes, reflecting tropical upwelling and extra-tropical downwelling of air masses within the Brewer-Dobson circulation. However, N_2O vertical gradients in the UTLS are smaller than those of CH_4 due to the longer lifetime of N_2O.

The different instruments show a very similar annual zonal mean structure, including a characteristic two-peak feature in the US [*e.g.*, *Jones and Pyle*, 1984], which stems from the upwelling within the Brewer-Dobson circulation that is located off the equator in the respective summer hemisphere. The appearance of these 'rabbit ears' [*Randel et al.*, 1998] is modulated by the QBO and the feature is much more pronounced when looking at monthly mean fields (see **Figure A4.4.1a** in *Appendix A4*). ACE-FTS exhibits a somewhat 'noisier' zonal mean field than the other instruments. Note that the 'noise' in the ACE-FTS climatology is not due to limitations in the

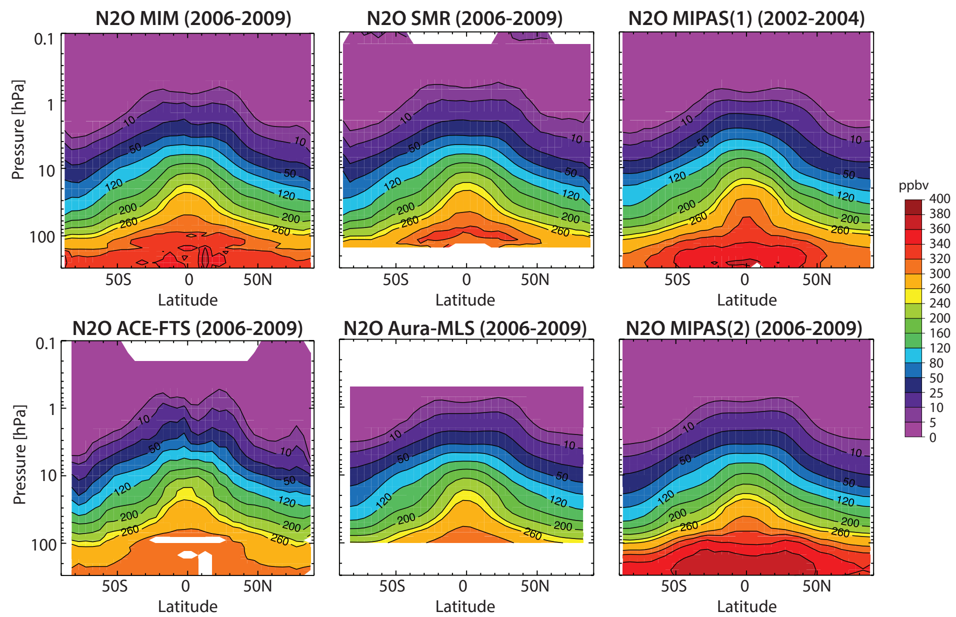


Figure 4.4.1a: Cross sections of annual zonal mean N_2O for 2006-2009. *Annual zonal mean N_2O cross sections are shown for the MIM in the leftmost upper panel along with SMR, MIPAS(1), ACE-FTS, Aura-MLS, and MIPAS(2). Note, MIPAS(1) is excluded from the MIM so not to bias the MIM towards this instrument.*

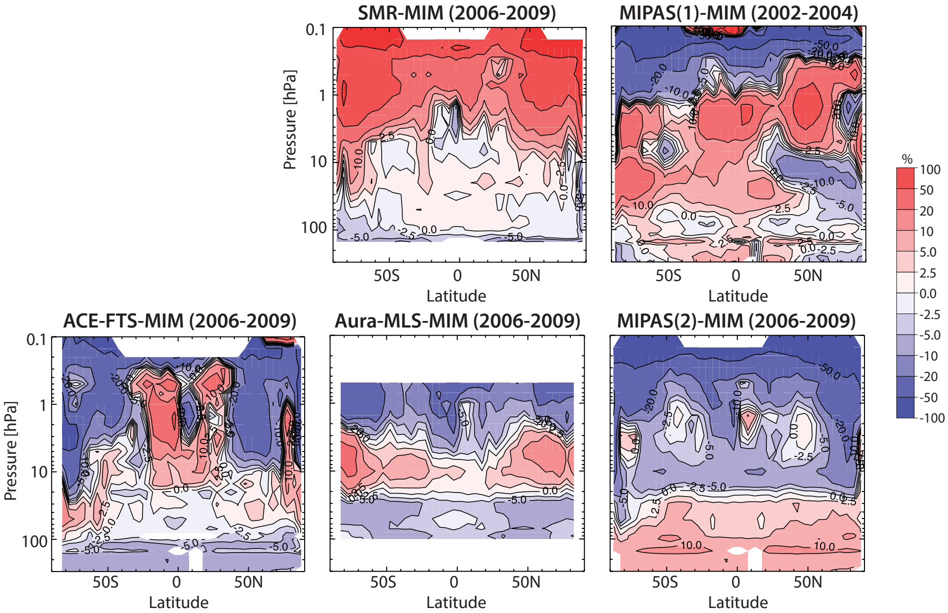


Figure 4.4.1b: Cross sections of annual zonal mean N_2O differences for 2006-2009. *Shown are the relative differences between the individual instruments' (SMR, MIPAS(1), ACE-FTS, Aura-MLS, and MIPAS(2)) annual zonal mean N_2O distributions and the MIM.*

retrieval. The single-scan precision of ACE-FTS is much better than (or at least comparable to) that of other instruments. The 'noise' in the ACE-FTS climatology is rather due to the instrument's limited sampling. This results in a smaller number of profile measurements that can be used to average out geophysical variability in the atmosphere.

Figure 4.4.1b shows the relative differences of the different instruments with respect to the MIM. For all instruments, the differences from the MIM are very small throughout the UTLS and MS, with maximum values of ±5% (~5-15 ppbv). In the US and LM, the absolute differences are small (~1-5 ppbv), but relative differences grow to very large values of up to ±100%. Note that these large relative differences are mostly due to the exponentially decreasing N_2O values that approach the detection limits of the instruments. SMR shows a systematic positive bias in the USLM compared to all the other instruments. The ACE-FTS shows strong positive deviations from the MIM in the tropical MS and US that are not seen in the monthly mean evaluations shown in **Figure A4.4.1b** in *Appendix A4*, and therefore are likely to be a sampling artefact. The structures seen in the ACE-FTS differences can be explained by sampling the effect of the seasonal change in the Brewer-Dobson circulation strength on the N_2O distribution only during phases of strong upwelling, *i.e.*, February-April and August-October.

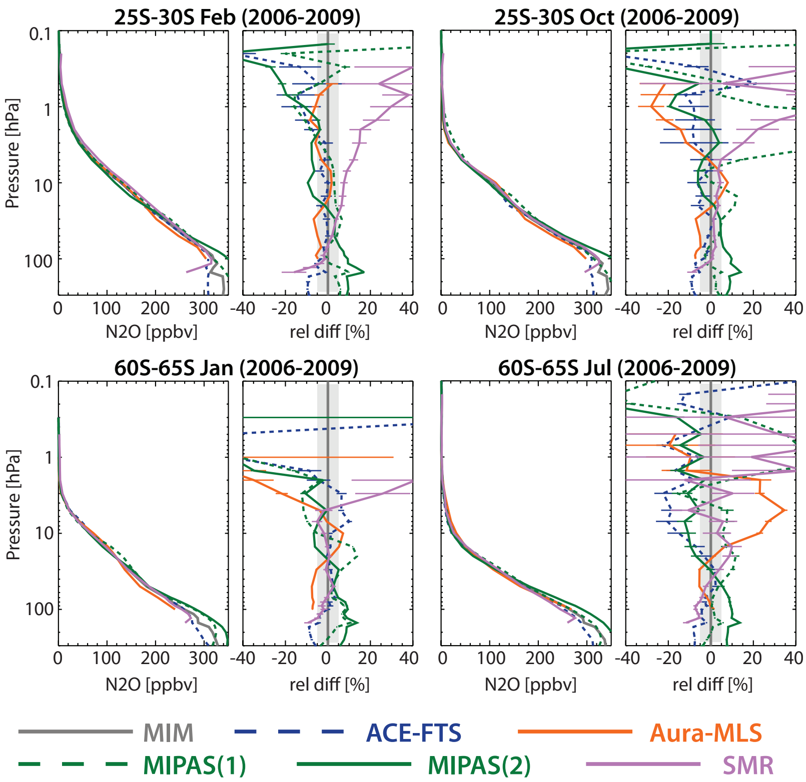


Figure 4.4.2: Vertical profiles of monthly zonal mean N_2O for the Southern Hemisphere. *Vertical N_2O profiles for 25°S-30°S February and October (upper panels), and for 60°S-65°S January and July (lower panels) are shown together with the instrument differences from the MIM for SMR, MIPAS(1), ACE-FTS, Aura-MLS, and MIPAS(2) and for the period 2006-2009.*

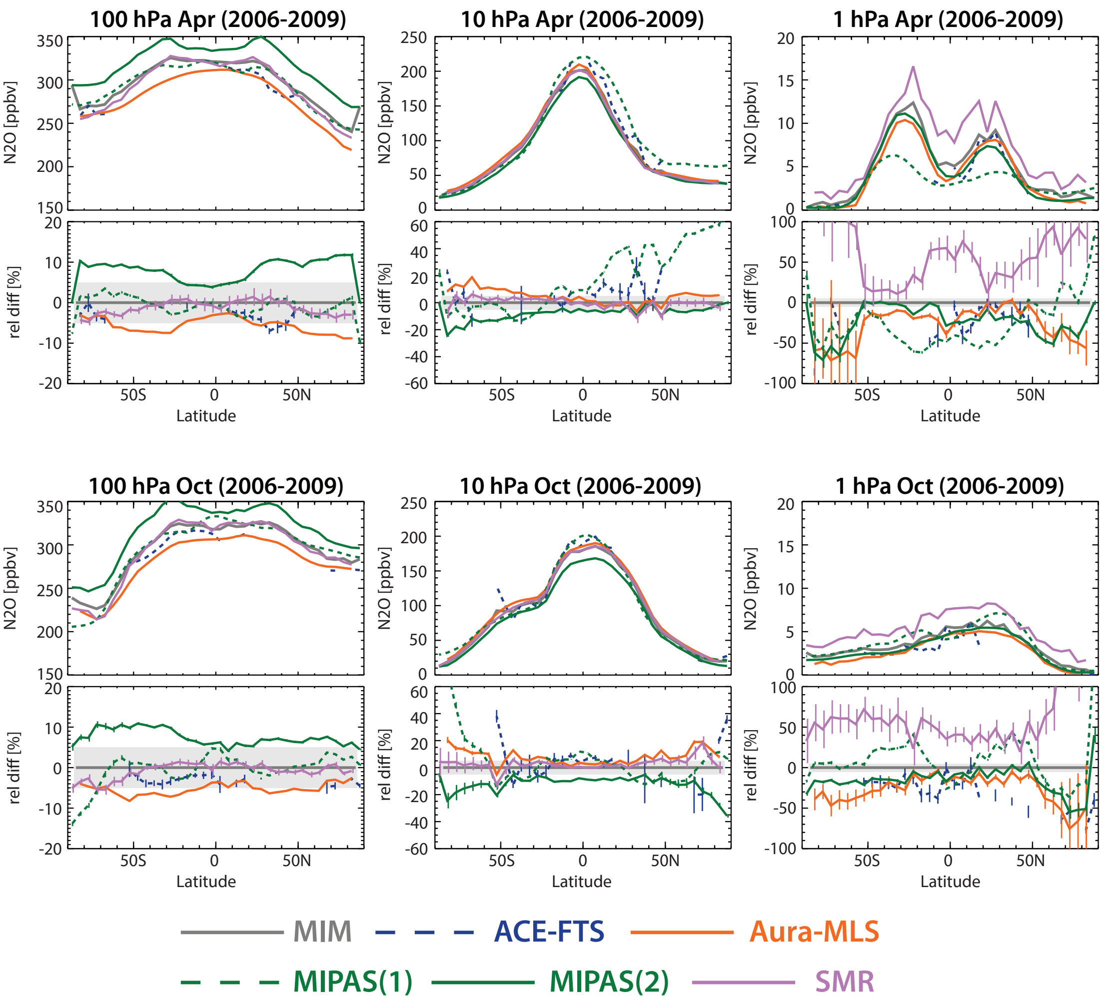


Figure 4.4.3: Meridional profiles of monthly zonal mean N_2O for 2006-2009. *Meridional N_2O profiles are shown at 100, 10, and 1 hPa for April (upper row) and October (lower row). Differences between the individual instruments (SMR, MIPAS(1), ACE-FTS, Aura-MLS, and MIPAS(2)) and the MIM profiles are shown in the lower panels.*

Vertical profiles and their relative differences are shown in **Figure 4.4.2** for the Southern Hemisphere. Note that the results are similar for the Northern Hemisphere (which can be found in **Figure A4.4.2** in *Appendix A4*). The monthly zonal averages reveal similar relative differences to those derived from the annual averages (compare also **Figures A4.4.1a** and **A4.4.1b** in *Appendix A4*). The monthly relative differences in the UTLS and MS found in the vertical profiles reach values of up to ±10-15% and increase above 10 hPa. MIPAS(2) shows much higher N_2O values below about 50 hPa than the other instruments. Above about 5-10 hPa, MIPAS(2) is closer to Aura-MLS and ACE-FTS, while SMR exhibits largest (positive) departures from the MIM. Aura-MLS shows positive deviations from the MIM around 10 hPa.

Figure 4.4.3 shows the monthly meridional zonal mean N_2O profiles and their relative differences. At 100 hPa, differences are within ±10% over all latitudes. We find that MIPAS(2) is systematically larger and Aura-MLS is systematically lower at this level. Note the very good agreement between ACE-FTS and the other instruments despite its infrequent sampling. Good agreement between all the instruments is also seen at 10 hPa with relative differences mostly within ±5-10%. MIPAS(2) is here generally lower than the other instruments by 10-15%. MIPAS(1) shows larger deviations at higher latitudes of the respective spring hemisphere, which is most likely due to sampling different years (2002-2004). ACE-FTS is somewhat noisier with relative differences of around ±20%. As noted above, this is due to its limited spatio-temporal sampling, and not due to a lack of precision in the profile measurements. At 1 hPa in April, the meridional profile of N_2O shows two local maxima in the subtropics dubbed 'rabbit ears' [*Randel et al.* [1998], see above). At this level, SMR is fairly noisy compared to the other instruments and exhibits a positive bias over all latitudes between 2 ppbv (October) and 4-5 ppbv (April), corresponding to up to 100% of the small mean N_2O mixing ratios measured at these altitudes.

4.4.3 N_2O evaluations: Seasonal cycles

Seasonal cycles in N_2O are often used as process-oriented diagnostics in model-measurement comparison efforts (*e.g.*, Chapter 5 of *SPARC* [2010]). In order to quantify the observational range or uncertainty, the seasonal cycles at 100 and 50 hPa in both the tropics and extra-tropics are compared in **Figure 4.4.4**. The mean values of the seasonal

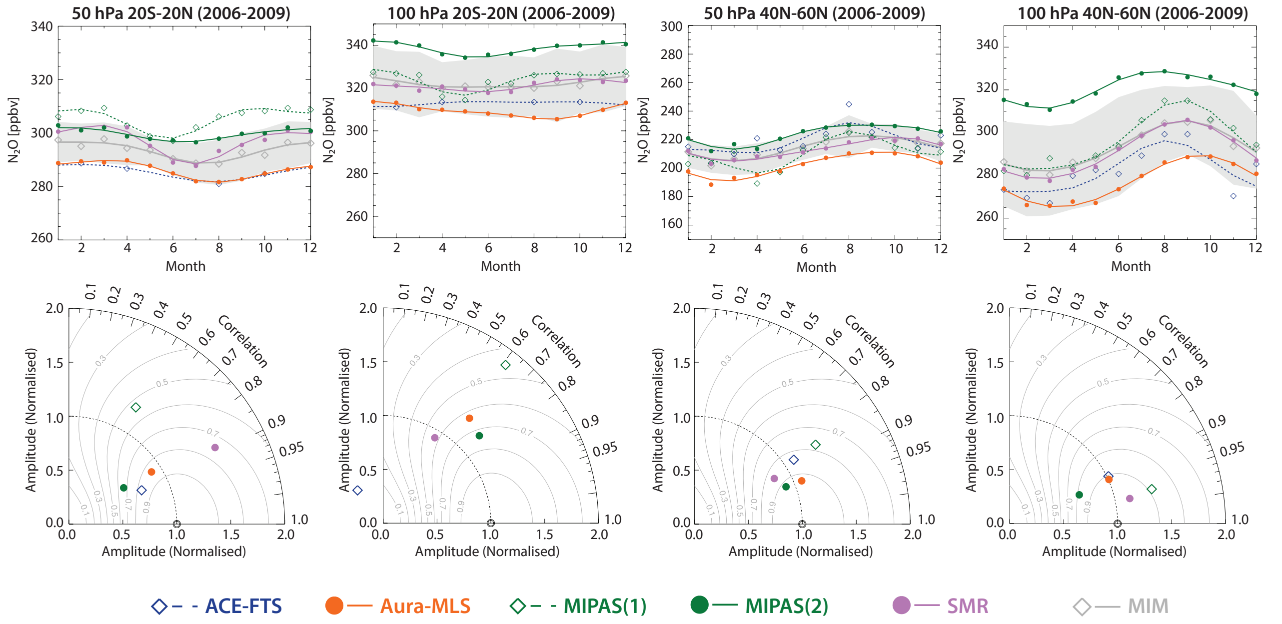


Figure 4.4.4: Seasonal cycle of N_2O in the tropics and at NH mid-latitudes at 100 and 50 hPa. *Seasonal cycles (upper panels) and corresponding Taylor diagrams (bottom panels) for monthly zonal mean N_2O are shown for 20°S-20°N (left two panels) and 40°N-60°N (right two panels) at 50 and 100 hPa, averaged over 2006-2009. The grey shading indicates ±1σ about the MIM.*

cycles are well defined to about ±5-10% at both pressure levels and in both the tropics and extra-tropics, consistent with the annual zonal mean evaluations. Taylor diagrams yield in addition information on the shape of the seasonal cycle. The amplitude in the seasonal cycle (seen in the Taylor diagram in the departures from 1 on the radial axis or the dashed line) is better defined in the extra-tropics than in the tropics. More generally, MIPAS(2) shows a somewhat lower amplitude than the other instruments at 50 hPa in the tropics and 100 hPa in the extra-tropics, while MIPAS(1) shows a somewhat too high amplitude in all regions. In the tropics, the instruments show rather large differences in the phase of the seasonal cycle (as seen in the Taylor diagram in lower correlation values on the azimuthal axis). Note that some of the differences in phase and amplitude of the seasonal cycles may be explained by differences in the vertical resolution of the measurements, in particular in regions with strong vertical gradients and large seasonal variability.

4.4.4 N_2O evaluations: Interannual variability

Finally, the interannual variability of monthly zonal mean N_2O is analysed using deseasonalised anomalies as shown in **Figure 4.4.5** for the tropics. At 100 hPa, the different instruments show no clear seasonality in N_2O near the tropical tropopause, with inter-instrument differences lying within 5-10 ppbv (~5%). SMR shows somewhat larger fluctuations than MIPAS or Aura-MLS at this level. Also, a strong negative anomaly is seen in SMR in the first half of 2004, which cannot be seen in MIPAS(1) or ACE-FTS. A similar negative anomaly is seen in SMR at 100 hPa in the extra-tropics (see **Figure A4.4.3** in *Appendix A4*), but is again not confirmed by MIPAS(1). However, when the MIPAS(1) and MIPAS(2) are treated as one combined time series (see **Figure A4.4.4** in *Appendix A4*), the feature is revealed at least in the extratropics and hence may indeed be real (note that de-seasonalizing the very short MIPAS(1) time series has likely removed the signature). **Figure A4.4.4** in *Appendix A4* reveals a discontinuity between MIPAS(1) and MIPAS(2) N_2O indicating that MIPAS(1) and MIPAS(2) have to be treated as independent datasets as is the case for CH_4 (see *Section 4.3*).

At 10 hPa, the different instruments show excellent agreement of the interannual variability, which is of the order of ±10%. An exception is ACE-FTS, which does not have the temporal coverage needed to follow the anomalies accurately enough. A strong QBO signal is apparent. Note, the QBO affects N_2O more strongly than CH_4, since N_2O exhibits stronger vertical gradients around this pressure level.

At 1 hPa, the QBO signal has disappeared, but the instruments capture large anomalies very well especially during January/February as seen in the time series. The evaluation of interannual variability indicates that SMR despite its large positive bias in the USLM apparent in **Figures 4.4.1-4.4.3** is useful for the construction of climate data records in this region.

The QBO signal is also apparent in the NH (see **Figure A4.4.3** in *Appendix A4*) and SH (not shown) at 10 hPa, with the good agreement amongst the instruments. **Figure A4.4.3** in *Appendix A4* also reveals that in the NH extra-tropics at 100 hPa, the instruments agree better than in the tropics, though still not as well as at 10 hPa and 1 hPa, which is due to the smaller gradients found in N_2O in this region.

4.4.5 Summary and conclusions: N_2O

N_2O climatologies from four limb-sounders (SMR, MIPAS, ACE-FTS and Aura-MLS) have been compared within the SPARC Data Initiative. MIPAS data before and after 2005, when the instrument switched from a high- to a low-spectral resolution measurement mode, have been evaluated separately (MIPAS(1) and MIPAS(2)). Note that Aura-MLS provides N_2O data in a slightly more limited height range. Overall findings on the systematic uncertainty in our knowledge of the N_2O mean state and important characteristics of the individual datasets are presented in the following summary including two synopsis plots as discussed in detail in *Section 3.3.5*.

Atmospheric mean state

The relative uncertainty in our knowledge of the atmospheric N_2O annual mean state as derived from the four satellite instruments is smallest in the LS and MS of both the tropics and extra-tropics with 1σ multi-instrument spreads of less than 4% and 6%, respectively (see **Figure 4.4.6**). Reasonably good knowledge is also obtained in the UT and extra-tropical LS at altitudes below 100 hPa, where the uncertainty is smaller than 15%. The relative uncertainty increases towards the USLM (with values larger than 50%). Note, absolute uncertainties are smallest in the USLM. N_2O mixing ratios decrease quickly with altitude in this region and reach values close to or below the detection limits of the instruments.

Performance by region

As seen in **Figure 4.7.7**, in the LM (0.1-1 hPa), considerable disagreement in terms of relative uncertainty is found in the tropics (with values up to ±50%), and large disagreement (with values up to ±100%) in the extra-tropics. The largest disagreement is found in SMR, which is a clear outlier and has a positive bias of a few ppbv (up to ±100%) consistent with earlier studies [*e.g.*, *Strong et al.*, 2008] in this region. The other instruments MIPAS, Aura-MLS and ACE-FTS agree well within ±10%.

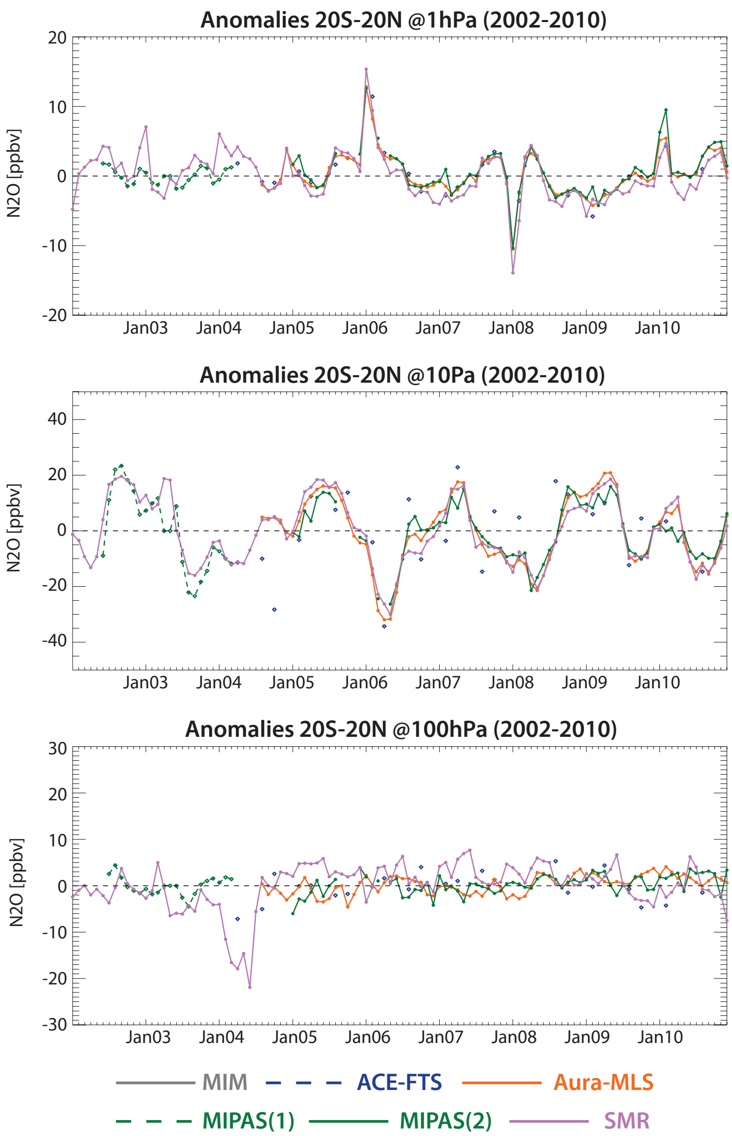


Figure 4.4.5: Time series of deseasonalised N_2O anomalies in the tropics. *Deseasonalised N_2O anomalies between 20°S-20°N are shown for the 1 hPa (upper panel), 10 hPa (middle panel), and 100 hPa (lower panel) levels.*

In the US (1-5 hPa), inter-instrument differences are somewhat larger for the instruments that agreed well in the LM, and with Aura-MLS and MIPAS(2) agreeing best with each other. SMR shows again largest differences from the MIM in the extra-tropics.

In the LS and MS (5-30 hPa and 30-100 hPa), the inter-instrument differences are mostly within ±5%, indicating very good agreement. However, somewhat larger MADs in the MS indicate that the deviations from the MIM are less well defined here than in the LS.

In the UT (100-300 hPa, which includes the extratropical lowermost stratosphere), the agreement between the instruments is good as well with inter-instrument differences being within ±10%. This good agreement can be explained by N_2O having smaller gradients across the UTLS region, which leads to smaller sampling-related biases in the monthly zonal means.

Instrument-specific conclusions

SMR shows an excellent performance in most diagnostics for N_2O with very small deviations from the MIM in the LS and MS. The deviations from the MIM increase towards higher altitudes and especially in the extra-tropics due to a positive bias of a few ppbv that becomes relevant where N_2O mixing ratios are low. Despite this bias, the instrument captures interannual variability well and hence may be used to construct climate data records after appropriate bias correction.

ACE-FTS measurements show very good agreement with the other instruments in the LS and MS, however its temporal and spatial coverage are not good enough to yield robust information on the seasonal cycles or interannual variability.

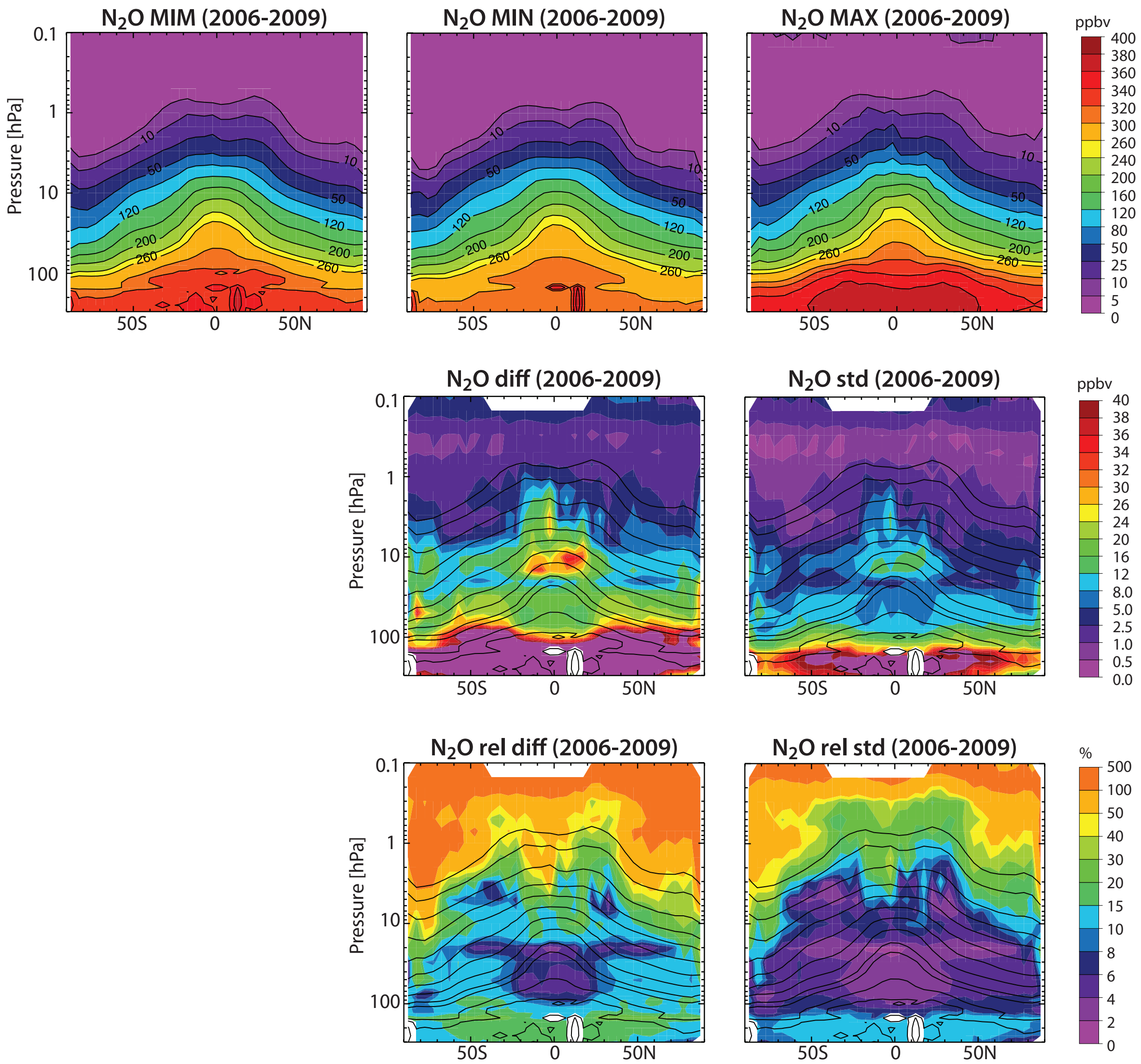


Figure 4.4.6: Summary of N_2O annual zonal mean state for 2006-2009. *Annual zonal mean cross sections for 2006-2009 of the MIM, minimum (MIN), and maximum (MAX) N_2O values are shown in the upper row. The maximum differences over all instruments (MAX-MIN) and the standard deviation over all instruments are shown in the middle row, the relative differences and relative standard deviations with respect to the MIM in the lower row. Black contours in lower panels repeat the MIM distribution. Instruments considered are SMR, MIPAS(1), ACE-FTS, MIPAS(2), and Aura-MLS. Note MIPAS(1) has been included despite the different time period it provided measurements for (2002-2004).*

MIPAS(2) shows largest positive deviations from the MIM in the UTLS and largest negative deviations in the MS, while **MIPAS(1)** exhibits much closer values to the MIM in most atmospheric regions. The differences between the two MIPAS datasets (or measurement periods) have to be taken into account when merging them into longer-term time series.

Aura-MLS data are limited to altitudes above the 100 hPa pressure level. The instrument performs very well in essentially all diagnostics; however its retrievals show a prominent structure with positive deviations from the MIM around 10 hPa and negative deviations below and above that level.

4.4.6 Recommendations: N_2O

N_2O seasonal cycles are often used for model-measurement comparisons. The seasonal cycles derived from the different datasets at 100 and 50 hPa show relatively good agreement in their mean values. In the extra-tropics, the different instruments' climatologies also agree in the amplitudes in the seasonal cycle. Some of the discrepancies may also be explained by the instruments' different vertical resolutions (for which model evaluations could in principle account for). Nevertheless, to gain more confidence in the N_2O seasonal cycles derived from satellite observations and to use them as model diagnostic, they will have to be validated against other independent observations if available.

Interannual variability is well captured by the different instruments except for ACE-FTS, indicating that once the biases are removed, the instruments show high enough quality for being merged into longer climate data records. Interannual variability is less pronounced and hence less well captured by the instruments in the lower stratosphere around 100 hPa, and especially in the tropics.

4.5 Trichlorofluoromethane – CCl_3F (CFC-11)

Trichlorofluoromethane (commonly named CFC-11) belongs to the chlorofluorocarbons (CFCs), and is an important component of the chlorine-containing ozone-depleting substances. CFC-11 is an anthropogenic compound with virtually no natural background and was emitted as a result of human activity through its widespread use as

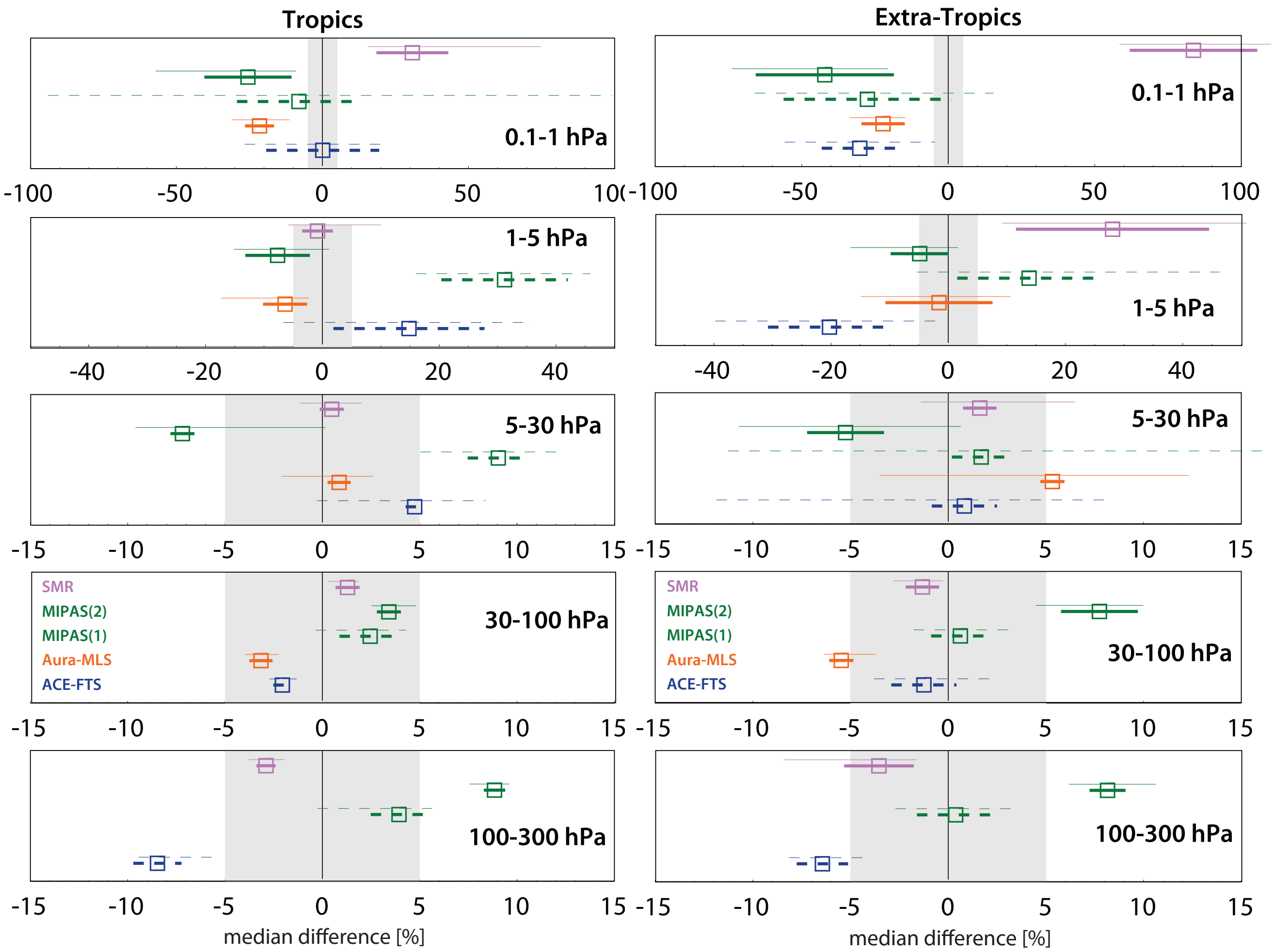


Figure 4.4.7: Inter-instrument differences in N_2O calculated for the tropics (left) (20°S–20°N) and (right) extra-tropics (40°S–80°S and 40°N–80°N) and for five different altitude regions from the UT up to the LM. *Shown are the median (squares), median absolute deviations (MAD, thick lines), and the mean ±1σ ranges (thin lines) of the relative differences between each individual instrument and the MIM calculated over a given latitude and altitude region. The reference period is 2006-2009.*

a refrigerant. Between the 1930s, the beginning of its industrial production, and the mid 1990s the atmospheric concentration of CFC-11 increased steadily. In compliance with the Montreal Protocol in the late 1980s and its subsequent amendments, its manufacture was banned in many countries due to its role in damaging the ozone layer. Consequently, global CFC-11 surface mixing ratios peaked in the mid 1990s and are now slowly decreasing [*WMO*, 2014]. Accordingly, a decrease in the total atmospheric burden of the long-lived CFC-11, with an atmospheric lifetime of 45 years, has been observed from ground-based total-column measurements at the Jungfraujoch station [*WMO*, 2011].

4.5.1 Availability of CFC-11 measurements

Vertically resolved satellite measurements of CFC-11 by the MIPAS instrument started in 2002. From 2004 onwards there are also ACE-FTS measurements available. Both time series extend over approximately 7 years. Additionally, HIRDLS measured CFC-11 from 2005 to 2007. While ACE-FTS and HIRDLS cover only the UTLS and up to 30 hPa into the MS, MIPAS measurements extend through the MS up to 5 hPa. **Tables 4.5.1** and **4.5.2** compile information on the availability of CFC-11 measurements, including time period, altitude range, vertical resolution, and references relevant for the data product used in this report.

4.5.2 CFC-11 evaluations: Zonal mean cross sections, vertical and meridional profiles

Annual zonal mean cross sections for the time period 2005-2007 are analysed to investigate mean biases between the various datasets. Additionally, vertical and meridional profiles are presented.

The annual zonal mean CFC-11 climatologies for 2005-2007 for MIPAS, ACE-FTS, HIRDLS and their MIM are shown in **Figure 4.5.1**. The maximum CFC-11 mixing ratios are found in the troposphere and in the TTL, where air is entrained from the troposphere into the stratosphere. For MIPAS and HIRDLS, the maximum mixing ratios in the TTL are occasionally larger (up to 0.275 ppbv) than those inferred from surface measurements (0.26 ppbv), suggesting a local bias of up to 5%. These discrepancies represent so far unexplained problems in the satellite datasets and dedicated instrument-specific validation studies are required in order explain them. Overall, MIPAS shows the largest mixing ratios in the TTL with a very flat isoline at 100 hPa extending from 30°S to 30°N and a uniform distribution below. Due to the long lifetime of CFC-11, such a uniform distribution in the TTL is expected, in contrast to the local maximum in the upper TTL as seen in the ACE-FTS and HIRDLS climatologies. Simulations with the Whole Atmosphere Community Climate Model (WACCM) in CTM mode driven by Goddard Earth Observing System Model, Version 5 (GEOS-5) data for 2005-2007 confirm the uniform CFC-11 distribution in the TTL as observed by MIPAS. Note that the local maximum in the HIRDLS V6 data does not exist in future versions of HIRDLS data (V7) due to corrections for UTLS aerosols. Above the tropopause, CFC-11 decreases rapidly with isolines roughly parallel to the north-south slope of the tropopause. HIRDLS shows steep gradients in the SH subtropics at the equatorward edge of the surf zone. Note that these steep vertical gradients are also present if the vertical resolution of the HIRDLS climatology is reduced (through smoothing), and are therefore in all likelihood not related to resolution aspects. Simulations with WACCM for 2005-2007 show CFC-11 contours with slopes that are very similar to ones observed by HIRDLS. For the ACE-FTS climatology, tropical CFC-11 does not decrease between 50 and 30 hPa and therefore the isolines in the inner tropics look quite different compared to the two other instruments. Note that this might be related to the fact that the retrieval has a fixed altitude limit at all latitudes (rather than extending to higher altitudes in the tropics), impacting the highest ACE-FTS levels in the climatology. Also, sampling rate for ACE-FTS in the tropics is much lower than for HIRDLS and MIPAS.

***Table 4.5.1:** Available CFC-11 measurement records from limb-sounding satellite instruments between 1978 and 2010.* *The red filling of the grid boxes indicates the temporal and vertical coverage of the respective instruments.*

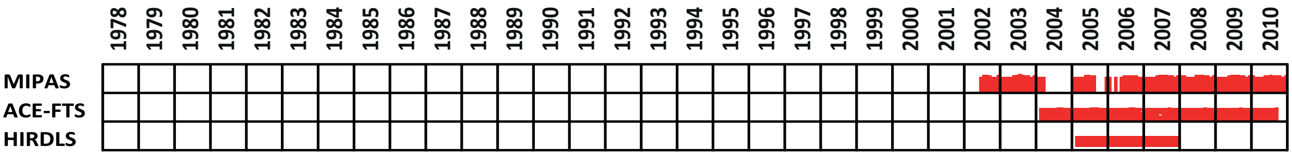


***Table 4.5.2:** Time period, vertical range, vertical resolution, references and other comments for CFC-11 measurements.*

Instrument	Time period	Vertical range	Vertical resolution	References	Additional comments
MIPAS MIPAS(1) V10 MIPAS(2) V220	 Mar 02 – Mar 04 Jan 05 – Apr 12	 ~300 – 1 hPa (10 – 50 km)	 4 km	*Kellmann et al.*, 2012	change in measurement mode in 2004
ACE-FTS V2.2	Mar 04 –	6 – 28 km	3 – 4 km	*Mahieu et al.*, 2008	
HIRDLS V6.0	Jan 05 – Mar 08	316 – 10 hPa (10 – 30 km)	1 km	*Gille and Gray*, 2011	

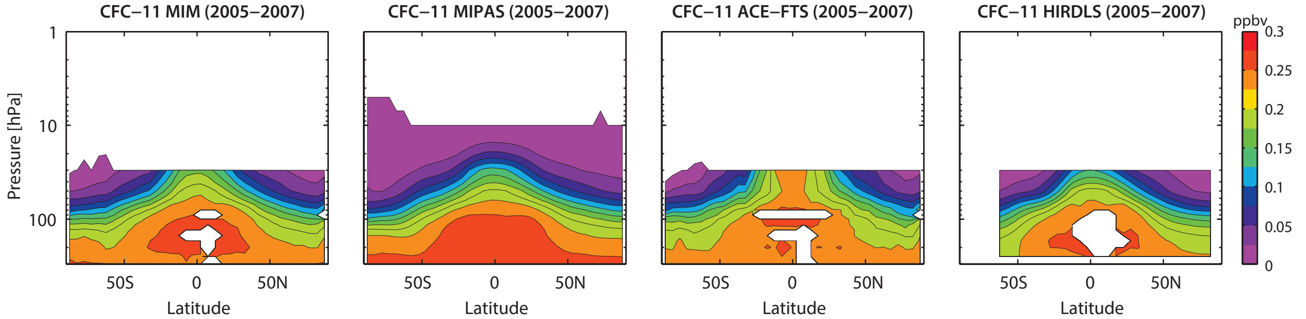


Figure 4.5.1: ***Cross sections of annual zonal mean CFC-11 for 2005-2007.*** *Annual zonal mean CFC-11 cross sections are shown for the MIM, MIPAS, ACE-FTS, and HIRDLS. The MIM is only displayed for regions where at least two instruments provide measurements.*

Differences of the individual datasets relative to the MIM are shown in **Figure 4.5.2**. The instruments agree well below 50 hPa in the tropics and below 100 hPa at higher latitudes, with differences to the MIM of up to ±5%. In particular, ACE-FTS and HIRDLS show excellent agreement with each other, with differences with respect to their MIM of only ±2.5% (see **Figure A4.5.1** in *Appendix A4*), while MIPAS exhibits larger differences when compared to the other two datasets. Above the tropopause, the relative differences increase slowly as the absolute CFC-11 abundance decreases. In the tropics above 50 hPa, ACE-FTS shows considerable disagreement with the other two datasets with differences to the MIM of up to +50% at the highest level (30 hPa). MIPAS and HIRDLS agree very well with each other, and if compared directly display differences with respect to their MIM of only up to ±5% (see **Figure A4.5.1** in *Appendix A4*). In the extra-tropical LS the situation reverses; MIPAS and ACE-FTS agree quite well while HIRDLS diverges from the other two datasets and exhibits differences relative to the MIM of up to -50%.

Monthly mean vertical CFC-11 profiles in tropical and mid-latitude regions are shown in **Figure 4.5.3**, together with their differences relative to the MIM. The profiles confirm that all three instruments agree very well below 100 hPa, with MIPAS values about 5-10% larger than the other two datasets. Above the tropopause, the monthly mean values show larger differences consistent with the annual mean values. The monthly mean profiles show that ACE-FTS in the tropics and HIRDLS in the mid-latitudes deviate strongly from the two other datasets in the respective regions. In both cases, the deviations become noticeable above the level where the vertical gradient changes and the background CFC-11 decreases more rapidly, which is about 70-50 hPa in the tropics and around 100 hPa in the mid-latitudes.

Figure 4.5.4 shows the latitudinal structure of the relative differences for the month August, as an example. For all levels, except for 200 hPa, the differences are lowest in the tropics and increase in the mid-latitudes and polar regions as one would expect based on the decreasing CFC-11 abundance. Eye-catching features are the relatively large ACE-FTS difference at 30 hPa in the tropics, also apparent in the differently shaped isolines mentioned earlier, and the steep gradients in HIRDLS CFC-11 between 20°S and 30°S. While the latitudinal gradients of tropical HIRDLS and MIPAS data are quite different, both datasets show a small plateau of nearly constant mixing ratios between 40°-50°S, however at different mixing ratio values. At 70 and 200 hPa, the differences in mid-latitudes and polar regions are considerably smaller than at 30 hPa. At 200 hPa, the largest differences can be observed in the respective winter hemisphere high latitudes, a characteristic which is confirmed by monthly mean evaluations for NH winter (see **Figure A4.5.2** in *Appendix A4*).

4.5.3 CFC-11 evaluations: Interannual variability

Tropical time series of monthly mean values and deseasonalised anomalies at 30 hPa (**Figure 4.5.5**) can be

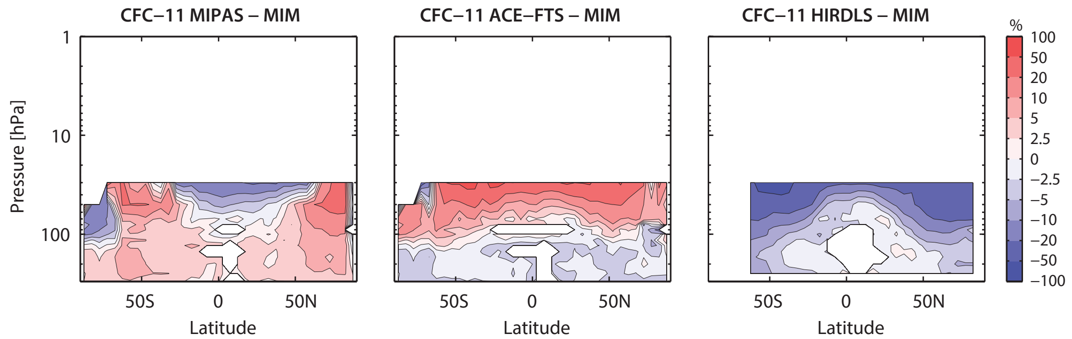


Figure 4.5.2: ***Cross sections of annual zonal mean CFC-11 differences for 2005-2007.*** *Annual zonal mean CFC-11 differences between the individual instruments (MIPAS, ACE-FTS, and HIRDLS) and the MIM are shown.*

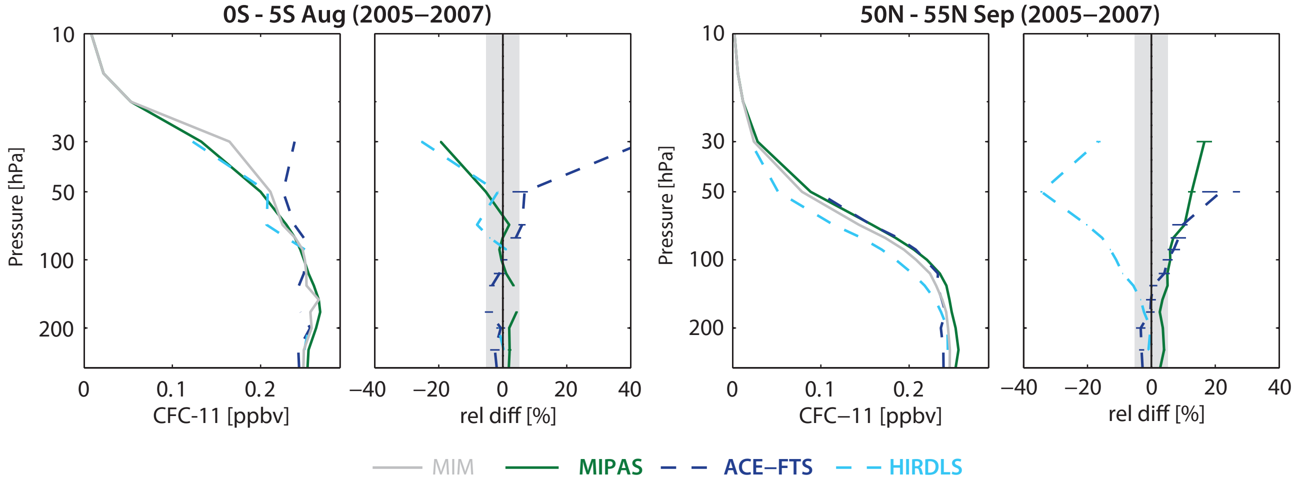


Figure 4.5.3: ***Profiles of monthly zonal mean CFC-11 for 2005-2007.*** *Vertical CFC-11 profiles for 0°S-5°S, August and 50°N-55°N September are shown together with their relative differences from the MIM. The grey shading indicates the ±5% difference range. Bars indicate the uncertainties in the relative differences.*

used to analyse seasonal and interannual variability. Most of the variability in the tropical time series is caused by interannual variations with only weak contributions from the annual cycle as the similarity of the seasonalised and deseasonalised time series reveals. The variability of the MIPAS CFC-11 time series is dominated by an approximately 2-year long cycle which is presumably linked to vertical velocity perturbations caused by the QBO. Perturbations of vertical transport can influence the distribution of trace gases with a significant vertical gradient and a long photochemical lifetime [*Randel*, 1990; *Salby et al.*, 1990], both characteristics of CFC-11. The other two datasets seem to also display the quasi-biennial cycle, although due to the shortness of the HIRDLS time series (three years) and the frequent data gaps in ACE-FTS, an unambiguous conclusion is impossible. The QBO signal is strong at the MS levels between 20 and 50 hPa and vanishes at around 70 hPa (not shown here). Interannual variability decreases with decreasing altitude, and at 200 hPa (not shown here) the long term change of CFC-11 is the dominant signal.

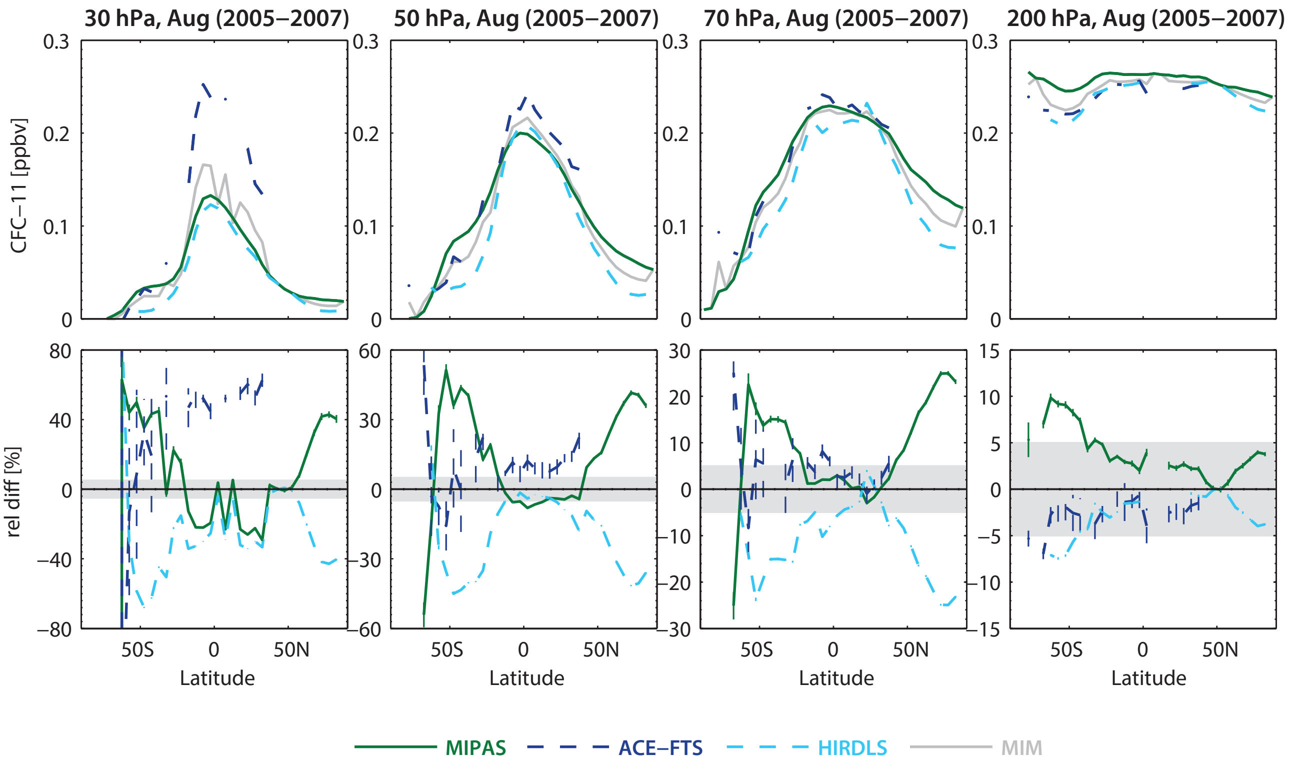


Figure 4.5.4: ***Meridional profiles of monthly zonal mean CFC-11 for 2005-2007.*** *Meridional CFC-11 profiles at 30, 50, 70 and 100 hPa for August are shown in the upper row. Relative differences between the individual instruments (MIPAS, ACE-FTS, and HIRDLS) and the MIM profiles are shown in the lower row. The grey shading indicates the ±5% difference range. Bars indicate the uncertainties in the relative differences.*

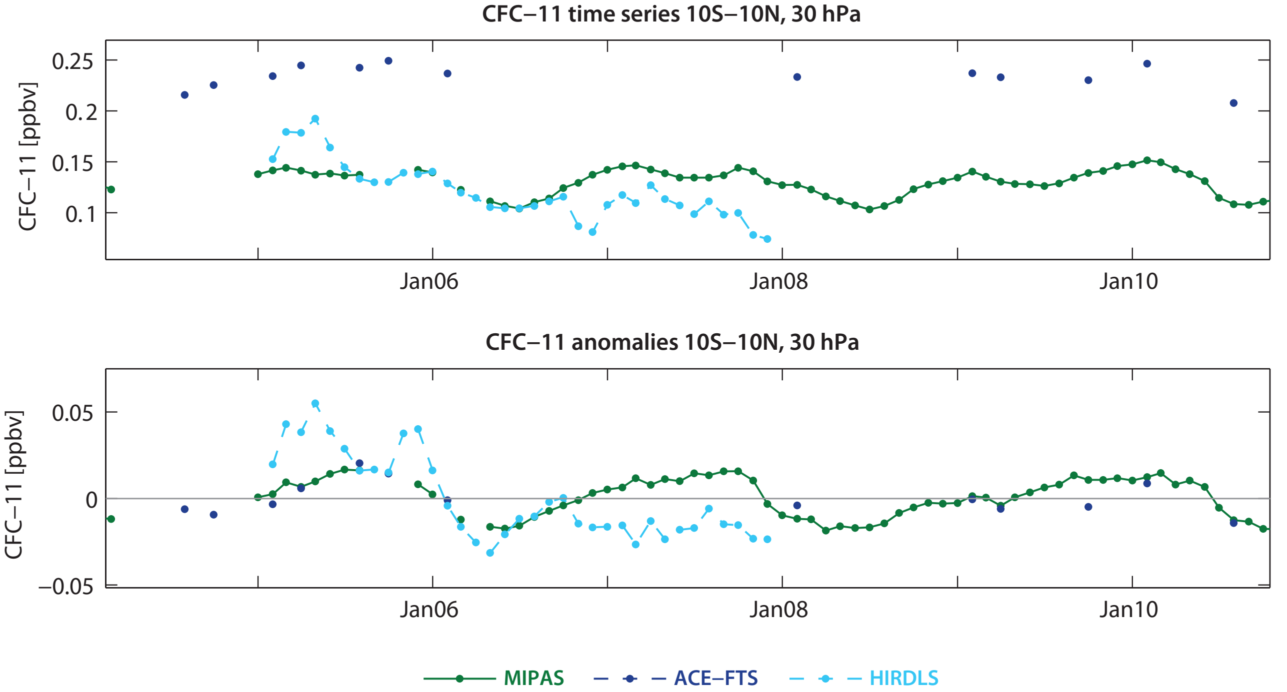


Figure 4.5.5: Time series of CFC-11 monthly zonal mean values and deseasonalised anomalies in the tropics. *Monthly mean values (upper panel) and deseasonalised anomalies (lower panel) of CFC-11 between 10°S – 10°N at 30 hPa.*

Figure 4.5.6 shows the CFC-11 time series of NH high latitude monthly mean values and deseasonalised anomalies at 100 hPa. The seasonal cycle (upper panel) with a minimum in late winter/early spring and a maximum in late summer is the dominant signal while interannual variations are small. The seasonal cycle, caused by descent of aged air in the winter polar vortex, is captured by all three datasets. HIRDLS shows overall lower values and also a smaller amplitude of the signal for the three years of overlap with ACE-FTS and MIPAS. Interannual anomalies (lower panel) are weak, however, most pronounced during NH winter as indicated by all three instruments. Evaluations of ACE-FTS and MIPAS time series at the SH high latitudes reveal similar results with signals in the seasonal cycle and peaks of interannual variability shifted by 6 months (see **Figure A4.5.3** in *Appendix A4*). Major difference to the NH is that ACE-FTS does not detect the seasonal cycle as it is observed by MIPAS.

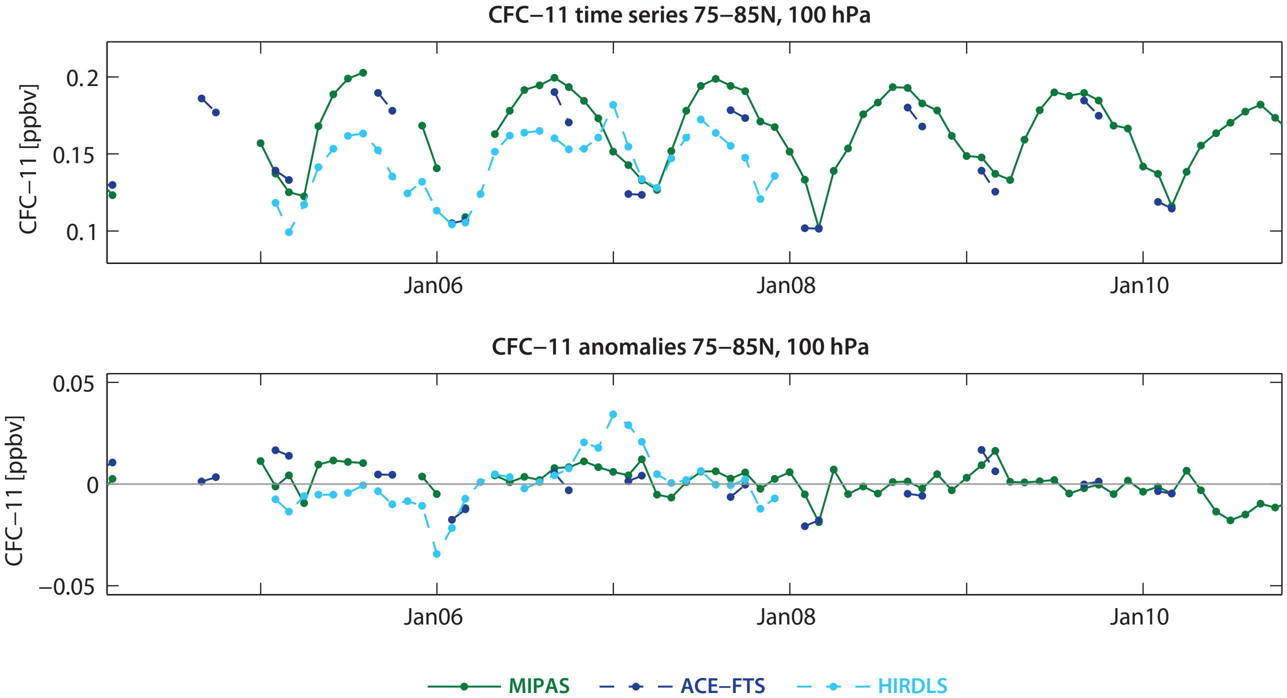


Figure 4.5.6: Time series of CFC-11 monthly zonal mean values and deseasonalised anomalies at NH high latitudes. *Monthly mean values (upper panel) and deseasonalised anomalies (lower panel) of CFC-11 between 75°N – 85°N at 100 hPa.*

4.5.4 Summary and conclusions: CFC-11

A comparison of three CFC-11 profile climatologies (MIPAS, ACE-FTS, HIRDLS) has been carried out. Overall findings on the systematic uncertainty in our knowledge of the CFC-11 mean state and important characteristics of the individual datasets are presented in the following summary including two synopsis plots. The first summary plot (**Figure 4.5.7**) provides information on the mean state and its uncertainty derived from the spread between the datasets. The second summary plot (**Figure 4.5.8**) shows specific inter-instrument differences in form of the deviations of the instrument climatologies to the MIM climatology. For each instrument and selected region the deviation to the MIM is given in form of the median (mean) difference over all grid points in this region. Additionally, for each instrument the spread of the differences over all grid points in this region is presented. Note that both pieces of information (average deviation and spread) are important for a meaningful assessment of inter-instrument differences. A detailed description of the summary plot evaluations can be found in *Section 3.3.5*.

Atmospheric mean state

The uncertainty in our knowledge of the atmospheric CFC-11 annual mean state is smallest below 50 hPa in the tropics and below 100 hPa in the extra-tropics. The evaluation of three datasets for the time period 2005-2007 reveals a 1σ multi-instrument spread in this region of less than ±5% (**Figure 4.5.7**). Maximum CFC-11 mixing ratios are found in the tropical TTL, with values up to 0.275 ppbv, potentially demonstrating a high bias compared to surface measurements. Since CFC-11 has a very long lifetime, the trace gas is expected to be distributed uniformly in the TTL as shown by MIPAS, and not to exhibit a local maximum in the upper TTL as seen in the ACE-FTS or HIRDLS climatologies. In the tropical LS, the spread between the datasets increases quickly with increasing altitude, reaching ±30% at 30 hPa. The absolute differences between the datasets are largest here, with deviations between 0.15 and 0.25 ppb due to high ACE-FTS values at 30 hPa, very likely related to retrieval issues. In the mid-latitude LS between 100 hPa and 50 hPa, mixing ratios decrease but absolute deviations increase slightly compared with the atmospheric region below 100 hPa. As a result, the relative spread is about 10%. Above 50 hPa, however, a large relative spread of up to ±50% exists for very low background values of up to 0.05 ppb.

Instrument-specific conclusions

The **MIPAS** climatology shows overall a very good agreement when compared to the other two instruments. In the region of low inter-instrument spread (below 100-50 hPa), MIPAS displays slightly higher values and in the region of large inter-instrument spread it is in the middle of the range. MIPAS has weaker meridional gradients at 200 hPa in the respective winter hemisphere than the other two instruments.

The **ACE-FTS** climatology shows a very good agreement with the other two datasets below 50 hPa. For tropical ACE-FTS there is no CFC-11 decrease between 50 and 30 hPa leading to a relatively large positive difference in the tropical LS (average of +25%). Similarly, in the mid-latitudes ACE-FTS does not decrease as fast as the comparison instruments with positive average deviations of +25%. While ACE-FTS shows similar seasonal variations as MIPAS and HIRDLS at the NH high latitudes, it does not display seasonal variations at high SH latitudes.

The **HIRDLS** climatology agrees very well with the other two datasets in the tropics below 50 hPa and in the mid-latitudes below 100 hPa. However, outside of this region HIRDLS displays considerably lower values especially in the mid-latitudes where average deviations range around -30% and individual deviations can be as large as -50%. These large deviations are related to relatively steep subtropical isolines.

A comparison of the key findings for CFC-11 and CFC-12 can be found at the end of *Section 4.6* on CFC-12.

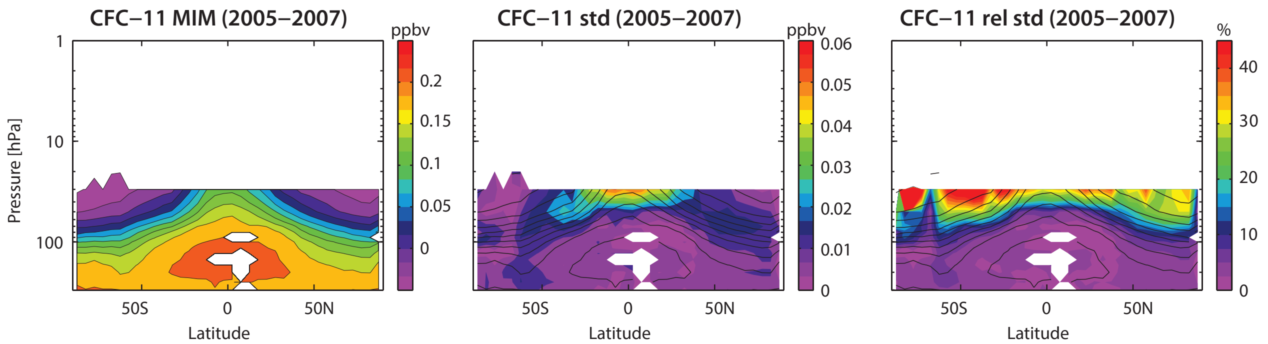


__Figure 4.5.7: Summary of CFC-11 annual zonal mean state for 2005-2007.__ Shown are the annual zonal mean cross section for the MIM of CFC-11 (left panel), the standard deviation over all three instruments (middle panel), and the relative standard deviation with respect to the MIM (right panel). Black contour lines in the two rightmost panels give the MIM distribution. Instruments included are MIPAS, ACE-FTS, and HIRDLS. The MIM and standard deviation are only displayed for regions where at least two instruments provide measurements.

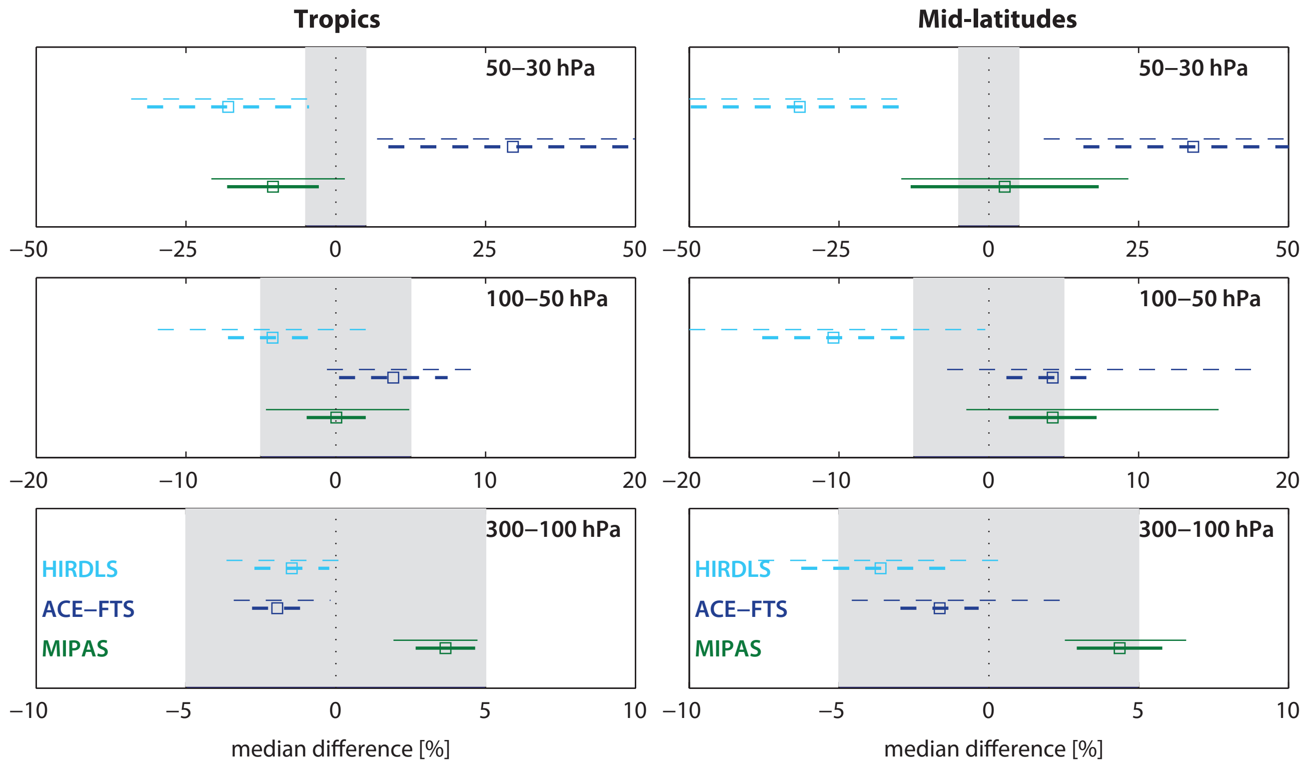


Figure 4.5.8: Summary CFC-11 differences for 2005-2007. *Over a given latitude and altitude region the median (squares), median absolute deviation (MAD, thick lines), and the standard deviation (thin lines) of the monthly mean relative differences between an individual instrument-climatology and the MIM are calculated. Results are shown for the tropics (30°S-30°N) and midlatitudes (30°S-60°S and 30°N-60°N) and for 3 different altitude regions from the UT up to the MS between 300 and 30 hPa for the reference period 2005-2007.*

4.6 Dichlorodifluoromethane – CCl_2F_2 (CFC-12)

Dichlorodifluoromethane is a CFC originally used as a refrigerant and aerosol spray propellant. As is the case for CFC-11, CFC-12 is an anthropogenic source gas, which is distributed and accumulated in the troposphere before being transported into the stratosphere. Once in the stratosphere, both gases are converted into reactive halogens and cause severe ozone depletion. As a consequence of the Montreal Protocol and its Amendments and Adjustments, CFC-12 abundance has plateaued in the atmosphere. However, due to its longer lifetime (100 years) and emissions from CFC-12 banks, the decline in CFC-12 abundance is delayed compared to CFC-11, which peaked in the early 90's [*WMO*, 2014].

4.6.1 Availability of CFC-12 measurements

Measurements of CFC-12 are available from MIPAS, ACE-FTS and HIRDLS, with the two first time series currently extending over 7 years and HIRDLS covering 3 years. MIPAS measurements extend up to 1 hPa while the other two instruments extend only to 15 hPa. **Tables 4.6.1** and **4.6.2** compile information on the availability of CFC-12 measurements, including time period, altitude range, vertical resolution, and references relevant for the data product used in this report.

4.6.2 CFC-12 evaluations: Zonal mean cross sections, vertical and meridional profiles

Annual zonal mean cross sections for the time period 2005-2007 are analysed to investigate mean biases between the various datasets. Additionally, vertical and meridional profiles are evaluated.

Figure 4.6.1 shows the annual zonal mean CFC-12 climatologies for 2005-2007 for all available measurements. Maximum CFC-12 values are reported in all three climatologies in the TTL, and for MIPAS in the extra-tropical UTLS, similar to what has been observed for CFC-11. For the MIPAS maximum (0.57 ppbv) and the HIRDLS maximum (0.56 ppbv), the tropical mixing ratios exceed maximum surface measurements (0.54 ppbv) indicating a high bias of the two satellite datasets below 100 hPa of up to 5%. While for ACE-FTS and MIPAS the tropical abundances fall below 0.5 ppbv at 50 hPa, for HIRDLS such values are found up to 30 hPa. The larger tropical CFC-12 values for HIRDLS are accompanied by steeper subtropical gradients similar to what has been observed for CFC-11 (see also discussion in *Section 4.6.1*). ACE-FTS shows elevated values at the highest retrieval level (15 hPa) when compared to other two datasets related to the imposed maximum retrieval altitude for all latitudes (as described in *Section 4.5*). Additionally, the solar occultation sounder has noisier isolines very likely related to sampling density with some kinks at the 130 hPa level.

Table 4.6.1: Available CFC-12 measurement records from limb-sounding satellite instruments between 1978 and 2010. *The red filling of the grid boxes indicates the temporal and vertical coverage of the respective instrument.*

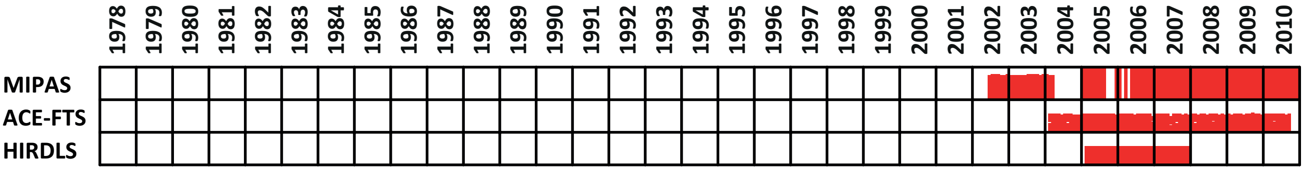


Table 4.6.2: Time period, vertical range, vertical resolution, references and other comments for CFC-12 measurements.

Instrument	Time period	Vertical range	Vertical resolution	References	Additional comments
MIPAS MIPAS(1) V10 MIPAS(2) V220	 Mar 02 – Mar 04 Jan 05 – Apr 12	 ~ 300 – 5 hPa (10 – 35 km)	 4 km	 *Kellmann et al.*, 2012	 change in measurement mode in 2004
ACE-FTS V2.2	Mar 04 –	5 – 22 km	3 – 4 km	*Mahieu et al.*, 2008	
HIRDLS V6.0	Jan 05 – Mar 08	316 – 26.1 hPa (10 – 24 km)	1 km	*Gille and Gray*, 2011	

The differences of all three datasets with respect to the MIM are displayed in **Figure 4.6.2**. Below 100 hPa, MIPAS and HIRDLS show excellent agreement with a positive departure from the MIM of up to +5%, while ACE-FTS has a negative departure from the MIM of up to -5%, in most cases, and -10%, occasionally. In general, these relatively small differences increase in the LS/MS with altitude, especially in the extratropics. Here, the differences change sign at around 100 hPa and ACE-FTS is larger when compared to other two datasets. In the NH, ACE-FTS is about 20% larger and MIPAS again shows excellent agreement with HIRDLS, whereas in the SH, differences between ACE-FTS (+50%) and HIRDLS (-50%) are large, and MIPAS is found in the middle range. In the tropics, largest CFC-12 abundances are reported by HIRDLS (+10%) as already noted above and smallest values are reported by MIPAS (-10%).

Monthly mean vertical CFC-12 profiles at higher latitudes in spring are shown in **Figure 4.6.3** together with their differences relative to the MIM. The NH profiles show a very good agreement for all three instruments with differences below ±10% over the entire vertical range and excellent agreement between MIPAS and HIRDLS. In the SH, all three instruments agree very well below 100 hPa. Differences increase above this level and in the MS relatively large differences for HIRDLS (negative) and ACE-FTS (positive) are the dominant signals, while MIPAS shows only a small positive departure from the MIM. The fact that CFC-12 from ACE-FTS at high altitudes does not decrease as fast as the comparison instruments is consistent with results from *Mahieu et al.* [2008].

In **Figure 4.6.4** meridional CFC-12 profiles at 30, 50, 70 and 200 hPa are shown. For the upper levels, HIRDLS shows steeper meridional gradients than the other two instruments, while MIPAS displays a small plateau between 40°S-50°S. Relative differences maximise at high latitudes where CFC-12 abundance is low. In the MS, HIRDLS exhibits larger values in the tropics and lower values in the extratropics compared to the other two instruments, while MIPAS and ACE-FTS agree mostly very well. Relative differences decrease with decreasing altitude and are quite small at 200 hPa (≤ 5%). The lower CFC-12 abundances measured here by ACE-FTS are consistent with previous studies [*Mahieu et al.*, 2008]. Surprisingly, the relative differences at 200 hPa are larger in the winter hemisphere high latitudes (similar to CFC-11), although there is no such strong meridional gradient as observed for the levels above. These differences result from the fact that ACE-FTS and HIRDLS decrease in poleward direction, while MIPAS values at high latitudes are very similar to the tropical abundances. Such different meridional gradients

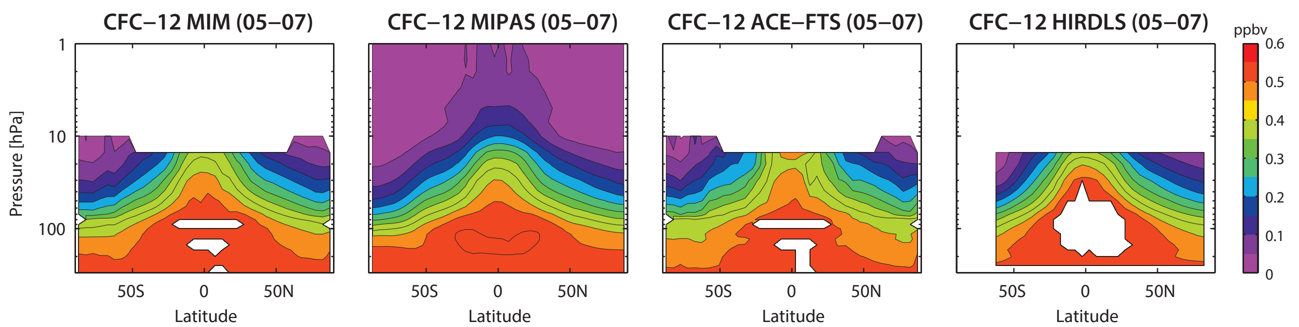


Figure 4.6.1: Cross sections of annual zonal mean CFC-12 for 2005-2007. *Annual zonal mean CFC-12 cross sections are shown for the MIM, MIPAS, ACE-FTS, and HIRDLS. The MIM is only displayed for regions where at least two instruments provide measurements.*

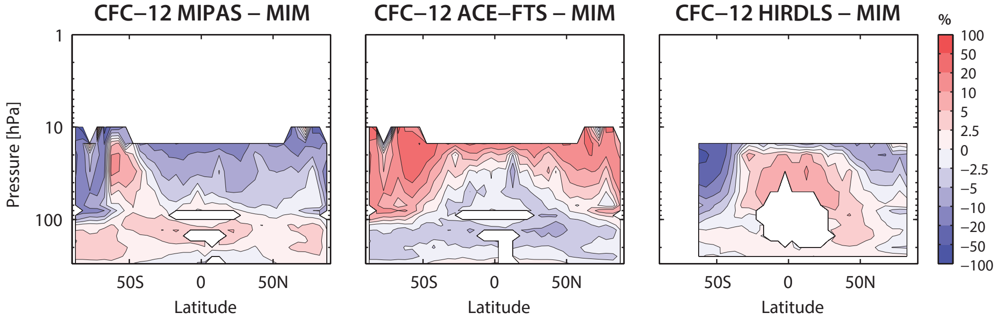


Figure 4.6.2: Cross sections of annual zonal mean CFC-12 differences for 2005-2007. Annual zonal mean CFC-12 differences between the individual instruments (MIPAS, ACE-FTS, and HIRDLS) and the MIM are shown.

at high latitudes are also observed for other months (see **Figure A4.6.2** in *Appendix A4* for December) and often the deviations are most pronounced in the respective winter/spring hemisphere.

4.6.3 CFC-12 evaluations: Interannual variability and seasonal cycle

Figure 4.6.5 shows the tropical time series of monthly mean values and deseasonalised anomalies at 20 hPa in order to analyse the seasonal and interannual variability. The tropical time series is dominated by interannual variations with only weak contributions from the annual cycle as a comparison of the two panels and the similarity of the seasonalised and deseasonalised time series reveals. As already observed for CFC-11, the MIPAS and HIRDLS time series show an approximately 2-year long cycle, which is assumed to be related to QBO transport variations. ACE-FTS measurements do not clearly reveal the same cycle, which might be related to noise near the top of the vertical range. Instead, ACE-FTS shows a stronger long-term change than the other two time series with a step-like decrease of 1 ppbv at the end of 2008. Note that below 70 hPa the QBO signal disappears and the month-to-month fluctuations together with the trend become the dominant mode of variability. In the UT, MIPAS data shows an offset separating the data before 2004 and after 2004, which are based on two different measurement modes. Note that this offset does not exist at higher latitudes. Since ACE-FTS measurements only started in 2004 a comparison of the early MIPAS data with another dataset (and therefore an attribution of the offset to the MIPAS measurement modes) is not possible.

At NH high latitudes (**Figure 4.6.6**), the dominant signal is the seasonal cycle with a minimum in late winter/early spring and a maximum in late summer related to the diabatic descent of aged air with the Brewer-Dobson circulation. HIRDLS and MIPAS show approximately the same seasonal cycle with the largest disagreement at the end of the HIRDLS measurement time period in autumn 2007, where HIRDLS shows a 3 months earlier decline of CFC-12 values. ACE-FTS measurements do not allow for a detailed analysis of the seasonal signal, but it becomes clear that there is no pronounced minimum in late winter in the ACE-FTS time series. Interannual anomalies are quite small for all datasets and peak in late winter/early spring. Although covering different time periods, MIPAS and HIRDLS interannual signals are roughly consistent with the largest disagreement in late 2007. Evaluations of ACE-FTS and MIPAS time series in the SH high latitudes reveal similar results with signals in the seasonal cycle and peaks of interannual variability shifted by 6 months (see **Figure A4.6.3** in *Appendix A4*).

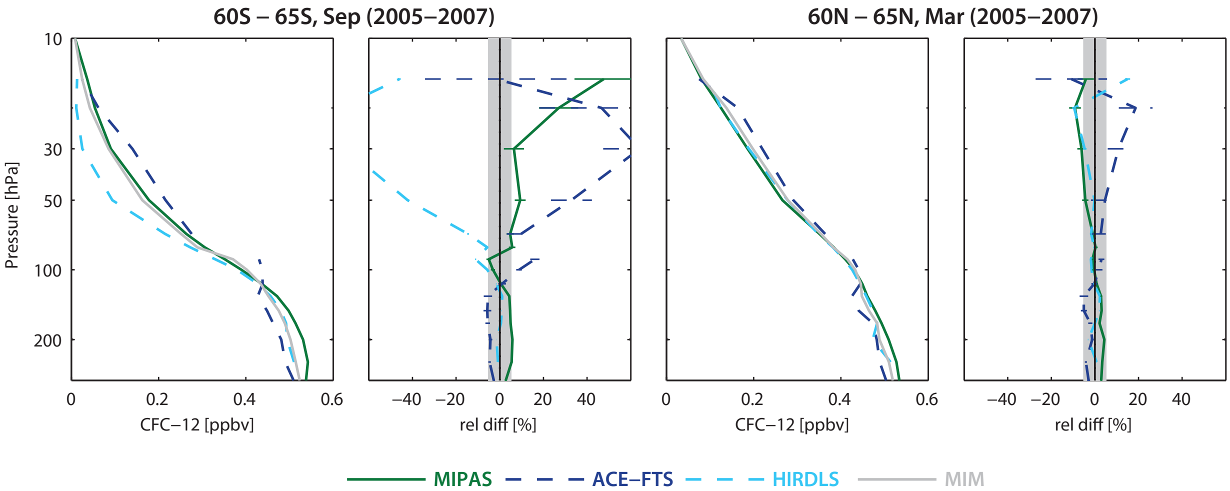


Figure 4.6.3: Profiles of zonal mean CFC-12 for 2005-2007. Zonal mean CFC-12 profiles for 60°S-65°S in September and 60°N-65°N in March are shown together with their relative differences from the MIM. The grey shading indicates the ±5% difference range. Bars indicate the uncertainties in the relative differences.

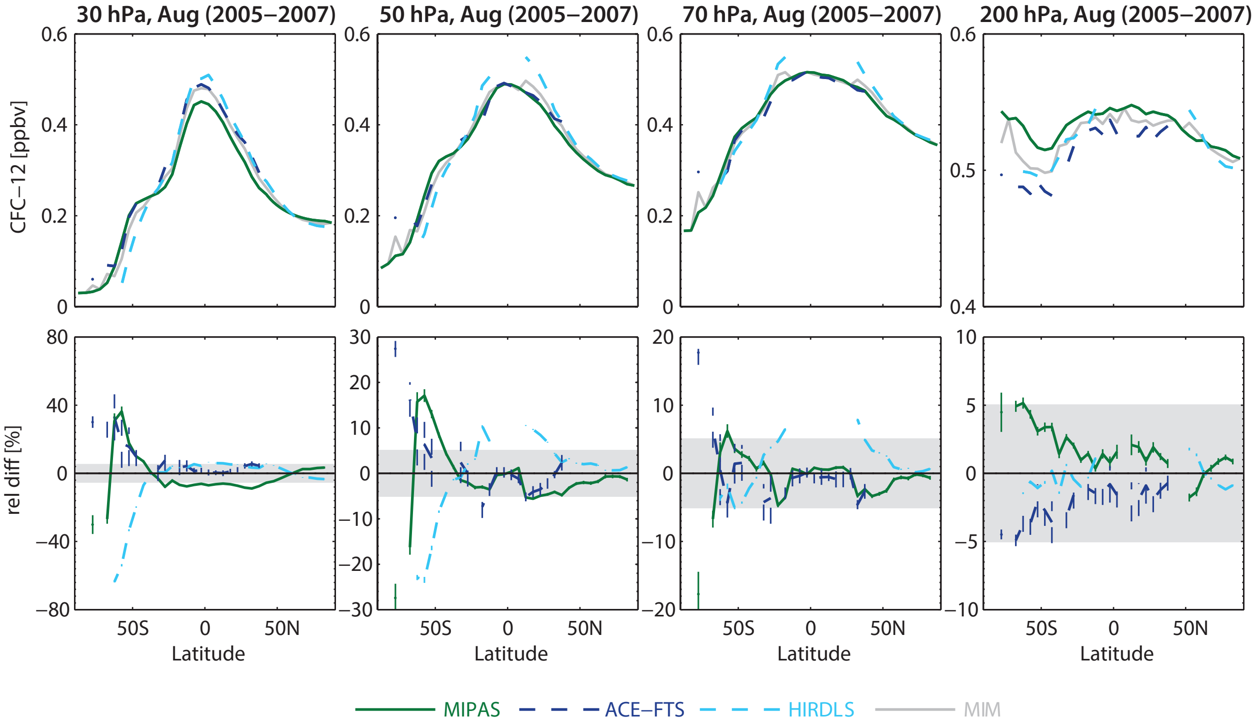


Figure 4.6.4: ***Meridional profiles of zonal mean CFC-12 for 2005-2007.*** *Meridional zonal mean CFC-12 profiles at 30, 50, 70, and 200 hPa for August are shown in the upper row. Relative differences between the individual instruments (MIPAS, ACE-FTS, and HIRDLS) and the MIM profiles are shown in the lower row. The grey shading indicates the ±5% difference range. Bars indicate the uncertainties in the relative differences.*

4.6.4 Summary and conclusions: CFC-12

A comparison of three CFC-12 profile climatologies (MIPAS, ACE-FTS, HIRDLS) has been carried out. Overall findings on the systematic uncertainty in our knowledge of the mean state of CFC-12, and important characteristics of the individual datasets are presented in the following summary, including two synopsis plots. The first summary plot (**Figure 4.6.7**) provides information on the mean state and its uncertainty derived from the spread between the datasets. The second summary plot (**Figure 4.6.8a** and **b**) shows specific inter-instrument differences in form of the deviations of the instrument climatologies from the MIM

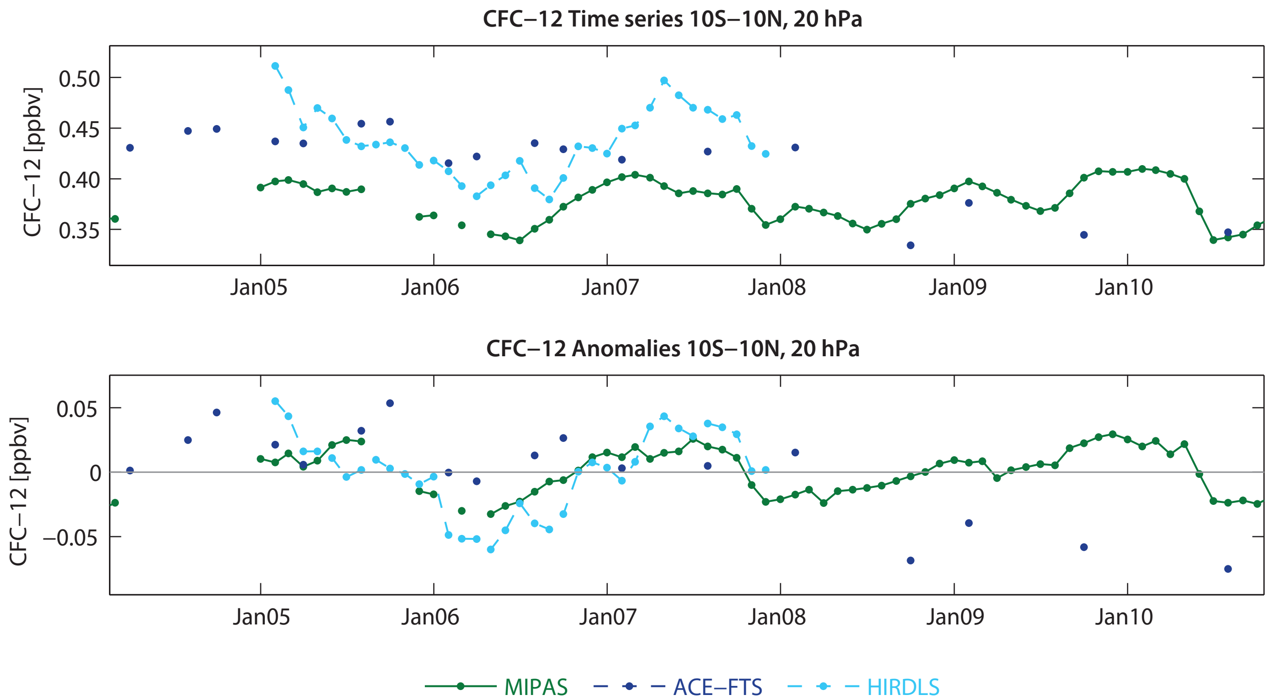


Figure 4.6.5: ***Time series of CFC-12 monthly mean values and deseasonalised anomalies in the tropics.*** *Monthly mean values (upper panel) and deseasonalised anomalies (lower panel) of CFC-12 between 10°S – 10°N at 20 hPa.*

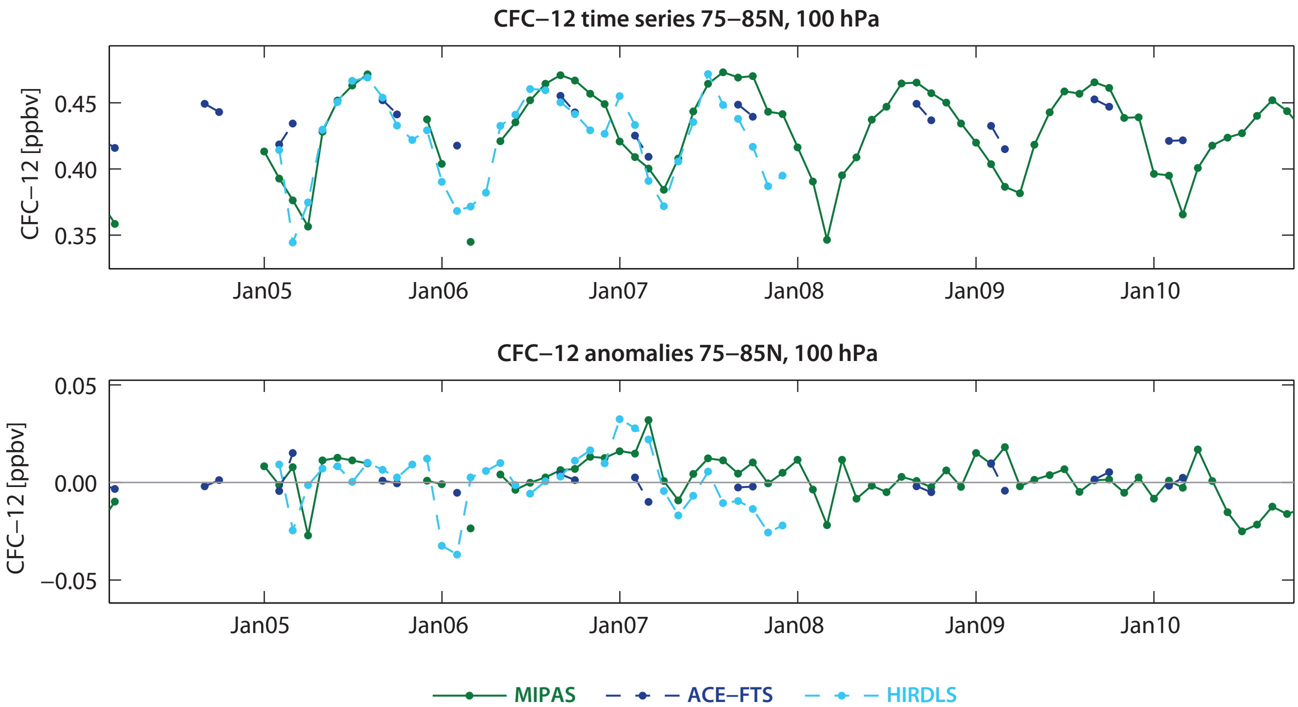


Figure 4.6.6: Time series of CFC-12 monthly mean values and deseasonalised anomalies at NH high latitudes. *Monthly mean values (upper panel) and deseasonalised anomalies (lower panel) of CFC-12 between 75°N – 85°N at 100 hPa.*

climatology. For each instrument and selected region, the deviation to the MIM is given in form of the median (mean) difference over all grid points in this region. Additionally, for each instrument the spread of the differences over all grid points in this region is shown. Note that both pieces of information (average deviation and spread) are important for a meaningful assessment of inter-instrument differences. A detailed description of the summary plot evaluations can be found in *Section 3.3.5*.

Atmospheric mean state

The uncertainty in our knowledge of the annual mean state of atmospheric CFC-12 is smallest below 100 hPa. The evaluation of three datasets for the time period 2005-2007 reveals a 1σ multi-instrument spread in this region of less than ±5%, and often even less than ±2.5% (**Figure 4.6.7**). Maximum CFC-12 mixing ratios are found in the TTL with values up to 0.6 ppbv, indicating a potential high bias compared to surface measurements. In the region between 100 and 20 hPa, good agreement between all datasets exists in the tropics, in the NH, and in the SH subtropics with a multi-instrument spread of less than ±10%. An exception to this good agreement is the SH extra-tropics. Here, considerable disagreement is found with a 1σ multi-instrument spread of up to ±50%. Note that the better agreement (±20%) south of 60°S is related to the fact that here only two datasets (ACE-FTS and MIPAS) are available, while north of 60°S the evaluations are based on all three datasets.

Instrument-specific conclusions

The **MIPAS** climatology is mostly in the middle range between ACE-FTS and HIRDLS and the only region where it

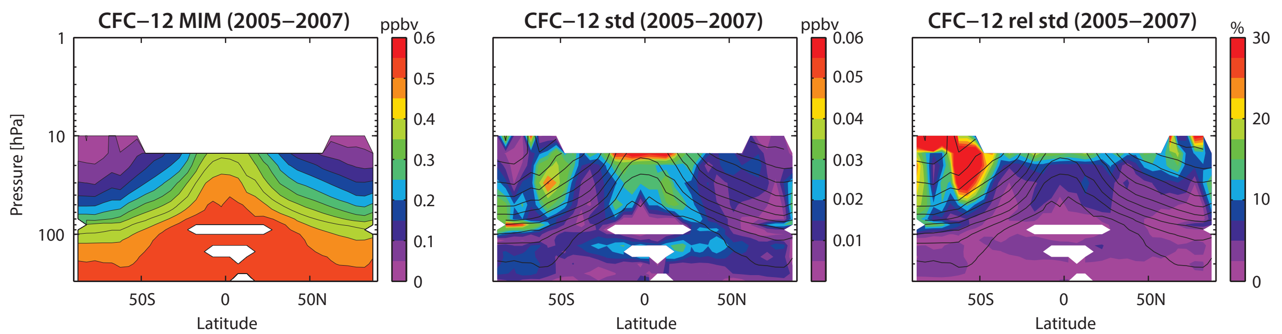


Figure 4.6.7: Summary of CFC-12 annual zonal mean state for 2005-2007. *Shown are the annual zonal mean cross section for the MIM of CFC-12 (left panel), the standard deviation over all three instruments (middle panel), and the relative standard deviation with respect to the MIM (right panel). Black contour lines in the right panels give the MIM distribution. Instruments included are MIPAS, ACE-FTS, and HIRDLS. The MIM and standard deviation are only displayed for regions where at least two instruments provide measurements.*

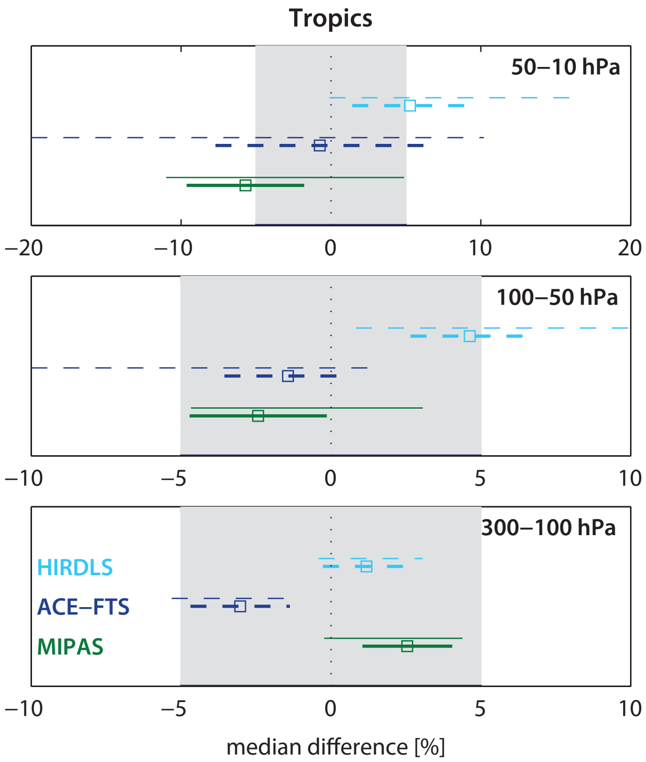


Figure 4.6.8a: Summary CFC-12 differences in the tropics for 2005-2007. *Over a given latitude and altitude region the median (squares), median absolute deviation (MAD, thick lines), and the standard deviation (thin lines) of the monthly mean relative differences between an individual instrument-climatology and the MIM are calculated. Results are shown for the tropics (30°S-30°N) for 3 different altitude regions between 300 and 10 hPa for the reference period 2005-2007.*

shows average deviations larger than +5% is above 50 hPa (**Figure 4.6.8**). While there is a very good overall agreement in the UT, MIPAS has different meridional gradients at 200 hPa than the other two instruments. In the winter hemisphere, MIPAS shows no or only a very weak decrease of values in the poleward direction. Furthermore, data in the UT shows an offset separating the data before 2004 and after 2004, which are based on two different measurement modes.

The **ACE-FTS** climatology shows very good agreement with the other two datasets below 50 hPa. Main features are negative average deviations of up to -2.5% below 100 hPa and

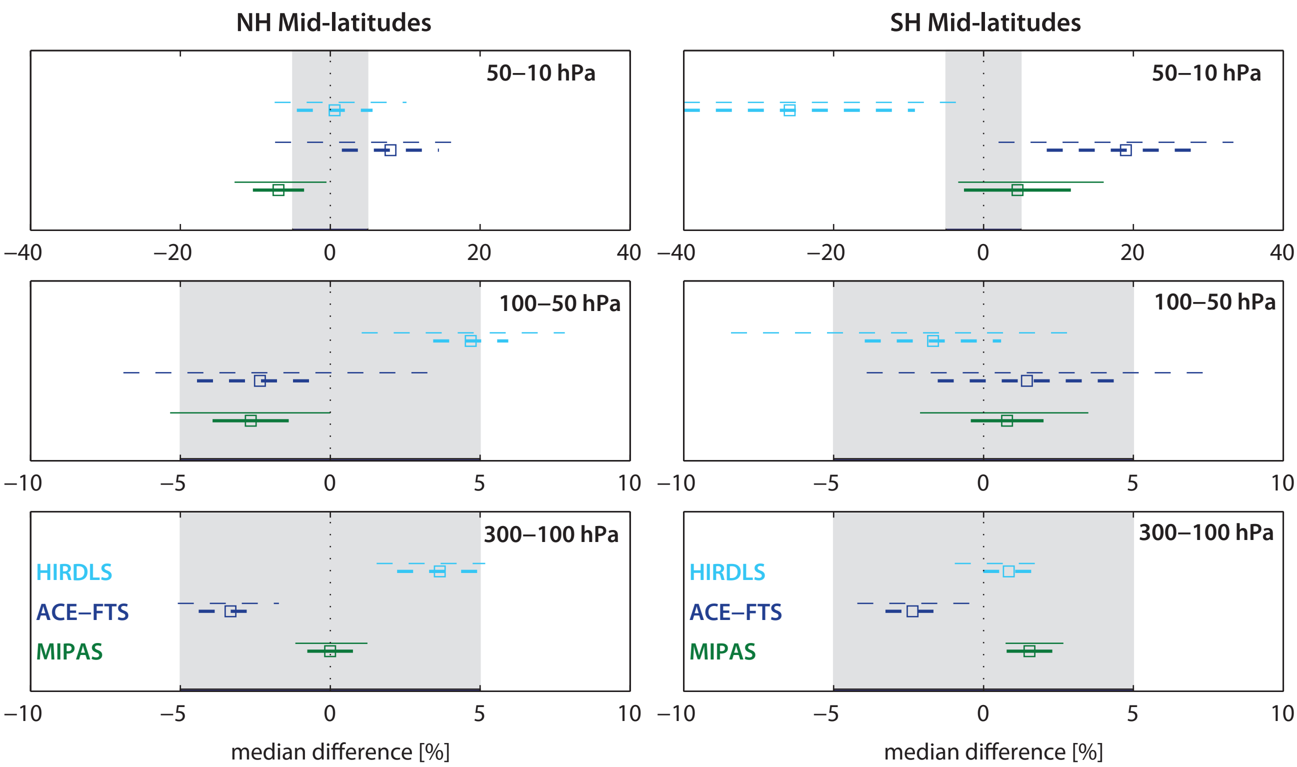


Figure 4.6.8b: Summary CFC-12 differences in mid-latitudes for 2005-2007. *Like Figure 4.6.8a but for NH mid-latitudes (30°N-60°N) and SH mid-latitudes (30°S-60°S).*

excellent agreement with MIPAS between 100 and 50 hPa. Above 50 hPa, ACE-FTS does not decrease as fast as the comparison instruments resulting in positive deviations, which are largest (average of +20%) in the SH. ACE-FTS shows some unrealistic elevated values at the highest retrieval level and no clear signals of seasonal cycle or interannual variability, which might be partially related to the data sampling density.

HIRDLS agrees very well with the other two datasets in most regions of the atmosphere with the largest deviations in the NH mid-latitudes below 50 hPa (+5%). An exception is the SH mid-latitudes above 50 hPa, where HIRDLS is considerably lower than the other instruments, with average deviations of up to -25% and individual deviations of up to -50%. Another important feature of the HIRDLS climatology is steep meridional gradients in the subtropics.

Comparison of key findings for CFC-11 and CFC-12

Overall, there is a better agreement of the CFC-12 climatologies than of the CFC-11 climatologies (*e.g.*, compare **Figures 4.5.4** and **4.6.4**). Differences between the performance in the NH and SH extra-tropical regions exist mostly for CFC-12, where a large inter-instrument spread is found in the SH above 50 hPa. However, for CFC-11 the vertical range extends only to 30 hPa making it more difficult to properly detect such hemispheric differences.

A large number of instrument-specific features can be observed for both trace gases. MIPAS CFC-11 and CFC-12 meridional gradients in the winter hemisphere high latitudes differ from ACE-FTS and HIRDLS in a similar way. ACE-FTS has problems at its highest retrieval level in the tropics for both trace gases, however, more pronounced for CFC-11. HIRDLS climatologies of CFC-11 and CFC-12 both show the steeper gradients in the subtropics, large negative deviations in the mid-latitudes and earlier decline of seasonal cycle in late 2007.

Finally, there are some instrument-specific features which differ considerably between the two CFCs. One example is the seasonal cycle at NH high latitudes, which ACE-FTS can detect for CFC-11 but not for CFC-12.

4.7 Carbon monoxide – CO

Carbon monoxide (CO) is an atmospheric constituent important for tropospheric air quality issues. CO is highly toxic at elevated concentrations. CO has an indirect radiative effect, since it scavenges OH, the cleaning agent of the atmosphere that otherwise would destroy the greenhouse gases CH_4 and O_3 [*Daniel and Solomon*, 1998]. The main sources of CO in the troposphere are the oxidation of methane and non-methane hydrocarbons, and incomplete combustion processes, such as biomass or fossil fuel burning. Due to its intermediate lifetime of about 3 months [*Seinfeld and Pandis*, 2006], CO is much more variable in the troposphere than other long-lived atmospheric constituents, and is therefore often used as a transport tracer of tropospheric air pollution or troposphere-stratosphere exchange in the UTLS region. For the latter purpose, O_3-CO tracer correlations have been frequently used in the past [*Hegglin et al.*, 2009, and references therein]. In the lower stratosphere, CO reaches a background value ranging between 8 and 15 ppbv [*Flocke et al.*, 1999], as determined by the equilibrium between methane oxidation (which forms CO) and CO oxidation (which destroys CO and forms CO_2). In the mesosphere and thermosphere, CO is produced by photolysis of CO_2, which leads to very high mesospheric abundances that are transported into the stratosphere during winter through downwelling within the polar vortex [*Allen et al.*, 2000].

4.7.1 Availability of CO measurements

Only a small set of CO measurements from limb-sounders are available to the SPARC Data Initiative, mainly from the newer generation of instruments (SMR, MIPAS, ACE-FTS, and Aura-MLS). Other datasets not compared within the SPARC Data Initiative are available from SAMS on Nimbus 7 [*Taylor*, 1987], which constitute the earliest measurements (although with a very high noise level), followed by measurements from ATMOS on the Space Shuttle [*Gunson et al.*, 1996], and from ISAMS on UARS [*Taylor et al.*, 1993]. SMR offers a data product at pressure levels smaller than 75 hPa, which is currently limited to one year starting in October 2003 due to a problem with the hardware that stabilises the frequency of the employed 576 GHz heterodyne radiometer [see *Dupuy et al.*, 2004]. A longer time series, corrected for this problem, is being prepared, but was not ready to be included in this assessment.

Tables 4.7.1 and **4.7.2** compile information on the CO data products used in this report, including time period, height range, vertical resolution, and relevant references.

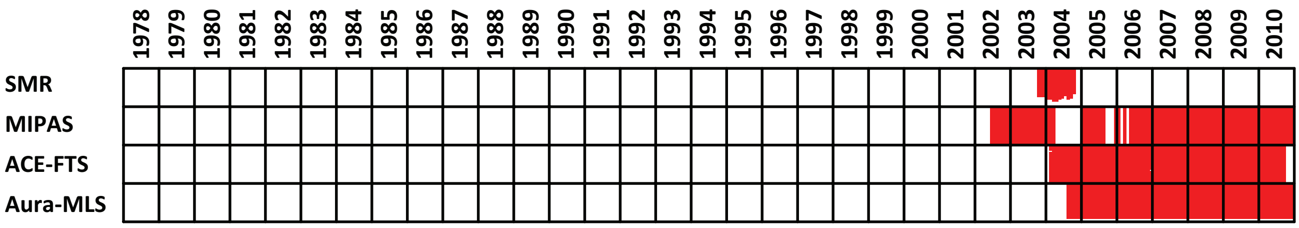
4.7.2 CO evaluations: Zonal mean cross sections, vertical and meridional profiles

Annual zonal mean cross sections for the time period 2006-2009 are analysed to investigate mean differences between the various datasets. SMR and MIPAS(1) are compared to this time period although their measurements were taken during 2003-2004 and 2002-2004, respectively. Additionally, vertical and meridional profiles are evaluated in order to focus on particular height or latitude regions and months.

MIPAS(2), ACE-FTS, Aura-MLS (2006-2009), MIPAS(1) (2002-2004) and SMR (2003-2004)

Figure 4.7.1a shows annual zonal mean CO climatologies for all available measurements. We did not use years prior to 2006 due to data gaps in MIPAS(2), which may influence the overall assessment. Note that SMR and MIPAS(1) are not included in the MIM calculation since the SMR climatology is averaged over one year starting in October 2003

Table 4.7.1: Available CO measurement records from limb-sounding satellite instruments participating in the SPARC Data Initiative between 1978 and 2010. *The red filling in each grid box indicates the temporal and vertical coverage (within the pressure range 300-0.1 hPa) of the respective instrument.*



and the MIPAS(1) climatology over 2002-2004 only. **Figure 4.7.1a** reveals large disagreement among the instruments on the annual zonal mean CO distribution. Nevertheless, common features in the distributions are values around 80-100 ppbv in the upper troposphere, strong vertical gradients across the tropopause, low values of around 15 ppbv in the LS and MS, and strongly increasing values toward the USLM with maxima in the polar regions. As mentioned in the introduction, the high values in the USLM stem from the photodissociation of CO_2 and subsequent downward transport. The mid-infrared sensors MIPAS(2), MIPAS(1), and ACE-FTS agree best. The ACE-FTS measurements show somewhat noisier fields due to the instrument's lower sampling frequency, which limits the smoothness of the climatology especially in regions with strong gradients. SMR, despite its generally higher spatial sampling density and daily global coverage, also exhibits noise in the annual mean climatology, which is due to the fact that CO was retrieved for only ~2 days per month during a limited time period from October 2003 to October 2004. The SMR product furthermore, does not reproduce the low background values of 8-15 ppbv expected in the lower stratosphere, while they are seen the MIPAS and ACE-FTS climatologies. Aura-MLS, on the other hand, shows stratospheric background CO values (<10 ppbv) that are somewhat lower than those from MIPAS and ACE-FTS [see also *Pumphrey et al.*, 2007]. Aura-MLS also shows other features in the climatology that do not agree with the MIPAS and ACE-FTS climatologies. These are local minima in the CO abundance in the SH lowermost stratosphere (around 200 hPa) and in the tropical LM (around 0.2 hPa). In addition, the Aura-MLS and SMR climatologies do not show downward sloping trace gas isopleths (from the tropics to the polar regions) in the LS as typically observed in other long-lived trace gases or the MIPAS and ACE-FTS CO climatologies.

At higher latitudes and altitudes (USLM), CO exhibits much larger vertical gradients than most other trace gas species due to its lower mesospheric source, and very large seasonal and inter-annual variability. Inconsistencies seen in the annual zonal mean distributions of the SMR and MIPAS(1) CO climatologies, may at least partially stem from differences in temporal and spatial sampling. In general, the instruments capture the pronounced seasonal features in the CO distribution well (see **Figures A4.7.1a** and **A4.7.1b** in *Appendix A4*), however, the shortcomings and uncertainties in absolute values as derived from the annual mean can also be seen in the monthly zonal mean evaluations. The same conclusions follow from the evaluation of the latitude-time evolution (see *Section 4.7.4*), as well from a monthly comparison between MIPAS(1) and SMR during late 2003 and early 2004 when their instrumental records overlap and sampling bias is minimised (see **Figures A4.7.2a** and **A4.7.2b** in *Appendix A4*).

The relative differences between the instruments and the MIM are displayed in **Figure 4.7.1.b**. The smallest departures from the MIM are found in the MIPAS climatologies, and are of the order of ±10% through most of the atmosphere, except in the polar MS, where relative differences for MIPAS increase to +20% (and more so for MIPAS(1) than MIPAS(2), likely due to sampling as mentioned above). The ACE-FTS and Aura-MLS climatologies show the opposite behaviour: the ACE-FTS (Aura-MLS) exhibits negative (positive) relative differences from the MIM throughout the LS, US, and LM, and positive (negative) relative differences in the MS. These differences are, however, no larger than

Table 4.7.2: Time period, vertical range, vertical resolution, references and other comments for CO measurements.

Instrument	Time period	Vertical range	Vertical resolution	References	Additional comments
SMR V2	Nov 01 –	~17 – 110 km	3 – 4 km	*Dupuy et al.*, 2004	only period Oct 03 – Oct 04 used in this report.
MIPAS MIPAS(1) V10 MIPAS(2) V220	 Mar 02 – Mar 04 Jan 05 – Apr 12	6 – 70km (cloud top – 70 km)	3.5 – 8 km	*Funke et al.*, 2009	change in measurement mode in 2005, CO data only available from Jul 2002 onward
ACE-FTS V2.2	Mar 04 –	5 – 105 km	3 – 4 km	*Clerbaux et al.*, 2008 *Hegglin et al.*, 2008	
Aura-MLS V3	Aug 04 –	215 – 0.0046 hPa	~ 4 km (UTLS) 3 km (above)	*Pumphrey et al.*, 2007 *Livesey et al.*, 2008	

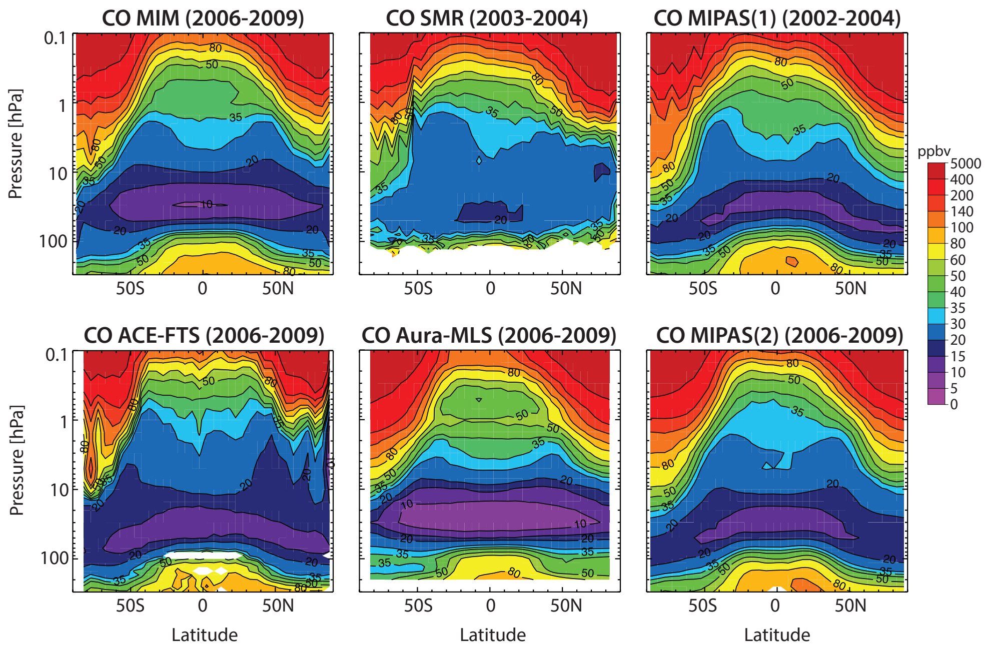


Figure 4.7.1a: Cross sections of annual zonal mean CO for 2006-2009. *Cross sections of CO are shown for the MIM, SMR, MIPAS(1), ACE-FTS, MIPAS(2), and Aura-MLS. Note that SMR is averaged over the period October 2003-October 2004 and MIPAS(1) over July 2002-March 2004. These datasets are not included in the MIM.*

±20%. Overall, MIPAS seems more similar to ACE-FTS than Aura-MLS. The largest relative differences are found in the SMR climatology, with values indicating a positive departure from the MIM. The values reach +100% in the tropical LS and Northern Hemisphere polar LS.

The differences of SMR and MIPAS(1) from the MIM in the MS and USLM are largely consistent with each other. MIPAS(1) (July 2002–March 2004) and SMR (October 2003-October 2004) were averaged over a similar time period, including the stratospheric warming event in January 2004 that led to the well-documented strong downward transport of mesospheric air at NH high latitudes. This difference indicates that comparisons using different time periods are affected by natural variability (at least in the USLM), and that part of the differences from the MIM can be attributed to the temporal sampling biases. However, a direct comparison between the two instruments for particular months still shows differences of over 40% in the global mean LS and NH USLM (with smaller differences around 10-15% in the tropical US and SH USLM;

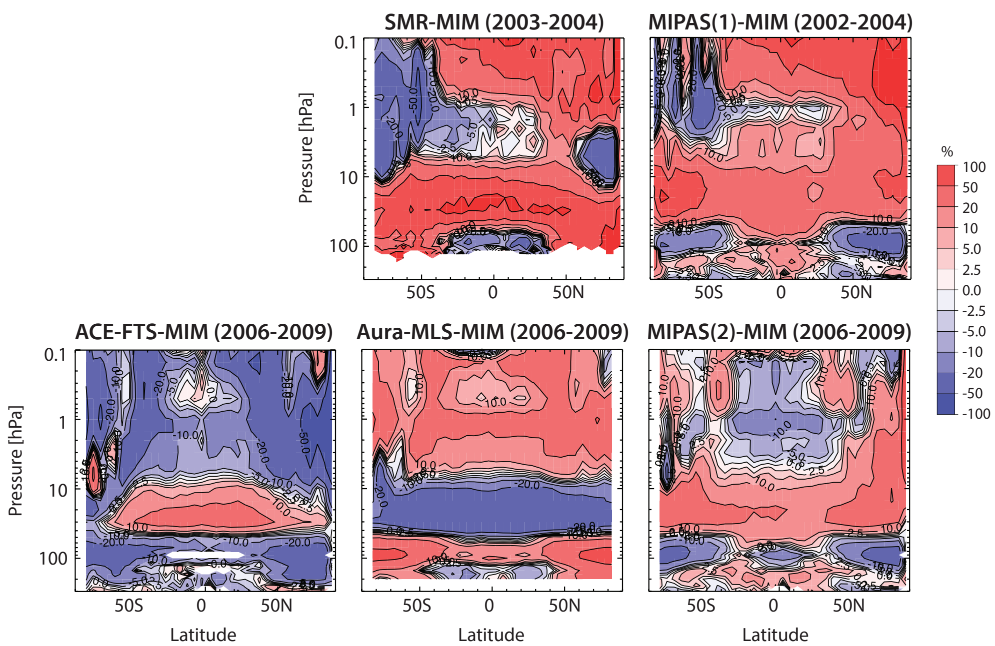


Figure 4.7.1b: Cross sections of annual zonal mean CO differences for 2006-2009. *Cross sections of CO relative differences between the individual instruments (MIPAS(1), ACE-FTS, Aura-MLS, MIPAS(2) and SMR) and the MIM are shown. Note, SMR (October 2003-October 2004) and MIPAS(1) (July 2002-March 2004) data are not included in the MIM.*

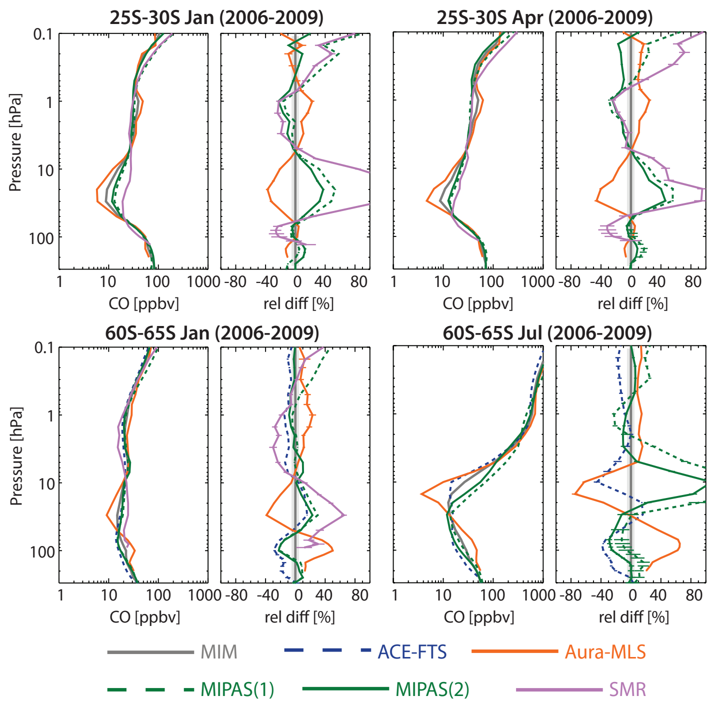


Figure 4.7.2a: Vertical profiles of monthly zonal mean CO for 2006-2009. *Vertical CO profiles for January and April 25°S-30°S (upper panels), and for January and July 60°S-65°S (lower panels) are shown together with instrument differences from the MIM. Note, SMR and MIPAS(1) measurements are taken in 2003-2004 and 2002-2004, respectively, and SMR does not provide data during July 2004 at SH high latitudes.*

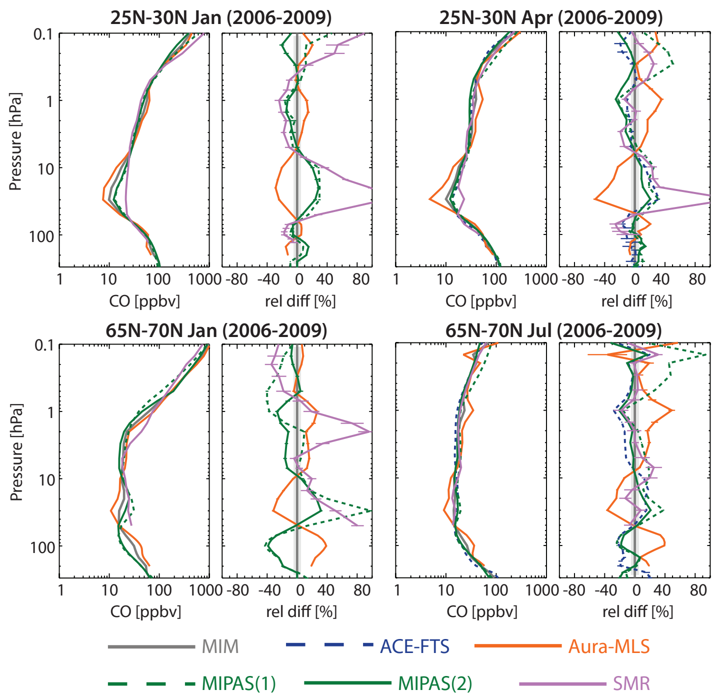


Figure 4.7.2b: Vertical profiles of monthly zonal mean CO for 2006-2009. *Vertical CO profiles for January and April 25°N-30°N (upper panels), and for January and July 65°N-70°N (lower panels) are shown together with instrument differences from the MIM. Note, SMR and MIPAS(1) measurements are taken in 2003-2004 and 2002-2004, respectively.*

see **Figures A4.7.2a** and **A4.7.2b** in *Appendix A4*). Also, as is shown in the following evaluations, the differences between SMR and MIPAS(1) are mostly larger than the differences between MIPAS(1) and MIPAS(2) (even though they sample different years), indicating systematic differences between the SMR and MIPAS datasets (with these findings also being reflected in the summary plot **Figure 4.7.8**).

4.7.3 CO evaluations: Vertical and meridional profiles

The vertical profiles shown in **Figures 4.7.2a** and **4.7.2b** reveal further details in the structure in the differences of the monthly mean cross sections (see also **Figures A4.7.1** and **A4.7.2** in *Appendix A4).*

In the SH (**Figure 4.7.2a**), MIPAS and Aura-MLS agree well in the tropical UTLS, however, their values diverge in the MS, and are closer to each other again in the US and LM. Where ACE-FTS is available for comparison, it mostly follows the shape of the MIPAS profiles, indicating that MIPAS and ACE-FTS produce the most consistent results. In the extra-tropics, Aura-MLS CO profiles show large deviations from the MIM in the UTLS and MS, but relatively good agreement in the USLM. The SMR profiles seem to agree in the shape with those of the MIPAS and ACE-FTS climatologies, but show significantly larger values in the LS and MS (between about 50 hPa and 5 hPa). In the USLM, SMR CO is slightly larger than MIPAS(1) CO, and both are larger than the other instruments, indicative of the sampling bias mentioned above.

The above findings are similar in the NH, for the most part (**Figure 4.7.2b**). ACE-FTS, where available, agrees well with the MIPAS profiles. Aura-MLS exhibits a wave-like structure in its differences to the MIM that is mostly opposite of the structure found in the differences between MIPAS (or ACE-FTS) and the MIM.

CO meridional mean profiles for April and October at different pressure levels are shown in **Figure 4.7.3**. The figure emphasises the very large relative differences of the measurements from the MIM, which are on average about ±30-40%. The best agreement is found on the 5 hPa level, where apart from the regions with strong downwelling, relative differences from the MIM are within ±20%.

4.7.4 CO evaluations: Latitude-time evolution

Figure 4.7.4a and **4.7.4b** show the climatological latitude-time evolution of CO at 1 and 100 hPa, respectively. Note, as indicated in the figure caption, SMR and MIPAS(1) are averaged over a different time period than the other instruments, and therefore not included in the MIM (average over 2006-2009). ACE-FTS also shows rather noisy fields due to its limited sampling, however the available information is helpful for validating the other instruments. SMR is not included in the 100 hPa evaluation, since this level is at the lower boundary of its measurement range.

At 1 hPa, SMR, MIPAS, and Aura-MLS agree on the downwelling within the polar vortex reasonably well, both in time and amplitude. However, outside of the polar vortex where minimum CO values occur, Aura-MLS shows much higher average values than the other instruments. While the latitude-time evolution of ACE-FTS CO is poorly defined, especially in the tropical region where its sampling density is lowest, it seems to indicate as well that Aura-MLS shows too high CO. This finding is consistent with *Pumphrey et al.* [2007] who found a positive bias against correlative measurements of 25-50% in the USLM.

At 100 hPa, MIPAS(1) and MIPAS(2) exhibit mostly the same structure, however, with MIPAS(1) having slightly higher tropical values, which may be due to a trend in UT CO over the first decade starting in 2000 [*Worden et al.*, 2013] or simply due to interannual variability. The rather limited information obtained from ACE-FTS supports the MIPAS findings in terms of both magnitude and structure. Aura-MLS on the other hand exhibits much higher CO mixing ratios, a somewhat different seasonality, and also much smaller gradients across the subtropical region towards higher latitudes than the other two instruments.

4.7.5 CO evaluations: Seasonal cycles

The seasonal cycle in zonal mean CO is shown in **Figure 4.7.5** for different levels and latitude bands. In the

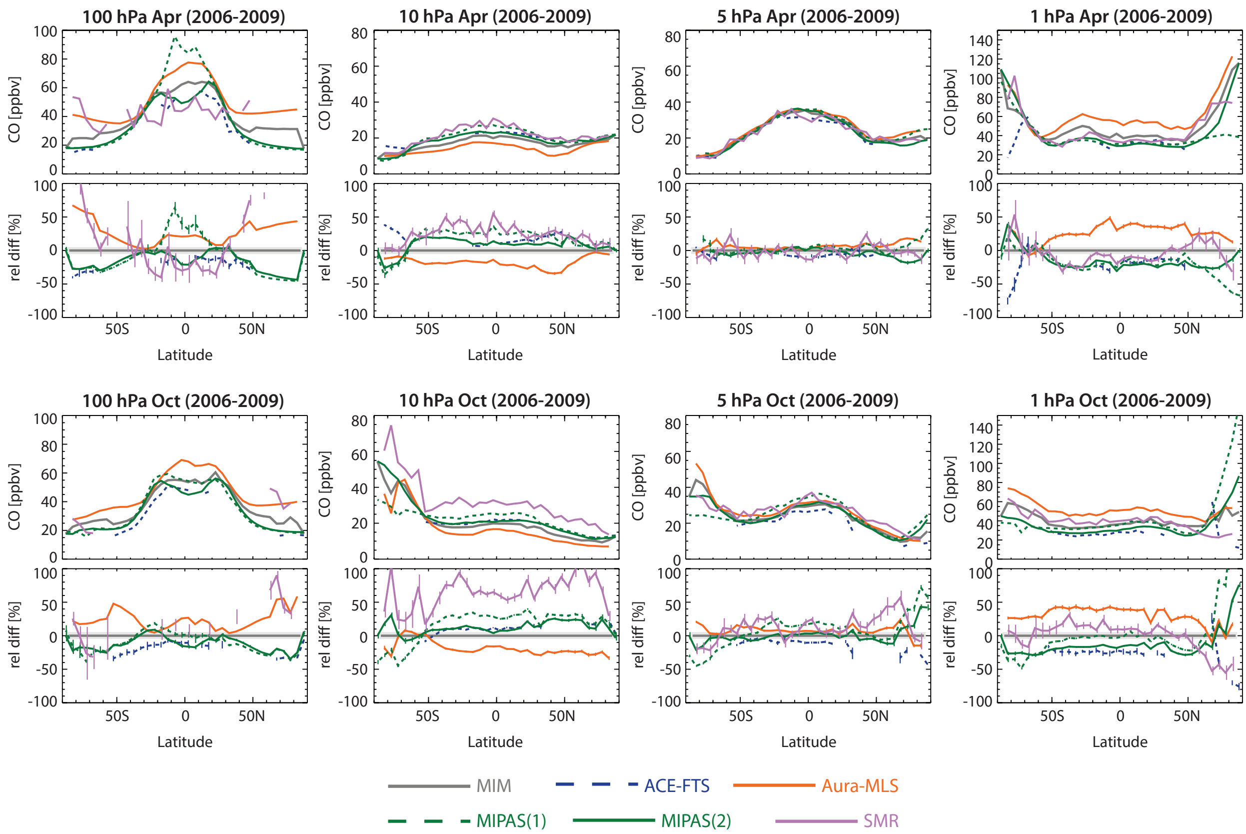


Figure 4.7.3: Meridional profiles of monthly zonal mean CO for 2006-2009. *Meridional CO profiles at 100, 10, 5, and 1 hPa for April (upper row) and October (lower row) averaged over 2006-2009. Differences between the individual instruments (SMR, MIPAS(1), MIPAS(2), ACE-FTS, and Aura-MLS) and the MIM profiles are shown in the lower panel. Note, SMR and MIPAS(1) measurements are taken in 2003-2004 and 2002-2004, respectively.*

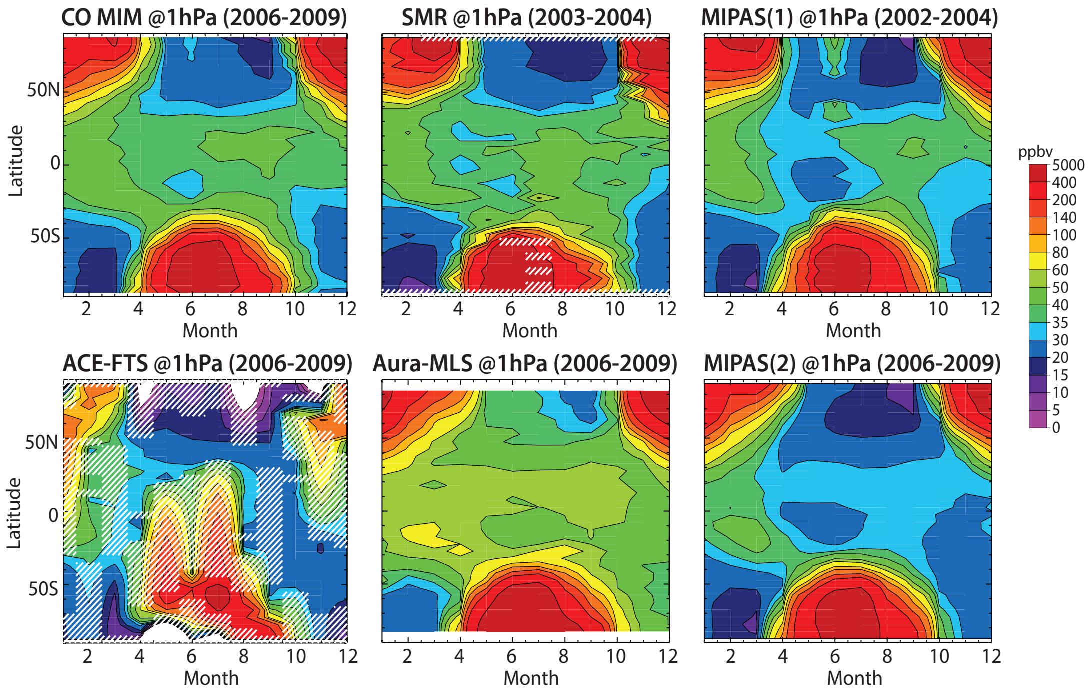


***Figure 4.7.4a: Latitude–time evolution of CO at 1hPa.** The latitude-time evolution of CO at 1 hPa is shown for the MIM (2006-2009 average) in the upper leftmost panel and the instruments SMR, MIPAS(1), ACE-FTS, Aura-MLS and MIPAS(2). SMR and the ACE-FTS show interpolated fields, with hatched regions indicating where no measurements are available. Note that SMR and MIPAS(1) are averaged over a different time period as indicated in the Figure title, and therefore are not included in the MIM.*

tropics, a semi-annual cycle with small amplitude is seen at 200 hPa. Here, MIPAS, ACE-FTS and Aura-MLS show very similar cycle phases and amplitudes and all agree within ±6%. MIPAS(1) and MIPAS(2) show mean values consistent with each other, however lie on the high side of the MIM, while Aura-MLS lies about in the middle, and ACE-FTS below the MIM. At 100 hPa, the inter-instrument differences become larger (more than ±15%). MIPAS(2) and ACE-FTS agree very well, while Aura-MLS and MIPAS(1) are on the high side of the MIM, and also show a somewhat larger amplitude than MIPAS(2) and the ACE-FTS. SMR is at the lower boundary of its measurement range and shows a seasonal cycle that is opposite of those of the other instruments. Note, that while the measurements at this level are not recommended to be used, similar problems are also seen for SMR at 70 and 50 hPa (not shown), which stems from a decreasing sensitivity at pressures of about 50 hPa and larger.

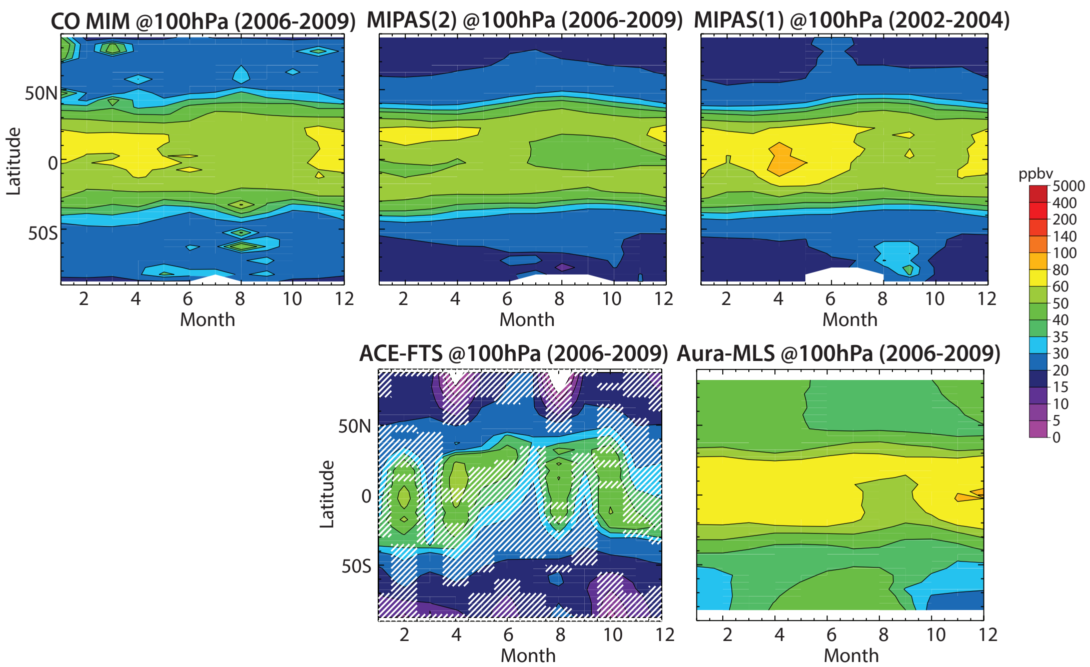


***Figure 4.7.4b:** Same as Figure 4.7.4a, but for 100 hPa.*

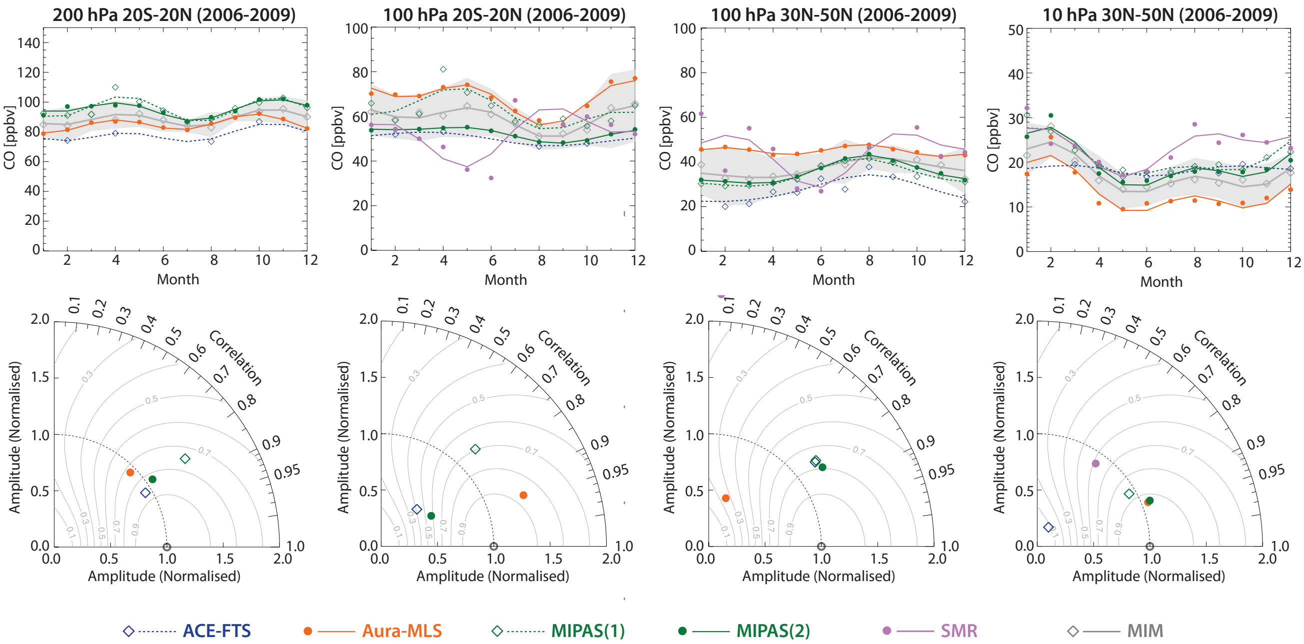


Figure 4.7.5: Seasonal cycle of CO. Seasonal cycles (upper panels) and corresponding Taylor diagrams (lower panels) of monthly zonal mean CO are shown for the tropics (20°S-20°N) at 200 and 100 hPa (two left columns), and for the extra-tropics (30°N-50°N) at 100 and 10 hPa (two right columns).

In the extra-tropics, MIPAS and ACE-FTS agree well on the phase and amplitude of the seasonal cycle at 100 hPa, although ACE-FTS shows slighly smaller mean values. SMR again shows the wrong seasonal cycle, and Aura-MLS is on the high side of the MIM with too small an amplitude. At 10 hPa, the seasonal cycles of MIPAS and Aura-MLS agree well in terms of phase and amplitude, however, the mean values of Aura-MLS here are lower than those of the MIM. SMR exhibits a more similar evolution of the seasonal cycle but with higher values in the second half of the year (based on data for 2004 only). Note that MIPAS(1), which covers approximately the same time period as SMR, does exhibit a seasonal cycle that is closer to MIPAS(2) and Aura-MLS, indicating that sampling may not be the only issue of SMR. The seasonal cycle of ACE-FTS is too flat, potentially attributable to its limited sampling.

4.7.6 CO evaluations: Interannual variability

Another important aspect of instrument performance, apart from the representation of the climatological mean structure, is the instruments' capability to demonstrate interannual variability. **Figure 4.7.6** shows anomalies for the different instruments in different atmospheric regions and at different pressure levels for 2005-2010. Note that SMR is not included in this evaluation since there is only one year of data, which is too short for deseasonalizing the data.

The anomalies in **Figure 4.7.6** reveal that in the global MS and tropical UT, MIPAS and Aura-MLS agree very well. This is a somewhat surprising result given the large discrepancies between the annual zonal mean structure of these two instruments. Furthermore, the two instruments seem to exhibit slightly different trends; MIPAS lies above (below) Aura-MLS in 2005 (2010). While ACE-FTS measurements are much sparser, it also follows the MIM and its overall tendencies quite well (at least in the extra-tropics). The interannual variability relative to the absolute amount of CO is relatively small in the tropics at both levels (~ ±8%) where variability is mostly determined by variability in the source processes of tropospheric CO, but large at 10 hPa in the extra-tropics (±30%) where high CO mixing ratios are dominated by the photo-dissociation of CO_2 in the mesosphere and downward transport within the polar vortex.

4.7.7 Summary and conclusions: CO

CO climatologies from four limb-sounders (SMR, MIPAS, ACE-FTS and Aura-MLS) have been compared within the SPARC Data Initiative. MIPAS data before/after 2005 have been evaluated separately (using MIPAS(1) and MIPAS(2)). Note that SMR currently provides CO data only over a short period of time and with limited temporal sampling. Overall findings on the systematic uncertainty in our knowledge of the CO mean state and important characteristics of the individual datasets are presented in the following summary including two synopsis plots as discussed in the previous trace gas sections and detailed in *Section 3.3.5*.

Atmospheric mean state

The CO climatologies obtained from the four satellite instruments show large relative differences from the MIM, and do not agree on some key features in the annual zonal mean distribution. The biases derived from the annual mean are somewhat lower in the monthly zonal mean evaluations, and can be further reduced when periods are

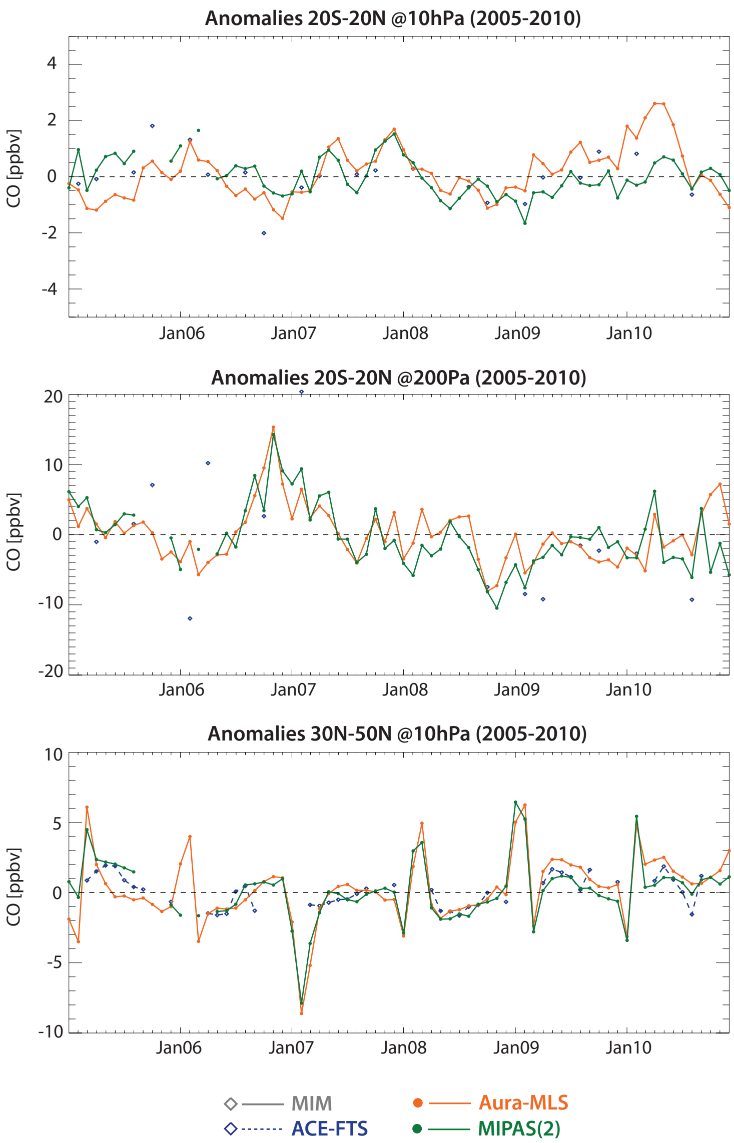


Figure 4.7.6: Time series of deseasonalised CO anomalies for 2005-2010. *Deseasonalised CO anomalies are shown for 20°S-20°N at 10 hPa (upper panel) and 200 hPa (middle panel), and for 30°N-50°N at 10 hPa (lower panel). Note that MIPAS here consists of MIPAS(2) data only.*

chosen during which instruments overlap (*e.g.*, SMR and MIPAS(1) in 2003 and 2004). It is notable that despite the disagreement in the annual and monthly zonal means, the instruments capture the pronounced seasonal features and interannual variability in the CO distribution quite well.

The uncertainty in our knowledge of the atmospheric CO annual mean state as derived from the Aura-MLS, ACE-FTS, MIPAS(1), MIPAS(2), and SMR satellite instruments and as averaged over 2002-2009 is smallest in the global UT with a 1σ multi-instrument spread of less than 4% (see **Figure 4.7.7**). Good knowledge is obtained in the tropical MS around 10 hPa and USLM around 5 hPa, where the uncertainty is about 10%. The uncertainty is largest in the extra-tropical LS around 100 hPa and throughout the stratosphere at NH high-latitudes (with 1σ values of more than 50%). Rather large uncertainty is also found in the LSMS between approximately 50-20 hPa, which may be explained by the large dynamic range of CO mixing ratios in the atmosphere that can cause retrieval problems, with the instruments having to detect relatively small CO mixing ratios in this region below very high values in the US and mesosphere. Part of the uncertainty in the USLM is due to strong interannual variability in this region that can lead to substantial sampling biases.

Performance by region

As seen in **Figure 4.7.8.**, in the USLM (0.1-5 hPa), ACE-FTS, MIPAS(2), and Aura-MLS show good agreement in the tropics, with relatively small MADs, indicating well defined differences. In the extra-tropics, ACE-FTS shows larger negative differences from the MIM, which is in part potentially attributable to a sampling bias due to the pronounced vertical and horizontal gradients in CO mixing ratios that are larger than for other trace gases in this region. The positive deviations from the MIM seen in both the tropics and extra-tropics in Aura-MLS data are consistent with *Pumphrey et al.* [2007] who found a positive bias in Aura-MLS against correlative measurements of 25-50% in the USLM. SMR and MIPAS(1) show much larger positive deviations from the MIM than the other three instruments in the LM, but similar values in the US. The differences are partially attributable to sampling during a different time period; a period during which the downwelling of CO-rich air from the mesosphere was stronger than usual. However, the MADs are very large with values of up to ±30%, indicating that the deviations from the MIM are not well defined within the region.

- In the MS (5-30 hPa), ACE-FTS agrees well with the two MIPAS datasets in the tropics, while Aura-MLS is lower (by 30%) and SMR higher than the MIM (>50%). In the extra-tropics, ACE-FTS lies in between Aura-MLS (on the negative side of the MIM) and MIPAS(2) (on the positive side of the MIM) and SMR is closer to MIPAS(1) and MIPAS(2).
- In the LS (30-100 hPa), the inter-instrument differences are around ±18% in the tropics and ±40% in the extra-tropics. Both SMR and Aura-MLS exhibit large positive deviations from the MIM in the extra-tropics, while ACE-FTS and MIPAS agree very well.
- In the UT (100-300 hPa), the agreement is best, especially in the tropics with all instruments lying within ±5%. In the extra-tropics, where natural variability is larger, ACE-FTS (Aura-MLS) is on the low (high) side of the MIM. SMR shows a high bias at 100 hPa, while its measurements do not reach below 100 hPa.

Instrument-specific conclusions

The **SMR** instrument provides currently only one year of CO data. SMR performs well in the tropical USLM. However, throughout the extra-tropical UTLS and MS it exhibits values that are mostly too high. Here, it shows the largest

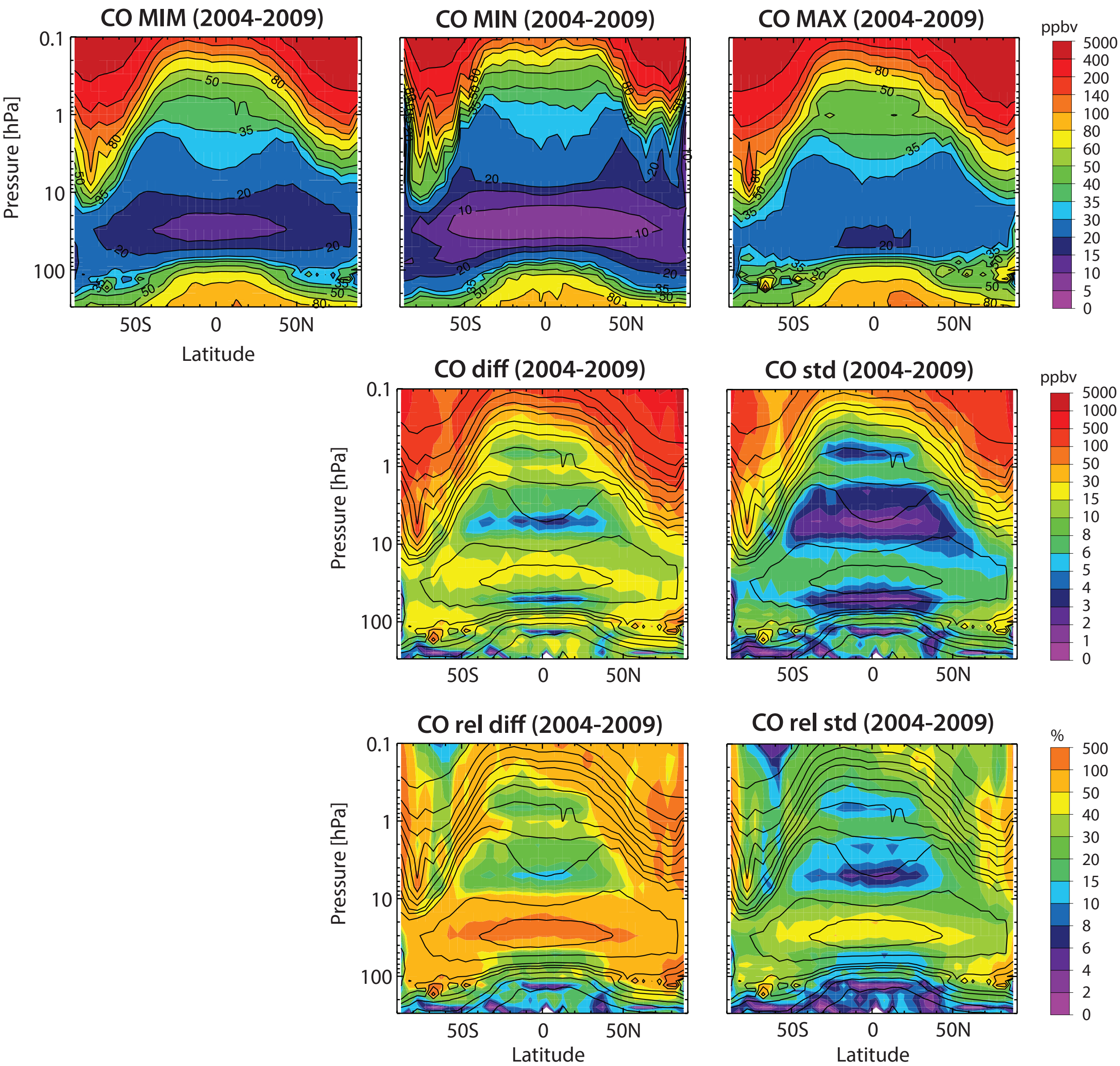


Figure 4.7.7: Summary of CO annual zonal mean state for 2002-2009. *Shown are the annual zonal mean cross-sections of the MIM, minimum (MIN), and maximum (MAX) CO values (upper row), the maximum differences over all instruments (MAX-MIN) and the standard deviation over all instruments (middle row), and the relative differences and relative standard deviations with respect to the MIM (lower row). Black contours in lower panels repeat the MIM distribution. Instruments considered are ACE-FTS, Aura-MLS, MIPAS(1), MIPAS(2),and SMR.*

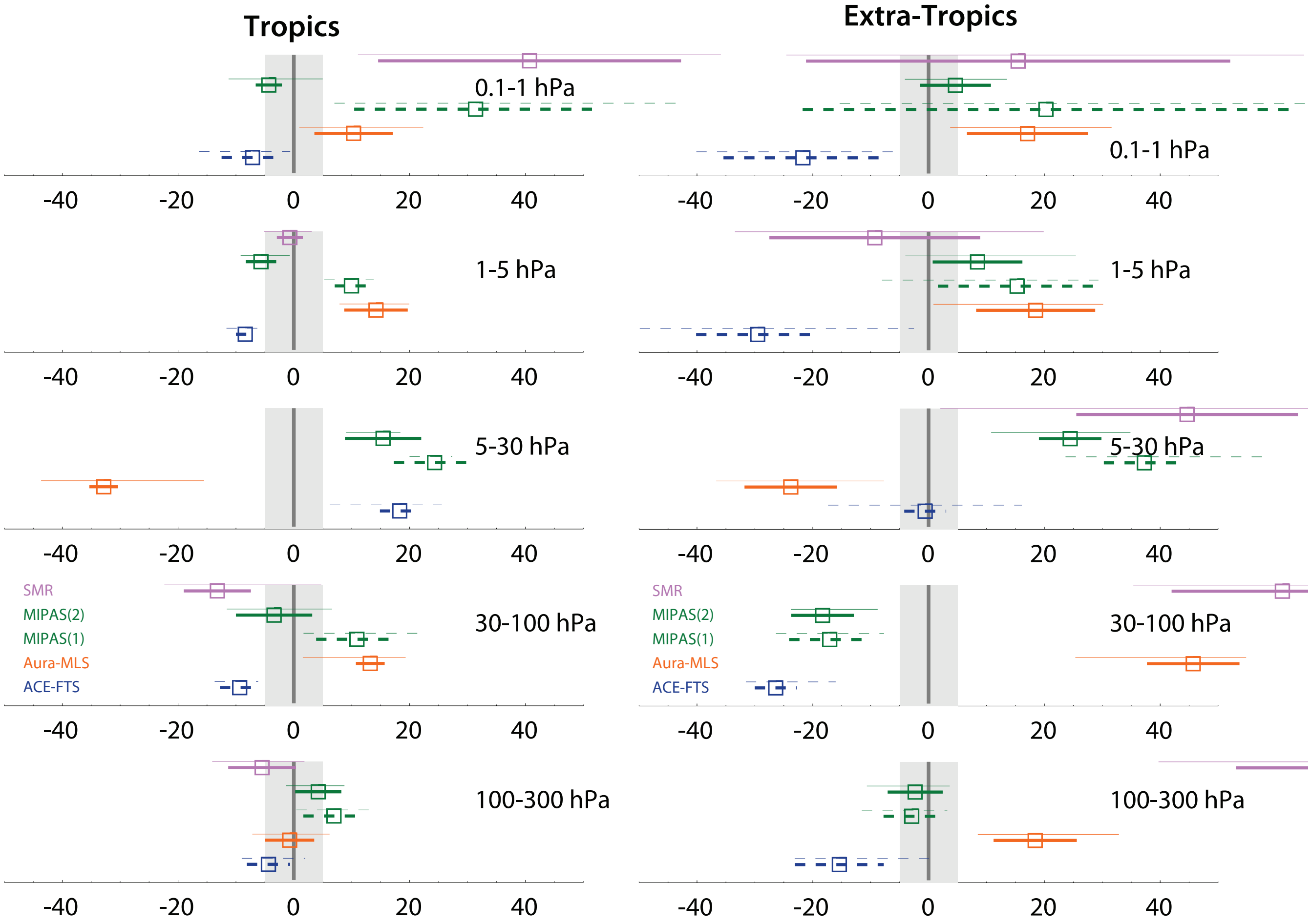


Figure 4.7.8: Inter-instrument differences in CO calculated for the tropics (left) (20°S–20°N) and extra-tropics (right) (40°S–80°S and 40°N–80°N), and for altitude regions from the UT up to the LM. *Shown are the median (squares), median absolute deviations (MAD, thick lines), and the mean ±1σ ranges (thin lines) of the relative differences between each individual instrument and the MIM calculated over a given latitude and altitude region. The reference period is 2002-2009. Note, SMR and MIPAS(1) data are not included in the MIM calculation. The median difference of SMR in the tropics between 5 and 30 hPa is outside the depicted range (at +80%).*

relative differences, up to ±50% from the MIM. Towards the lower boundary of its measurement range (between 100-70 hPa) in both the tropics and the extra-tropics, SMR exhibits seasonal cycles in CO that look different from the seasonal cycles of the other instruments. Note that due to the quickly decreasing measurement response below altitudes around 70 hPa, which is responsible for the poor performance of SMR in the presented evaluations in the UTLS, the SMR SPARC Data Initiative climatologies are now updated to exclude data below 70 hPa.

ACE-FTS agrees best with MIPAS on both structure and mean value in the CO distribution, especially in the tropics. In the extra-tropics, ACE-FTS shows consistently lower values than the MIM. However, a larger sampling bias over regions with larger CO gradients may be the reason for the discrepancies found due to the climatological validation approach used in these evaluations.

Both MIPAS versions are consistent for the most part, although **MIPAS(1)** shows consistently higher values than **MIPAS(2)**. The discrepancies are larger in the USLM than in the lower atmosphere, which is at least partially explained by the different time periods spanned by the measurements. The USLM exhibits particularly large interannual variability and differences in temporal and spatial sampling can lead to a large sampling error. MIPAS nominal CO data have been cross-validated with ACE-FTS observations [*Clerbaux et al.*, 2008; *Hoffmann et al.*, 2011]. Differences between the two instruments are typically within ±25%. This result is consistent with, although more conservative than, the differences found in our climatological validation approach, at least in the tropics where natural variability is small. MIPAS also agrees very well (within 10%) with ground-based microwave observations [*Forkman et al.*, 2012].

The **Aura-MLS** CO climatology exhibits an apparently unphysical behaviour in the LS, where CO isopleths are not sloping downwards towards higher latitudes as found in MIPAS and ACE-FTS, and as is expected for longer-lived tracers whose distribution is controlled by the Brewer-Dobson circulation. Aura-MLS shows lower CO values than the other instruments in the 10-30 hPa region, and higher values above and below that region. The mean climatology biases are also reflected in the seasonal cycles, which exhibit too low (high) values at 10 (100) hPa. It is notable that despite the structural problems in the CO mean distribution, Aura-MLS reproduces interannual variability very well. Also, it performs well in the tropical UT where the scientific interest is high.

4.7.8 Recommendations: CO

While the instruments show rather large discrepancies in zonal monthly and annual mean evaluations, they agree very well on interannual variability in both the tropical UTLS and MS. It is, hence, recommended that diagnostics be used for model-measurement comparisons that focus on temporal anomalies from the mean state in order to eliminate inter-instrument biases in the CO mean distribution.

4.8 Hydrogen fluoride – HF

HF is primarily produced through the photodissociation of anthropogenic CFCs and hydrochlorofluorocarbons (HCFCs). Once produced, HF is the dominant reservoir of free fluorine atoms and has a very long lifetime, allowing it to accumulate in the stratosphere. The removal of HF happens through downward transport into the troposphere and subsequent rainout, or by upward transport to the mesosphere where it is destroyed by photolysis. Due to its very long lifetime, HF can be used as a tracer for diagnosing transport by the Brewer-Dobson circulation, and for separating dynamics and chemistry in polar regions. Since HF is a direct product of CFCs and HCFCs, it is considered a useful tracer for monitoring anthropogenic changes of the stratospheric composition.

4.8.1 Availability of HF measurements

Measurements of HF are available from 1991 to 2005 from HALOE, and from 2004 onward from ACE-FTS. The two datasets overlap for 2004-2005. **Tables 4.8.1** and **4.8.2** compile information on the availability of HF measurements, including time period, altitude range, vertical resolution, and references relevant for the data product used in this report.

4.8.2 HF evaluations: Zonal mean cross sections, vertical and meridional profiles

Annual zonal mean cross sections for the time period 2004-2005 are analysed to investigate mean differences between the two datasets. Additionally, vertical and meridional profiles are evaluated. Note that although only two datasets are available, the comparison of both datasets to their MIM (and not a direct comparison) will be used to be consistent with other parts of the report.

Figure 4.8.1 shows the annual zonal mean HF climatologies for 2004-2005 for HALOE and ACE-FTS. HF increases with altitude due to the combination of its stratospheric source and very long lifetime. The HF isopleths slope downwards towards higher latitudes as a result of tropical upwelling and extra-tropical downwelling by the Brewer-Dobson circulation. The annual mean HF distributions observed by HALOE and ACE-FTS show the same overall shape. HALOE isopleths display some kinks at 50°S-60°S and 50°N-60°N that are, at least partially, related to the HALOE sampling pattern. The change of the latitudinal coverage from month to month can cause such discontinuities. Note that HALOE coverage was reduced after 2002. Similar kinks can be observed in the ACE-FTS isopleths at around 80°S.

The relative differences of HALOE and ACE-FTS annual means from the MIM are displayed in **Figure 4.8.2**. Above 50 hPa (10 hPa at the equator), HALOE detects less HF than ACE-FTS, with differences from the MIM of up to ±5%, but up to ±10% in some areas. The only exception to this good agreement is in the SH high latitudes where differences from the MIM can be as high as ±20% (differences between the instrument climatologies can become as large as 40%). The fact that HALOE observes less HF in the MS/US is consistent with existing comparisons with other instruments such as Atmospheric and Oceanic Sensors (ATMOS) [*Russell et al.*, 1996a]. The UTLS and the tropical MS are the only regions where ACE-FTS measures less HF than HALOE, with differences from the MIM mostly below ±10%, although exceeding ±50% in some parts of the UT. In each individual

***Table 4.8.1:** Available HF measurement records from limb-sounding satellite instruments between 1978 and 2010.* *The red filling of the grid boxes indicates the temporal and vertical coverage of the respective instruments.*

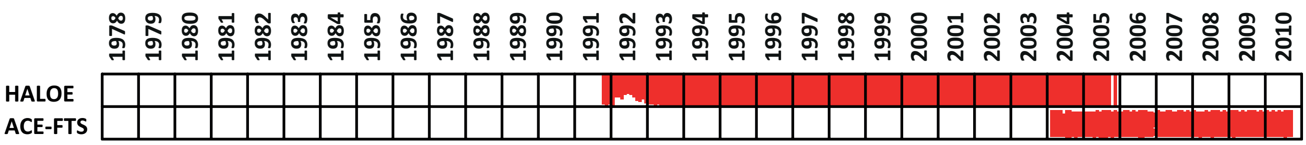


***Table 4.8.2:** Time period, vertical range, vertical resolution, references and other comments for HF measurements.*

Instrument	Time period	Vertical range	Vertical resolution	References	Additional comments
HALOE V19	Oct 91 – Nov 05	250 – 0.1 hPa 12 – 65 km	3.5 km	*Grooß and Russell*, 2005	
ACE-FTS V2.2	Mar 04 –	250 – 0.5 hPa 12 – 55 km	3 – 4 km	*Mahieu et al.*, 2008	

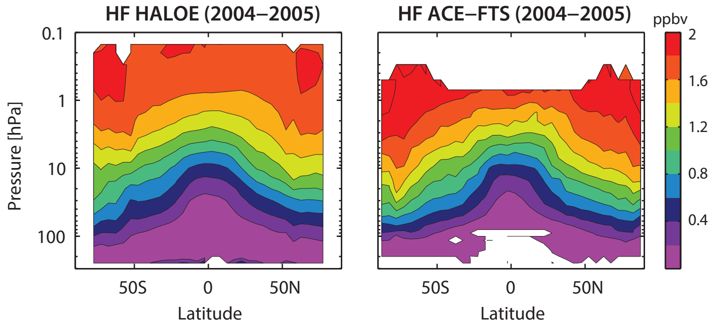


__Figure 4.8.1: Cross sections of annual zonal mean HF for 2004-2005.__ HF cross sections are shown for HALOE and ACE-FTS.

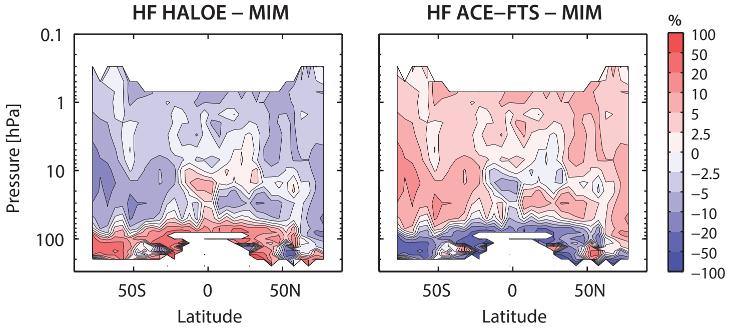


__Figure 4.8.2: Cross sections of annual zonal mean HF differences for 2004-2005.__ HF differences for HALOE and ACE-FTS with respect to the MIM are shown. Note that, since the MIM consists of only two instruments, any issue with one dataset will fully be reflected by the difference of the other dataset.

latitude band, the two instruments measure during different months, which impacts the representativeness of the annual mean differences. In particular, the high latitude climatologies will be influenced by the different sampling of the vortex. However, the annual mean differences give a picture that is generally consistent with the monthly mean differences presented below (see **Figures 4.8.3** and **4.8.4** for profile comparisons) and in *Appendix A4* (see **Figures A4.8.1 – A4.8.4** for monthly mean cross sections).

Monthly mean vertical HF profiles are shown in **Figure 4.8.3** together with their differences from the MIM for SH high latitudes in March, and tropical latitudes in August. Above 50 hPa, the two instruments show good agreement with differences from the MIM of up to ±10%, while below 50 hPa differences increase up to ±20% in the high latitudes, and up to ±50% in the tropics. ACE-FTS is smaller in the UTLS and larger in the MS/US, consistent with the annual mean cross sections. Profiles in the polar regions during their respective summers show very good agreement, with differences above the tropopause mostly below ±5%, with HALOE (ACE-FTS) on the low (high) side (see **Figure A4.8.5** in *Appendix A4*). Note that many profiles are in the polar region averages, and fewer in the tropical regions.

In **Figure 4.8.4,** meridional HF profiles and their differences from the MIM are shown at 1, 10, 70 and 100 hPa. At the upper stratospheric levels, the relative differences show a meridional gradient with largest values in the tropics. While differences in the extra-tropics are mostly below ±5%, they reach values of up to ±10% (±20%) in the tropics at 1 hPa (10 hPa), which might be related to the sample size of the data in the tropical averages. At the lower stratospheric levels (70 and 100 hPa), the relative differences

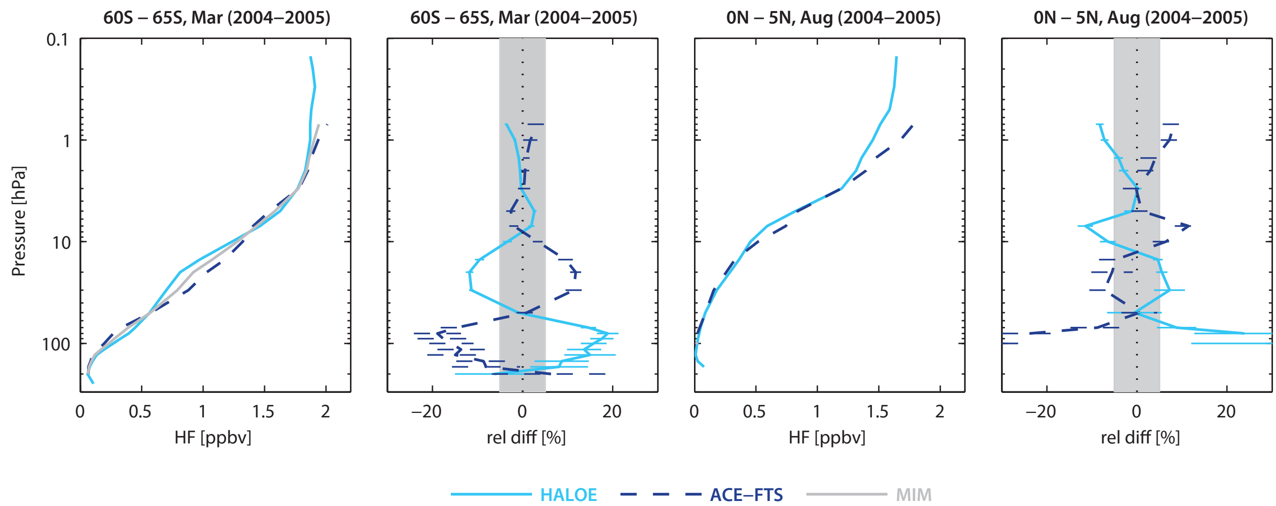


__Figure 4.8.3: Vertical profiles of monthly zonal mean HF for 2004-2005.__ Zonal mean HF profiles for 60°S-65°S, March (left panels) and 0°N-5°N, August (right panels) are shown together with their differences from the MIM. The grey shading indicates the ±5% difference range. Bars indicate the uncertainties in the relative differences.

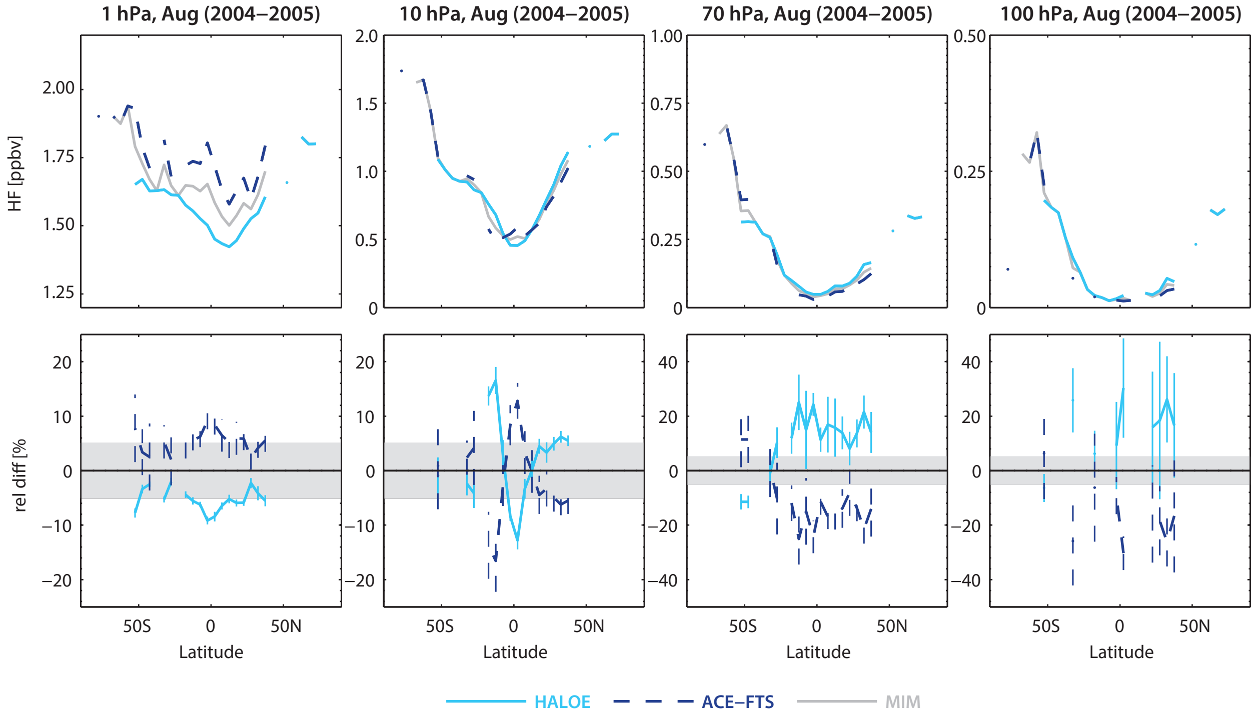


__Figure 4.8.4: Meridional profiles of monthly zonal mean HF for 2004-2005.__ HF profiles at 1, 10, 70 and 100 hPa for August are shown in the upper row. Differences of the individual instruments (HALOE and ACE-FTS) to the MIM are shown in the lower row. The grey shading indicates the ±5% difference range. Bars indicate the uncertainties in the relative differences.

from the MIM can be larger (±20%), although they show no strong meridional gradient. Overall, the monthly mean comparisons show slightly larger (smaller) differences between the instruments for the tropics (polar regions) than the annual mean comparison.

The HF time series from HALOE and ACE-FTS overlap for only two years, which makes a quantitative comparison of the seasonal cycle and interannual variability difficult. **Figure 4.8.5** shows the time series of monthly mean values from 1994 to 2010 for SH high latitudes at 1 hPa, and SH (NH) mid-latitudes at 10 hPa (100 hPa). Different time scales of variability dominate these three case studies. In the US at SH high latitudes, both time series show increasing values over their respective lifetimes, indicating a positive trend as the dominant signal. A seasonal cycle with increasing HF abundance over the summer is found in both the HALOE and ACE-FTS time series. In the mid-latitude region at 10 hPa, the signal of interannual variability dominates both time series, with stronger variations in the later time period ACE-FTS record. In the mid-latitude LS, the seasonal cycle is the strongest signal and both time series agree on its overall shape, with maximum values in the winter. A more detailed comparison of the overlap period, however, shows stronger month-to-month variations in ACE-FTS and therefore considerable disagreement between the two LS time series for individual months.

4.8.3 Summary and conclusions: HF

A comparison of two HF profile climatologies (HALOE, ACE-FTS) has been carried out. Overall findings on the systematic uncertainty in our knowledge of the mean state of atmospheric HF, and important characteristics of the individual datasets are presented in the following summary, including two synopsis plots. The first summary plot (**Figure 4.8.6**) provides information on the mean state and its uncertainty derived from the spread between the datasets. The second summary plot (**Figure 4.8.7**) shows specific inter-instrument differences in the form of deviations of the two instrument climatologies from their MIM climatology. For each instrument and selected region, the deviation from the MIM is given as the median (mean) difference over all grid points in this region. Additionally, for each instrument the spread of the differences over all grid points in this region is presented. Note that both pieces of information (average deviation and spread) are important for a meaningful assessment of inter-instrument differences. A detailed description of the summary plot evaluations can be found in *Section 3.3.5*.

Atmospheric mean state

The uncertainty in our knowledge of the annual mean state of atmospheric HF as derived from two satellite datasets is smallest above 100 hPa, with a 1σ multi-instrument spread in this region of less than ±10% (less than ±5% above 10 hPa (**Figure 4.8.6**)). One exception is in the SH high latitudes where the two annual mean climatologies give a spread of ±15% in the MS. The larger disagreement in the SH high latitudes is mainly caused by the fact that the annual averages are based on different months, and therefore the annual mean datasets for both instruments are impacted by sampling biases. The evaluation of individual monthly mean profiles and

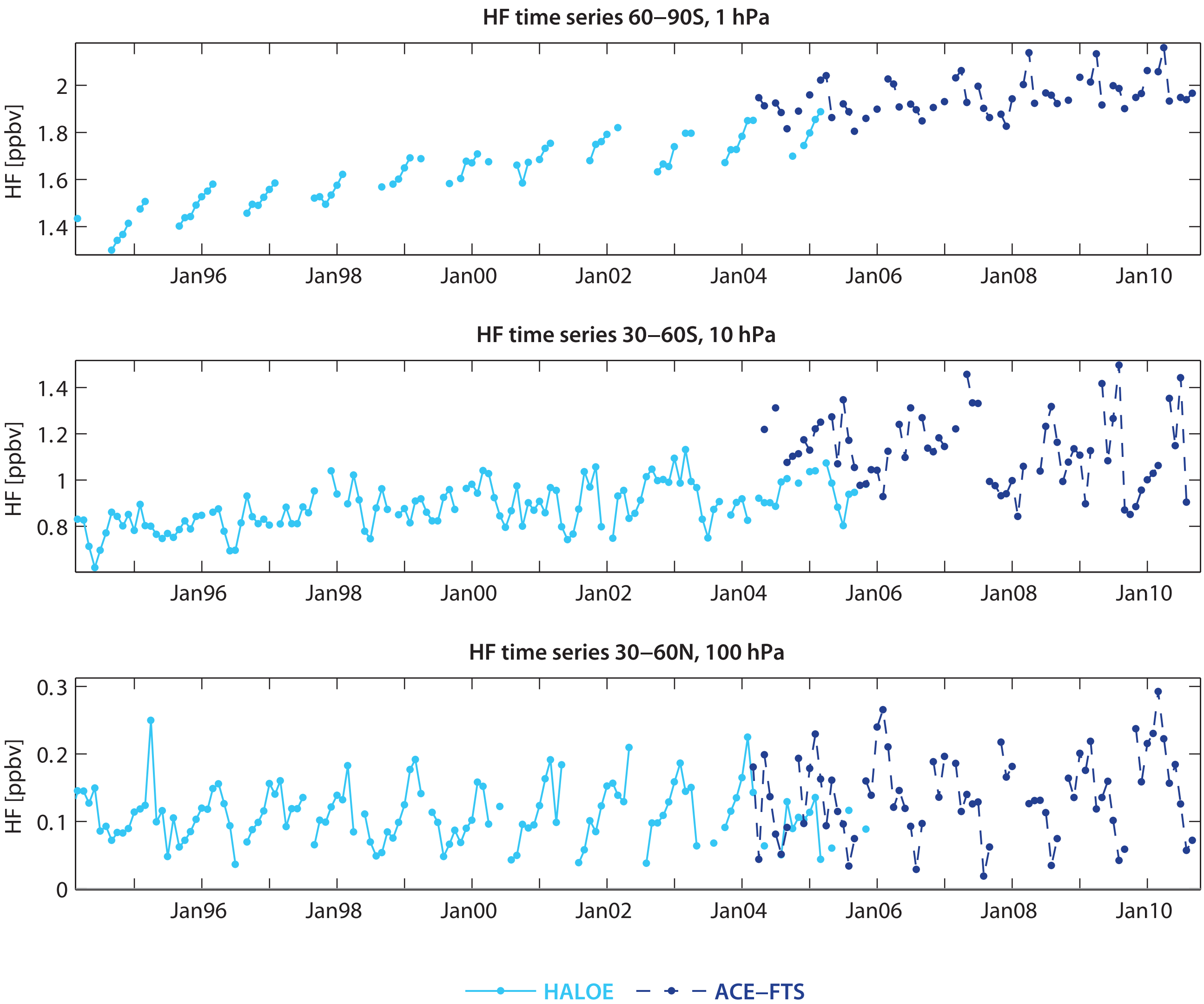


Figure 4.8.5: Time series of HF monthly mean values in mid- and high latitudes. *Monthly mean HF values between 60°S-90°S at 1 hPa (upper panel), 30°S-60°S at 10 hPa (middle panel), and 30°N-60°N at 100 hPa (lower panel).*

the summary plot of HF differences for high latitudes (see **Figure A4.8.6** in *Appendix A4*) show that differences in the NH and SH high latitude are of the same magnitude compared to differences at lower latitudes (**Figure 4.8.7**). Below 100 hPa, the HF annual mean state is less well known, with a 1σ multi-instrument spread of ±30% or larger.

Instrument-specific conclusions

ACE-FTS observes more HF than HALOE in the region above 50 hPa, although both datasets agree very well and show only small relative differences from the MIM (up

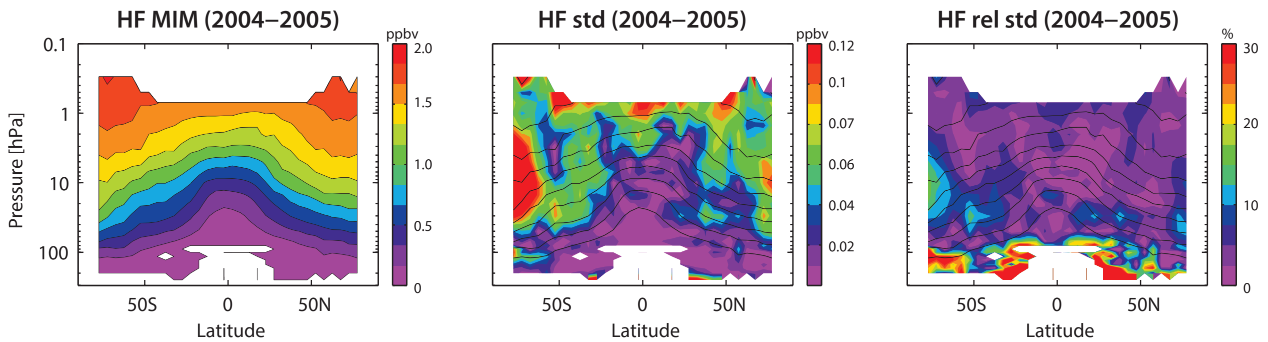


Figure 4.8.6: Summary of HF annual zonal mean state for 2004-2005. *Shown are the annual zonal mean cross section of the HF MIM (left panel), the standard deviation over both instruments (middle panel), and relative standard deviation with respect to the MIM (right panel). Black contour lines in the right panels give the MIM distribution. Instruments included are HALOE and ACE-FTS. The MIM and standard deviation are only displayed for regions where both instruments provide measurements.*

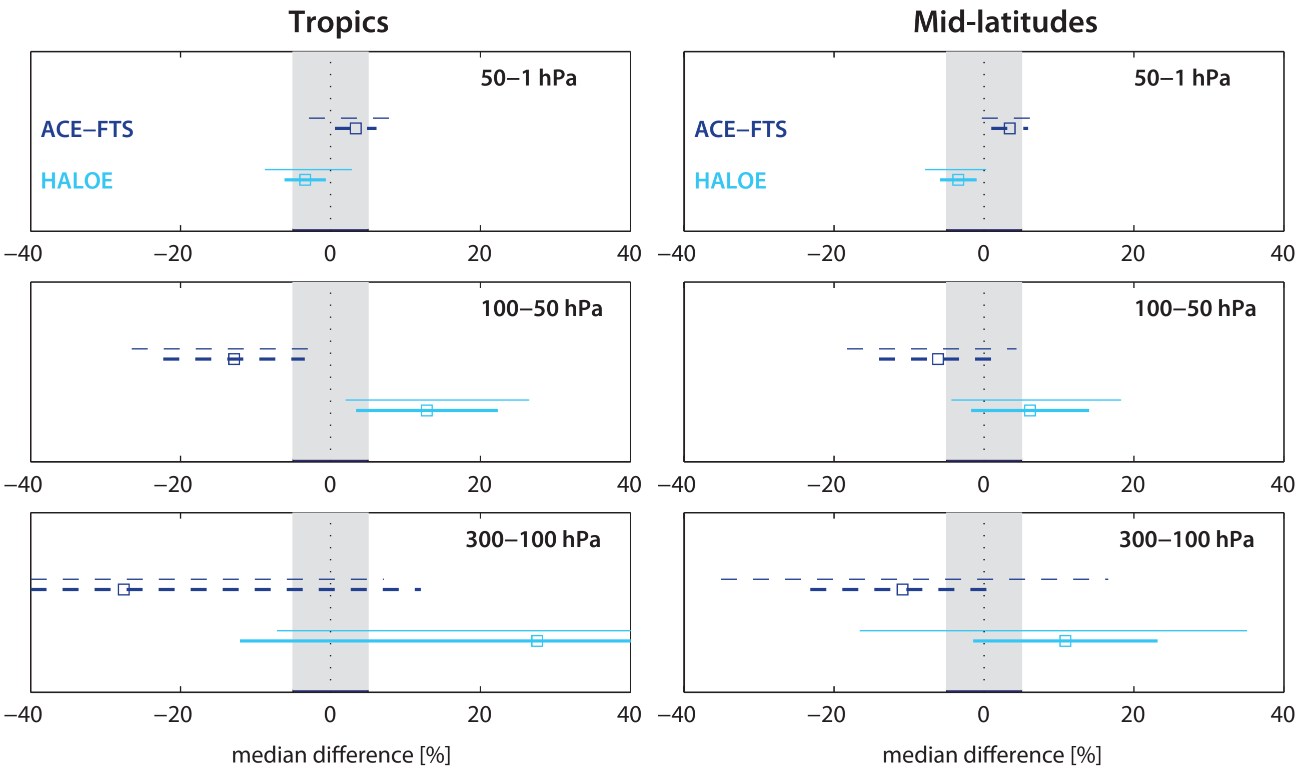


***Figure 4.8.7:** Summary HF differences for 2004-2005. Over a given latitude and altitude region the median (squares), median absolute deviation (MAD, thick lines), and the standard deviation (thin lines) of the monthly mean relative differences between an individual instrument-climatology and the MIM are calculated. Results are shown for the tropics (30°S-30°N) and mid-latitudes (30°S-60°S and 30°N-60°N) and for 3 different altitude regions from the UT up to the MS between 300 and 1 hPa for the reference period 2004-2005.*

to ±5%). Below 50 hPa, HALOE detects more HF than ACE-FTS, with differences of up to ±10% between 50 hPa and 100 hPa and below 100 hPa in the mid-latitudes. The largest disagreement between the two instruments is found in the tropical UT where mean differences are about ±25%, and individual differences (for single latitude bands/ pressure levels) can be as large as ±50% as indicated by the large regional spread (**Figure 4.8.7**). For the two-year-long overlap period, both datasets agree roughly on the seasonal and interannual variability, with some differences found in month-to-month variations. Annual mean cross sections show some kinks at 50°-60°N/S for HALOE and 70°S-80°S for ACE-FTS, which are thought to be related to sampling issues.

4.9 Sulfur hexafluoride – SF_6

SF_6 is a gas of tropospheric origin and is mainly used in large electrical equipment, from which it escapes into the atmosphere during maintenance. Once in the atmosphere it absorbs infrared radiation, and is one of the most efficient greenhouse gases. SF_6 is chemically inert in the troposphere and stratosphere, and is only removed through transport into the mesosphere where it is destroyed by photolysis or electron-capture reactions [*Morris et al.*, 1995; *Reddmann et al.*, 2001]. As a result, it has an atmospheric lifetime of hundreds to thousands of years [*Ko et al.*, 1993; *Ravishankara et al.*, 1993]. Growing anthropogenic SF_6 emissions over past few decades have led to an increase of SF_6 in the atmosphere [*Geller et al.*, 1997]. This fact, in combination with its long lifetime, makes SF_6 a suitable tracer to derive estimates of the mean age of stratospheric air [*Stiller et al.*, 2008], which can be used as a measure of the intensity of the Brewer-Dobson circulation [*Austin and Li*, 2006]. Due to recent model predictions of an intensification of the Brewer-Dobson circulation, observational evidence of long-term changes of age of air are a focus of ongoing research. In order to derive reliable proxies for trends in the stratospheric circulation from SF_6 data, one needs to account for the non-uniform SF_6 growth rates [*Garcia et al.*, 2010].

4.9.1 Availability of SF_6 measurements

Measurements of SF_6 are available from 2004 onward from ACE-FTS, and from 2005 onward from MIPAS. While ACE-FTS covers the UT to MS up to 7 hPa, MIPAS measurements extend through the US up to 0.7 hPa. **Tables 4.9.1** and **4.9.2** compile information on the availability of SF_6 measurements, including time period, altitude range, vertical resolution, and references relevant for the data products used in this report.

***Table 4.9.1:* Available SF_6 measurement records from limb-sounding satellite instruments between 1978 and 2010.** *The red filling of the grid boxes indicates the temporal and vertical coverage of the respective instruments.*

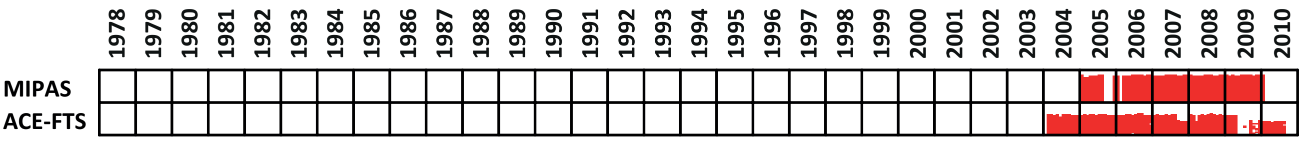


***Table 4.9.2:* Time period, vertical range, vertical resolution, references and other comments for SF_6 measurements.**

Instrument	Time period	Vertical range	Vertical resolution	References	Additional comments
MIPAS V201	Jan 05 – Apr 12	6 – 50 km	4 – 6 km	*Stiller et al.*, 2008	
ACE-FTS V2.2	Mar 04 –	6 – 35 km	3 – 4 km	*Brown et al.*, 2011	

4.9.2 SF_6 evaluations: Zonal mean cross sections, vertical and meridional profiles

Annual zonal mean cross sections for the time period 2005-2010 are analysed to investigate mean differences between the various datasets. Additionally, vertical and meridional profiles are evaluated. Note that although only two datasets are available, the comparison of both datasets to their MIM (and not a direct comparison) will be used for consistency with the rest of the report.

Figure 4.9.1 shows the annual zonal mean SF_6 climatologies for 2005-2010 for ACE-FTS and MIPAS. SF_6 decreases with altitude due to the combination of its very long lifetime, growing tropospheric emissions, and stratospheric transport time scales. The SF_6 isopleths slope downwards towards higher latitudes as a result of air mass transport by the Brewer-Dobson circulation. While MIPAS and ACE-FTS measurements show a similar annual mean SF_6 distribution overall, some clear differences exist. ACE-FTS shows much noisier isopleths, very likely as result of its sparser sampling, as well as more scatter in the retrieved profiles than for some other ACE-FTS products. Apart from the noisy structure, ACE-FTS isopleths, in particular at 4.5 and 5 ppbv, are less steep than the corresponding MIPAS isopleths. Another notable feature are the peaks in MIPAS SF_6 data in the UTLS at the 5.5 and 6 ppbv isopleths near 25°S/25°N. These peaks are visible in the annual mean climatologies; however, monthly mean evaluations (see **Figures A4.9.1**-**A4.9.8** in *Appendix A4*) demonstrate that they are most pronounced in the respective winter/spring hemisphere. This phenomenon is possibly related to the seasonality of mixing and upwelling in the tropical UTLS, and indicates younger air in this region [*Stiller et al.*, 2012]. The effect could also be intensified by temperature artefacts.

The relative differences of MIPAS and ACE-FTS annual means with respect to the MIM are displayed in **Figure 4.9.2**. Below 50 hPa, the two instruments show excellent agreement, with differences from the MIM mostly below ±2.5%. Above 50 hPa, the relative differences increase slightly but still agree within ±5%. Except for some small regions (*e.g.*, the UTLS in the SH), SF_6 measurements from ACE-FTS are higher than the ones from MIPAS. The largest differences (±10% to ±20%) can be observed at 30°N/S at the 10 hPa level, and at high latitudes in the SH at the

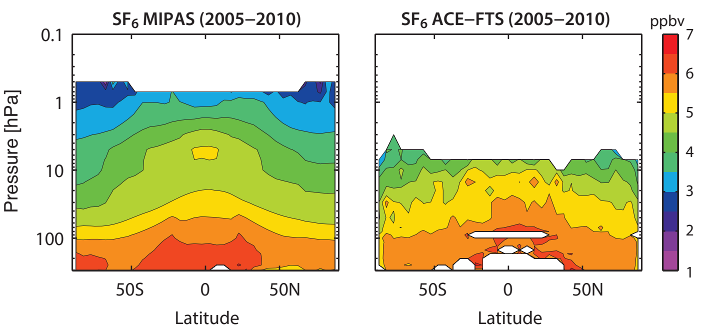


***Figure 4.9.1:* Cross sections of annual zonal mean SF_6 for 2005-2010.** *SF_6 cross sections are shown for MIPAS and ACE-FTS.*

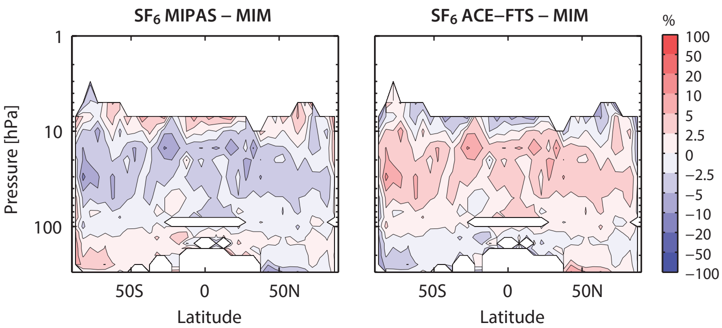


***Figure 4.9.2:* Cross sections of annual zonal mean SF_6 differences for 2005-2010.** *SF_6 differences between MIPAS, ACE-FTS and their MIM are shown. Note that, since the MIM consists of only two instruments, any issue with one dataset will fully be reflected by the difference of the other dataset.*

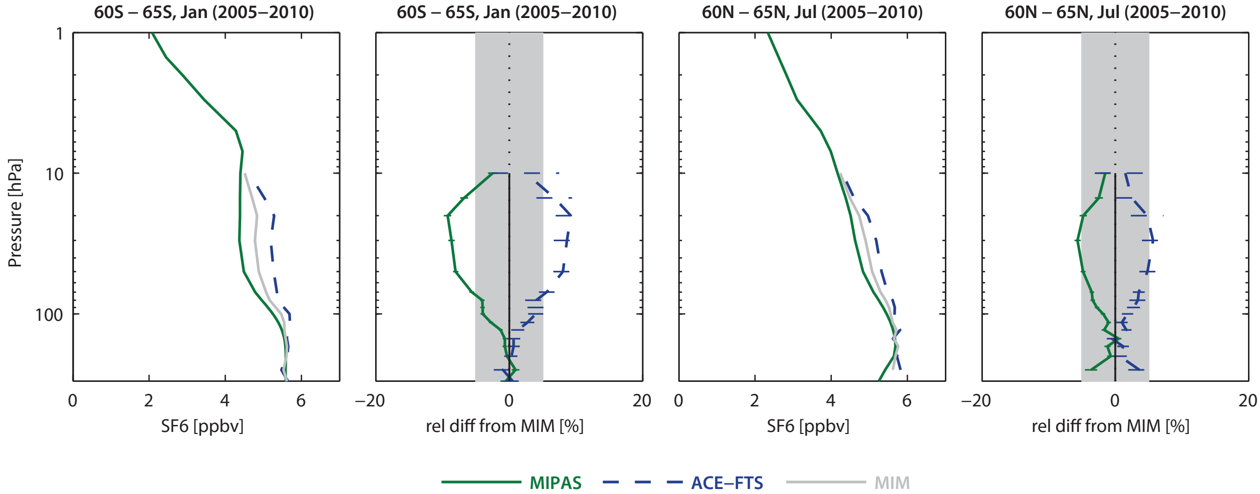


Figure 4.9.3: Vertical profiles of monthly zonal mean SF_6 for 2005-2010. *Zonal mean SF_6 profiles for 60°S-65°S, January (left panels) and 60°N-65°N, July (right panels) are shown together with their differences from the MIM. The grey shading indicates the ±5% difference range. Bars indicate the uncertainties in the relative differences.*

30 hPa level, related to isolated elevated values from ACE-FTS in these regions. While the monthly mean comparisons (**Figures A4.9.1**-**A4.9.8** in *Appendix A4*) are generally consistent with the annual mean comparison, slightly larger deviations between the instruments can be observed for some monthly mean evaluations (*e.g.*, January).

Monthly mean vertical SF_6 profiles are shown in **Figure 4.9.3** together with their differences from the MIM for SH/NH high latitudes in summer. The two datasets show excellent agreement at the lowest levels (~200 hPa) and at the upper levels at (~10 hPa) with differences of ±1%. In between these levels, the MIPAS profile has a different vertical gradient when compared to ACE-FTS, with a stronger SF_6 decrease below 50 hPa, and a weaker decrease above 20 hPa, resulting in maximum differences of ±5% at around 20 hPa. Meridional SF_6 profiles at 20 hPa for different months (see **Figure A4.9.9** in *Appendix A4*) confirm differences at this level of mostly ±5%, occasionally reaching ±10%, with larger ACE-FTS abundances everywhere except for very high NH latitudes in September.

4.9.3 SF_6 evaluations: Interannual variability and seasonal cycle

Figure 4.9.4 shows the time series of tropical monthly mean values as well as the deseasonalised anomalies from 2004 to 2010 at 20 hPa. Both datasets show increasing values over their respective lifetimes indicating a positive trend as the dominant signal. The seasonal cycle and interannual variability are rather weak for MIPAS, while ACE-FTS displays large month-to-month fluctuations. These fluctuations are the reason why *Brown et al.* [2011] used annual averages in their ACE-FTS trend study. Note that the low interannual anomalies in the MIPAS time series at the end of each calendar year are caused by the lack data available for these three months for the first year of the measurement period. The inter-annual anomalies of ACE-FTS are larger than the MIPAS anomalies at the beginning of the time period, but mostly lower than MIPAS at the end of the time period after 2008, pointing to a different long-term behaviour of the two datasets in this region.

The evaluation of monthly mean time series and anomalies in the NH mid-latitudes is shown in **Figure 4.9.5**. Here, MIPAS displays a weak seasonal cycle with maximum SF_6 abundance during the NH winter months. ACE-FTS on the other hand does not show a clear seasonal signal but is dominated by strong month-to-month fluctuations. The deseasonalised anomalies of the two datasets do not agree on the month-to-month or year-to-year scale, but show consistent results regarding their long-term changes with a clear increase of SF_6 during the displayed time period. In the SH high latitudes, ACE-FTS and MIPAS show the best agreement regarding the SF_6 seasonal cycle and interannual variations (**Figure 4.9.6**). MIPAS has a clear seasonal cycle with elevated values during the SH late summer/early autumn months. Note that SF_6 from MIPAS is also enhanced in September, which is in the middle of a time period of otherwise minimum SF_6 abundance. While frequent data gaps make it impossible to detect a clear seasonal cycle in ACE-FTS, the data indicate elevated values in winter consistent with the MIPAS signal. The interannual anomalies of the two datasets are roughly consistent and display the same long-term change with an increase of the SF_6 abundance.

4.9.4 Summary and conclusions: SF_6

A comparison of two SF_6 profile climatologies (MIPAS and ACE-FTS) has been carried out. Overall findings on the systematic uncertainty in our knowledge of the SF_6 mean state and important characteristics of the individual datasets are presented in the following summary including two synopsis plots. The first summary plot (**Figure 4.9.7**) provides information on the mean state and its uncertainty derived from the spread between the datasets. The second summary plot (**Figure 4.9.8**) shows specific inter-instrument differences in form of the deviations of the instrument climatologies to the MIM climatology. For each instrument and selected region, the deviation from the MIM is given as the median (mean) difference over all grid points in this region. Additionally, for each instrument the spread of the differences over all grid points in this region is presented. Note that both pieces of information (average deviation and spread) are important for a meaningful assessment of inter-instrument differences. A detailed description of the summary plot evaluations can be found in *Section 3.3.5*.

Atmospheric mean state

The uncertainty in our knowledge of the atmospheric SF_6 annual mean state as derived from these satellite datasets is small throughout the UT to the MS, with a 1σ multi-instrument spread of less than ±5%. The only exceptions are individual localised grid points where the spread reaches values of ±12%. The uncertainty in our knowledge of the SF_6 mean state is especially small below 50 hPa where the two instruments give a spread of ±2%. Note that ACE-FTS

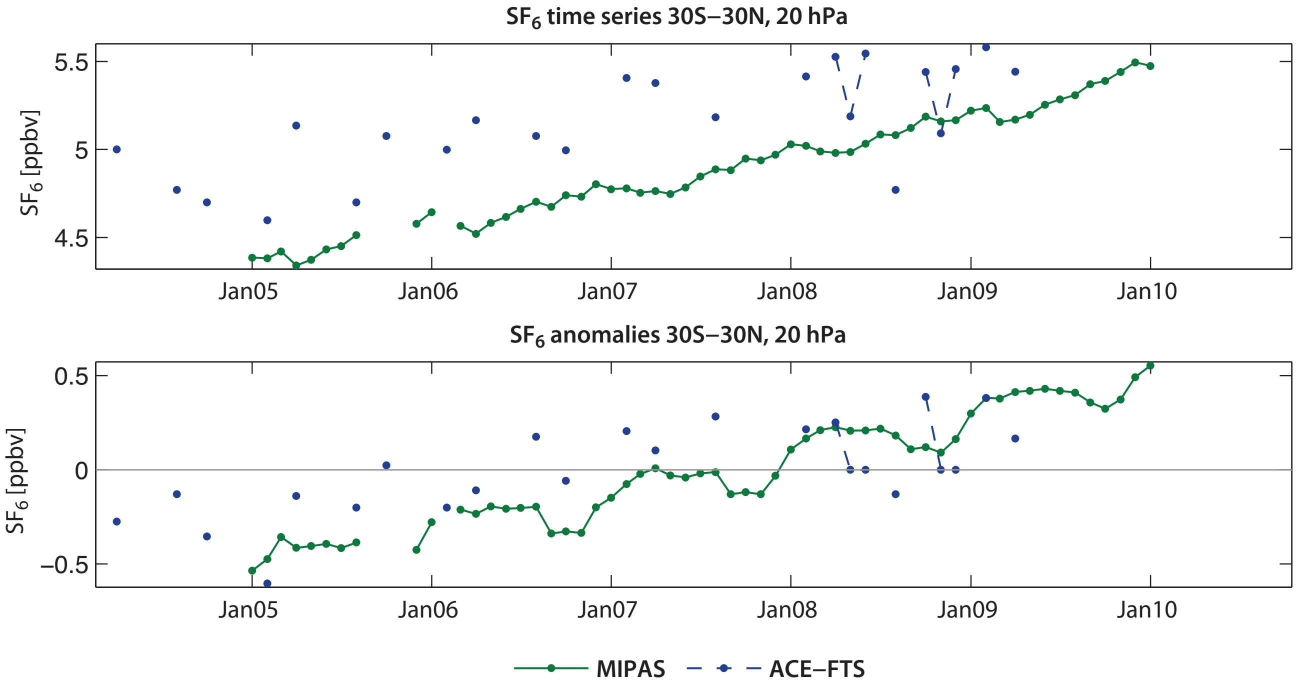


***Figure 4.9.4:** **Time series of SF_6 monthly mean values and deseasonalised anomalies in the tropics.** Monthly mean values (upper panel) and deseasonalised anomalies (lower panel) of SF_6 between 30°S – 30°N at 20 hPa.*

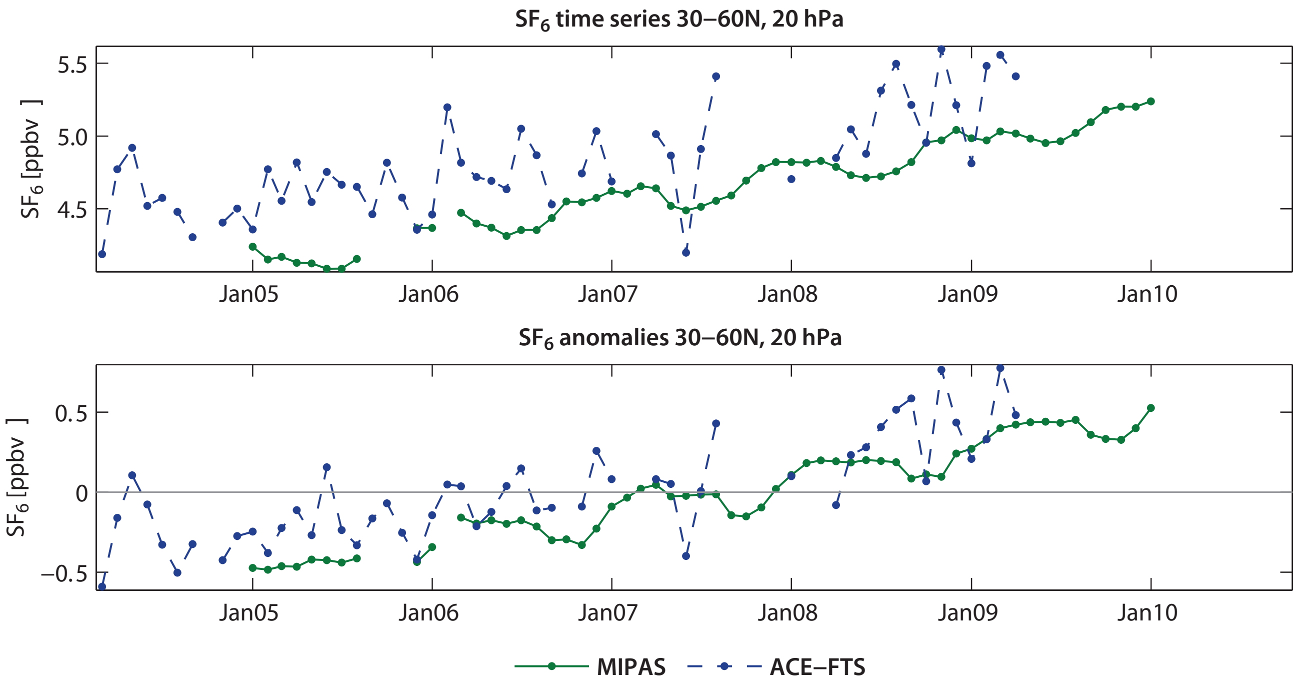


***Figure 4.9.5:** **Time series of SF_6 monthly mean values and deseasonalised anomalies in the NH mid-latitudes.** Monthly mean values (upper panel) and deseasonalised anomalies (lower panel) of SF_6 between 30°N – 60°N at 20 hPa.*

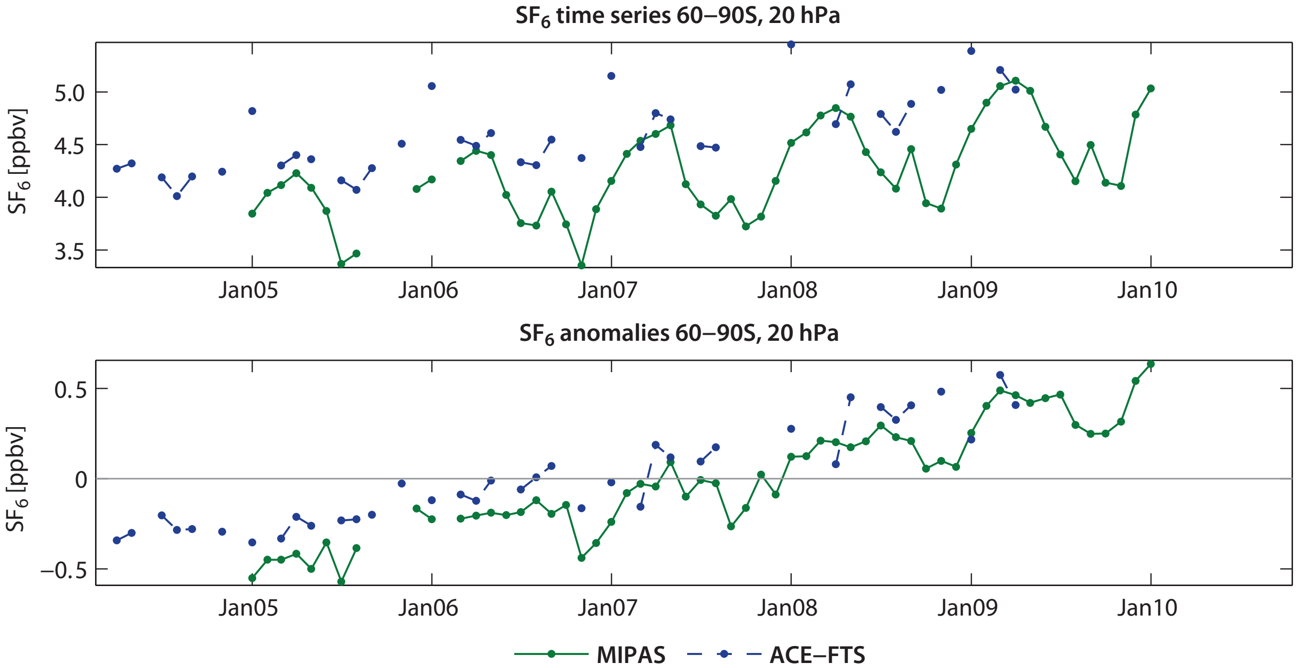


***Figure 4.9.6: Time series of SF$_6$ monthly mean values and deseasonalised anomalies at SH high latitudes.** Monthly mean values (upper panel) and deseasonalised anomalies (lower panel) of SF$_6$ between 60°S – 90°S at 20 hPa.*

and MIPAS measure SF$_6$ around the same band, and it is therefore possible that the two datasets share systematic error components.

Instrument-specific conclusions

MIPAS detects less SF$_6$ than ACE-FTS in most atmospheric regions, with small differences of around -2.5% with respect to their MIM. Above 10 hPa and in the SH extratropics below 100 hPa, MIPAS is larger than ACE-FTS. In the UTLS around 25°S/25°N, MIPAS shows some elevated mixing ratio peaks, which are most pronounced in the respective winter/spring hemisphere. In addition to SF$_6$, the phenomenon is also apparent in the MIPAS CFC-12 and, to a smaller degree, CFC-11 latitudinal profiles in the UTLS with the same seasonal dependence.

ACE-FTS detects more SF$_6$ than MIPAS (+2.5% difference from the MIM), which is consistent with the ACE-FTS trend comparisons made by *Brown et al.* [2011] with results from the SLIMCAT chemical transport model. ACE-FTS shows less steep and much noisier SF$_6$ isopleths when compared to MIPAS, likely as result of its sparser sampling and more scatter in the retrieved profiles used as input for the climatology. Furthermore, ACE-FTS does not decrease as fast with increasing altitude in the LS. The evaluation of the monthly zonal mean time series reveals that ACE-FTS shows pronounced month-to-month variations and no clear seasonal cycle.

4.10 Nitrogen monoxide – NO

Tropospheric NO is released from fossil fuel combustion and is a key air pollutant responsible for the formation of

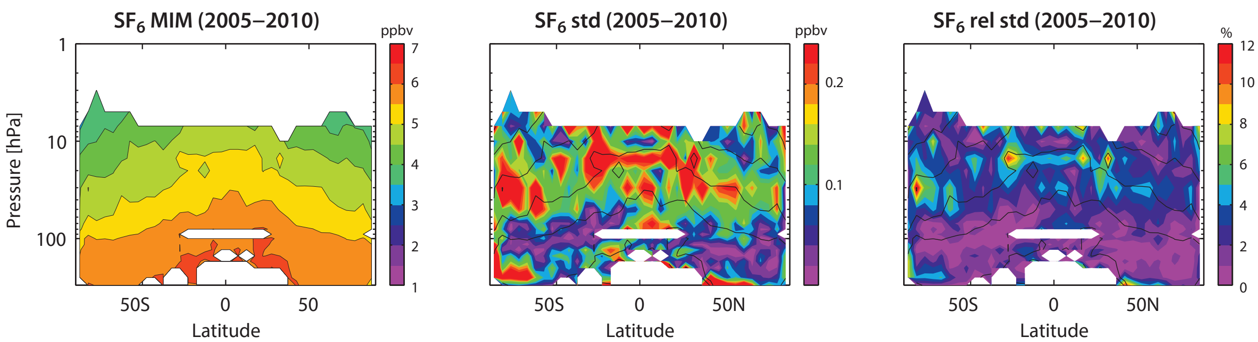


***Figure 4.9.7: Summary of SF$_6$ annual zonal mean state for 2005-2010.** Shown are the annual zonal mean cross section of the SF$_6$ MIM (left panel), the standard deviation over both instruments (middle panel), and the relative standard deviation with respect to the MIM (right panel). Black contour lines in the right panels give the MIM distribution. Instruments included are MIPAS and ACE-FTS. The MIM and standard deviation are only displayed for regions where both instruments provide measurements.*

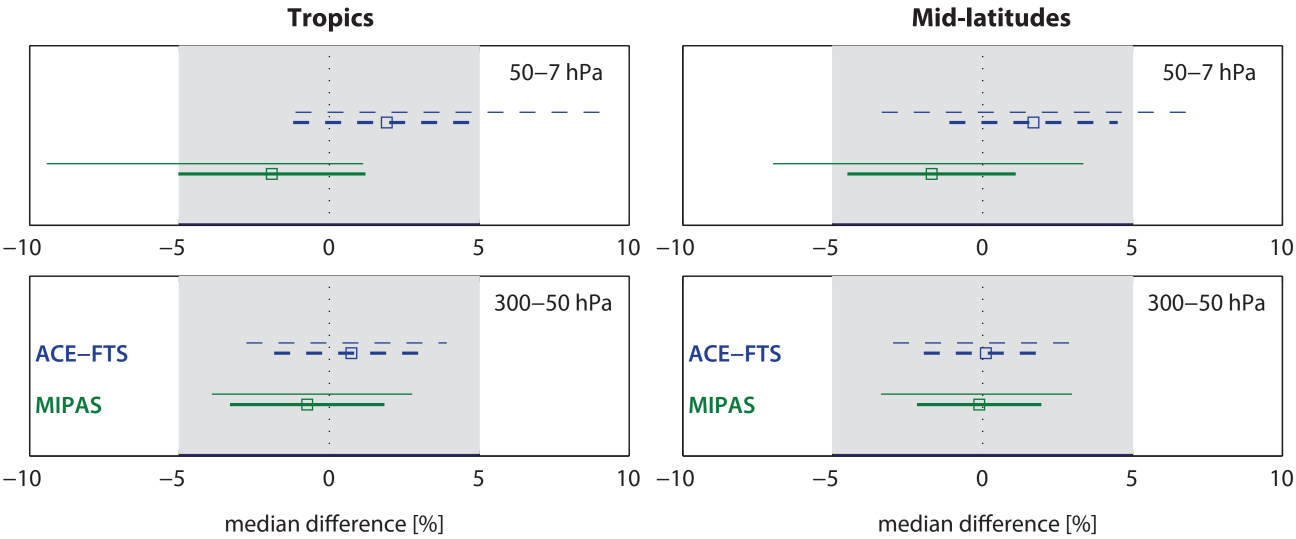


Figure 4.9.8: Summary SF_6 differences for 2005-2010. *Over a given latitude and altitude region the median (squares), median absolute deviation (MAD, thick lines), and the standard deviation (thin lines) of the monthly mean relative differences between an individual instrument-climatology and the MIM are calculated. Results are shown for the tropics (30°S-30°N) and mid-latitudes (30°S-60°S and 30°N-60°N) and for 2 different altitude regions from the UT up to the MS between 300 and 7 hPa for the reference period 2005-2010.*

smog and acid rain. In the stratosphere, NO is produced from the oxidation of N_2O, originating from soil emissions (see *Section 4.4*). Additionally, NO is an important component of aircraft exhaust generated by oxidation of N_2 at high temperatures within aircraft engines. NO has also a thermospheric source (due to particle precipitation and soft X-rays) which can indirectly contribute to stratospheric NO *via* descent during polar winters. In the stratosphere, there is a rapid exchange between NO and NO_2, which together from the reactive nitrogen chemical family NO_x (see *Section 4.12*). Through the catalytic NO_x cycle, NO is involved in the chemical ozone depletion.

Stratospheric NO has a strong diurnal cycle with large NO abundances during daytime, extremely low NO abundances during nighttime, and steep gradients at local sunrise (SR) and sunset (SS). **Figure 4.10.1** shows examples of the diurnal NO cycle as a function of LST for three different pressure levels as derived from a chemical box model [*McLinden et al.*, 2010]. A direct comparison of satellite-based NO measurements that correspond to different LSTs is not possible unless the dependence on the SZA is taken into account.

4.10.1 Availability of NO measurements

Measurements of NO are available from two solar occultation instruments, HALOE and ACE-FTS, which have overlapping records for 2004 and 2005. Solar occultation measurements are always made at SZA = 90° and can therefore be directly compared if separated into local sunrise and sunset. Furthermore NO measurements are availabe from the limb emission instruments MIPAS

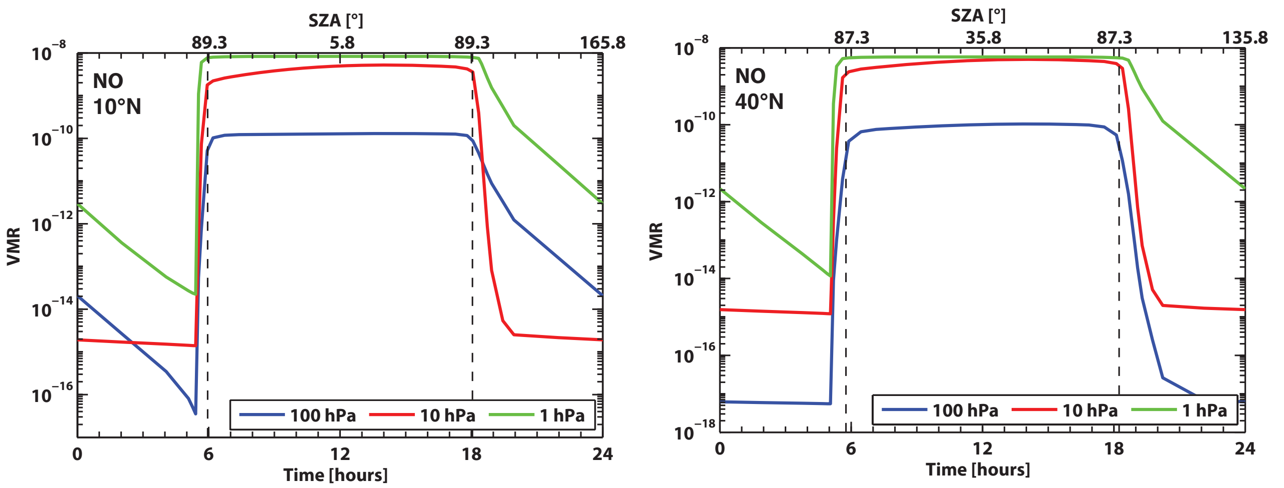
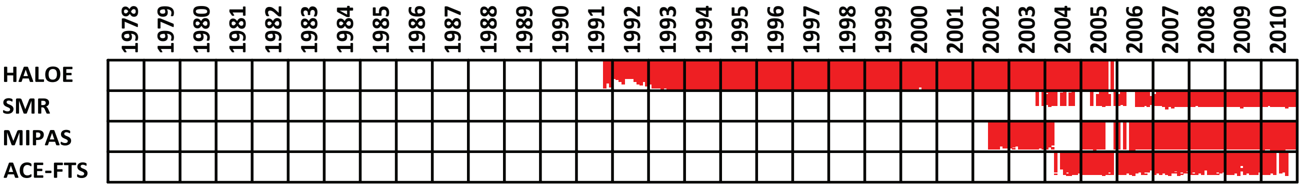


Figure 4.10.1: Diurnal NO cycle. *NO variations as function of LST are shown at 10°N and 40°N at 1, 10, and 100 hPa for March 15.*

Table 4.10.1: Available NO measurement records from limb-sounding satellite instruments between 1978 and 2010. The red filling of the grid boxes indicates the temporal and vertical coverage of the respective instruments.



and SMR. For a comparison of these two instruments with each other and with the solar occultation instruments, the difference in LST must be taken into account. This correction is done by scaling the SMR (corresponding to approximately 6am/pm LST) and ACE-FTS measurements with a chemical box model [*McLinden et al.*, 2010] to the LST of the MIPAS measurements at 10am/pm. **Tables 4.10.1** and **4.10.2** compile information on the availability of NO measurements, including time period, vertical range and resolution, and references relevant for the data products used in this report.

4.10.2 NO evaluations: Zonal mean cross sections and vertical profiles

Monthly zonal mean cross sections are analysed to investigate mean biases between the various datasets. Additionally, vertical profiles are evaluated. Note that if only two datasets are available, the comparison of both datasets to their MIM (and not a direct comparison) will be used to stay consistent with other parts of the report.

HALOE and ACE-FTS (2004-2005)

Figure 4.10.2 shows the monthly zonal mean NO local sunrise climatologies for February and August 2004-2005, and local sunset climatologies for April and July 2004-2005 for HALOE and ACE-FTS. The comparisons for sunrise and sunset are based on different months in order to ensure a maximum overlap between the two instruments. The local sunrise/sunset mixing ratios for NO are very small below 10 hPa, but increase above with a maximum at 1 hPa. While both datasets show the same overall structure of monthly mean NO fields, some clear differences exist. ACE-FTS has more moderate vertical gradients above 1 hPa when compared to HALOE. ACE-FTS observes very high mixing ratios above 1 hPa at high latitudes in the winter hemisphere, related to the descent of upper mesospheric and thermospheric NO_x produced by ionizing energetic particle precipitation. HALOE has no coverage in this latitude region. Note that HALOE includes a diurnal correction in its retrieval, which provides small corrections for the summer high latitudes [*McHugh et al.,* 2005].

The relative differences of HALOE and ACE-FTS from the MIM are displayed in **Figure 4.10.3**. In the UTLS and MS, HALOE shows larger values than ACE-FTS, while in the US and LM HALOE measures less NO. In the US, the relative differences are small (±5 to ±10%) but increase above and below this region (up to ±50%). The deviations are consistent for different months, for sunrise and sunset measurements, and between coincident profile comparisons [*Kerzenmacher et al.*, 2008].

Figure 4.10.4 shows monthly mean NO profiles together with their differences from the MIM. The comparison for the NH mid-latitudes (35°N-40°N) in August shows very good agreement between the two local sunrise datasets, with only small differences (up to ±5%) in the US and LM. These differences increase for levels above 0.5 hPa, where deviations increase due to the steeper vertical gradients of the HALOE NO field. For the other three cases shown in **Figure 4.10.4**, ACE-FTS has a flattened maximum when compared to HALOE resulting in differences of up to ±20% in some parts of the MS or US. For both local sunrise and local sunset conditions, ACE-FTS measures lower NO values than HALOE throughout the stratosphere, and higher NO values in the mesosphere. An exception is the situation in December at 45°N-50°N, with HALOE detecting higher NO values.

Table 4.10.2: Time period, vertical range, vertical resolution, references and other comments for NO measurements.

Instrument	Time period	Vertical range	Vertical resolution	References	Additional comments
HALOE V19	Oct 91 – Nov 05	up to 140 km	3.5 km	*Grooß and Russell,* 2005	
ACE-FTS V2.2	Mar 04 –	12 – 105 km	3 – 4 km	*Kerzenmacher et al.,* 2008	
SMR V2-1	Oct 03 –	30 – 60 km 80 – 110 km	4 – 6 km 6 – 8 km	*Sheese et al.,* 2013	Only 1 day per month prior to April 2007
MIPAS V15 V220	 Mar02 – Mar04 Jan05 – Apr12	 12 – 70 km	 3.5 – 5 km 2.5 – 6 km	 *Funke et al.,* 2005a *Funke et al.,* 2005b	2005: Change in spectral resolution

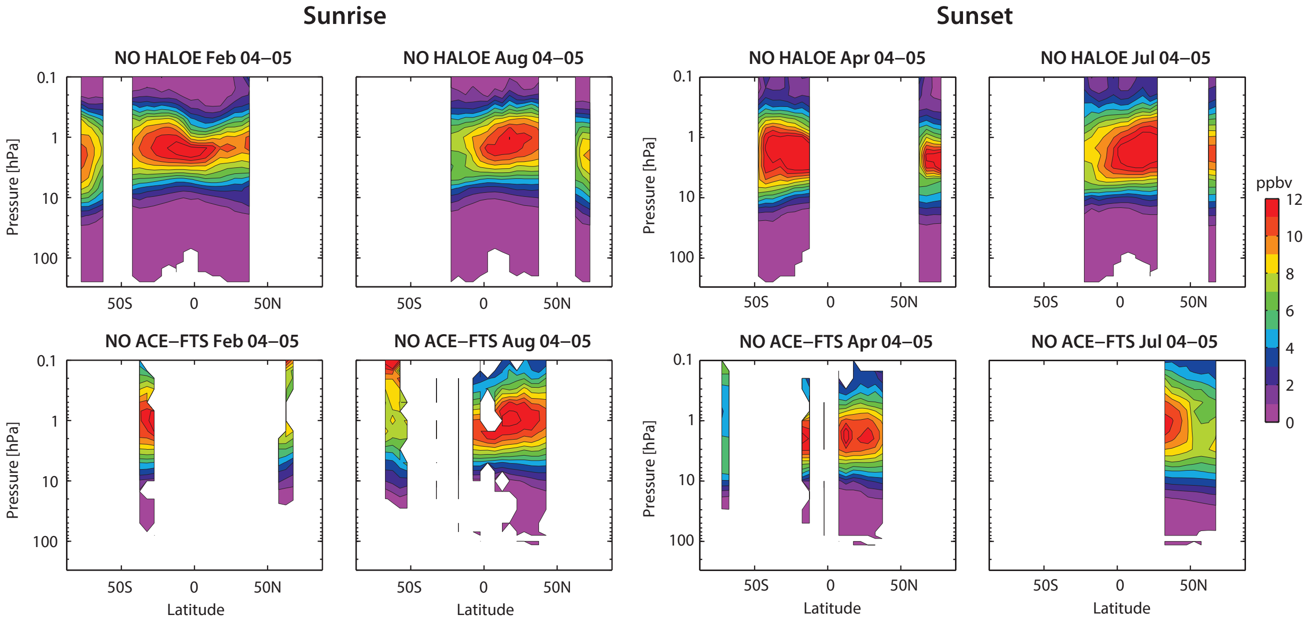


Figure 4.10.2: ***Cross sections of monthly zonal mean, local sunrise and sunset NO for 2004-2005.*** *Monthly zonal mean, local sunrise for February and August (column 1 and 2) and local sunset for April and July (column 3 and 4) NO cross sections are shown for HALOE (upper row) and ACE-FTS (lower row).*

ACE-FTS, MIPAS, and SMR (2005-2010)

In order to compare the two emission instruments, MIPAS measurements are split into am and pm climatologies corresponding to 10am LST and 10pm LST, respectively. Furthermore, SMR am measurements are scaled to 10am and SMR pm measurements are scaled to 10pm by using tabulated diurnal cycles from a chemical box model. For the scaling of SMR, the input climatologies are restricted to SZA's smaller than 87.5°, so that only the sunlit data are used for scaling. Due to the Odin orbit, measurements are performed at mid and high latitudes in the summer hemisphere only. In the tropics, the orbit provides measurements at twilight such that the ascending node observations occur near 6:00am LST and descending node observations occur near 6:00pm LST. The solar occultation dataset from ACE-FTS is also scaled to 10am and 10pm using the same box model, and can thus be compared to MIPAS and scaled SMR.

Figure 4.10.5 shows August monthly mean cross sections for the three datasets corresponding to 10am. Additionally, unscaled SMR am data and ACE-FTS local sunrise data are

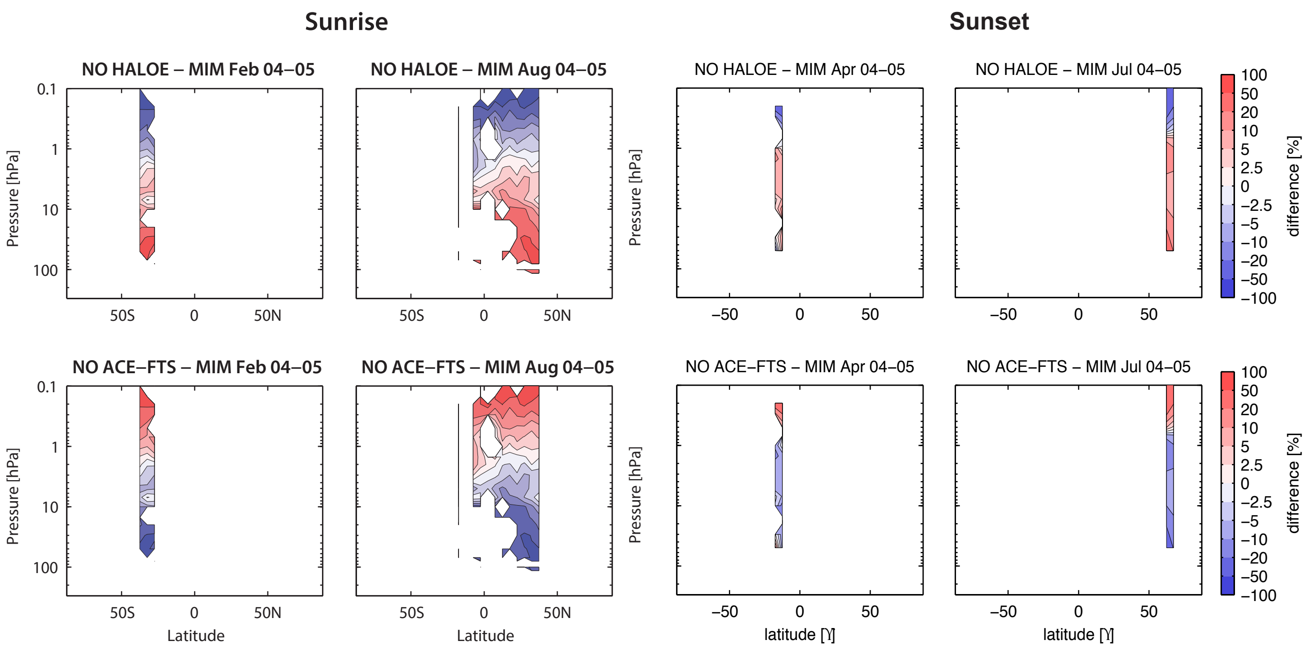


Figure 4.10.3: ***Cross sections of monthly zonal mean, local sunrise and sunset NO differences for 2004-2005.*** *Monthly zonal mean, local sunrise for February and August (column 1 and 2) and local sunset for April and July (column 3 and 4) NO differences between the individual instruments (HALOE and ACE-FTS) and their MIM are shown.*

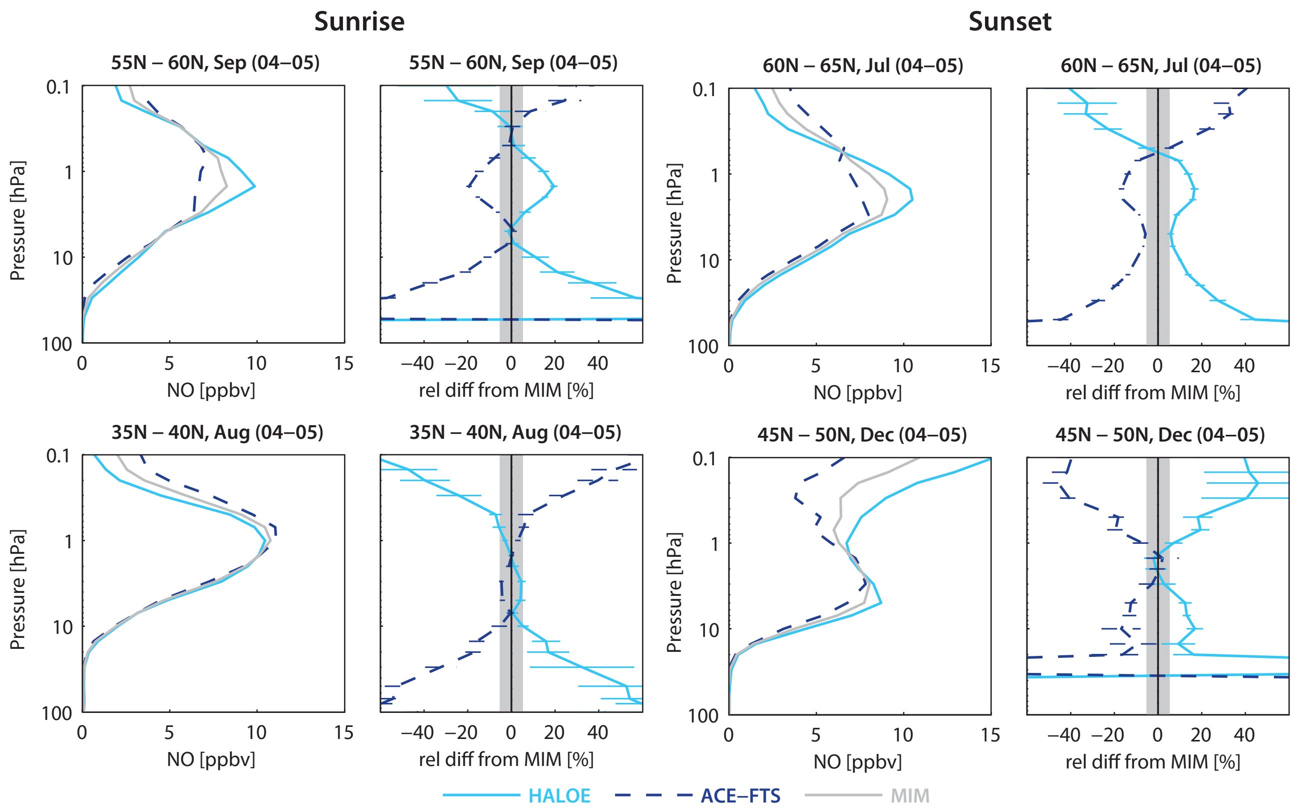


Figure 4.10.4:* *Profiles of monthly zonal mean, local sunrise and sunset NO for 2004-2005. *Zonal mean NO profiles are shown together with their differences from the MIM for local sunrise, 55°N-60°N, September and 35°N-40°N, August (column 1 and 2) and local sunset, 60°N-65°N, July and 45°N-50°N, December (column 3 and 4).*

shown, although not included in the MIM. Clearly, there are large differences between the datasets. In particular the scaled SMR climatology shows a different monthly mean NO distribution, with no meridional gradients from the tropics to the mid-latitudes, with overall larger NO abundances below 1 hPa, and steeper vertical gradients above this level.

Relative differences of the three datasets from the MIM corresponding to 10am are displayed in **Figure 4.10.6** (upper row). The comparison confirms that scaled SMR measures higher NO (except above 1 hPa), and that ACE-FTS and MIPAS agree better with each other than with SMR. Differences of SMR to the other two datasets are particularly high in the MS. ACE-FTS is mostly lower than MIPAS.

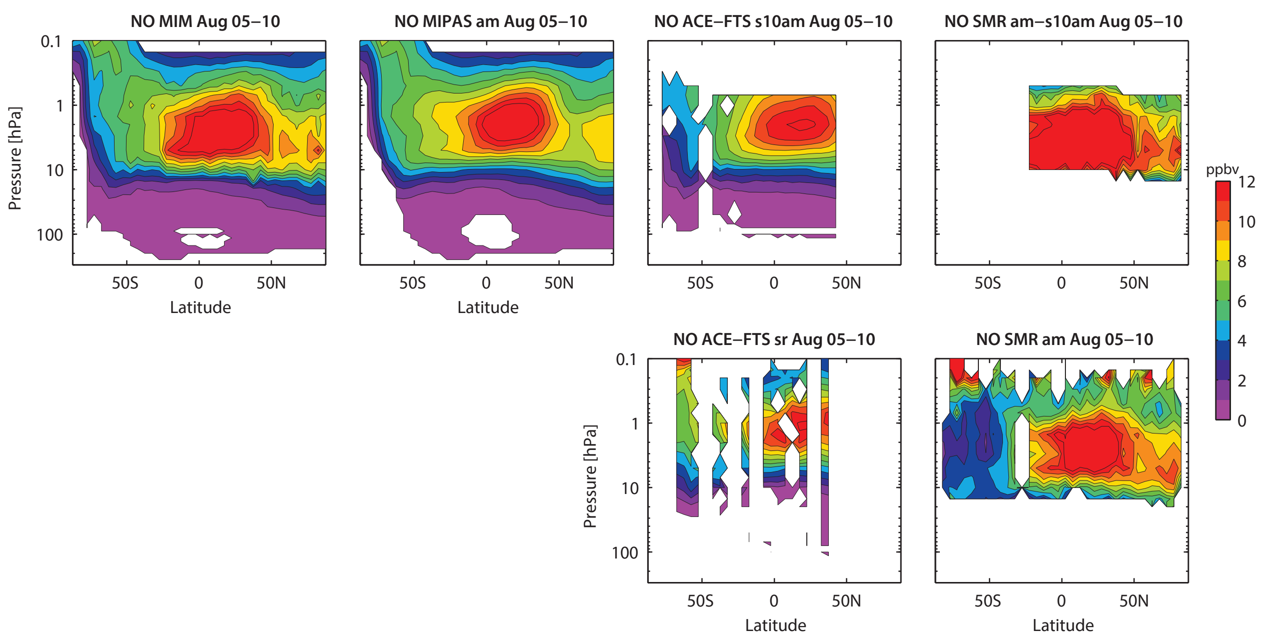


Figure 4.10.5:* *Cross sections of monthly zonal mean 10am NO for 2005-2010. *Monthly zonal mean 10am NO cross sections for August 2005-2010 are shown for the MIM, MIPAS (corresponding to 10am), ACE-FTS scaled to 10am (s10AM), and SMR am scaled to 10am data (am-s10am). Additionally, ACE-FTS local sunrise and SMR am data are shown but not included in the MIM.*

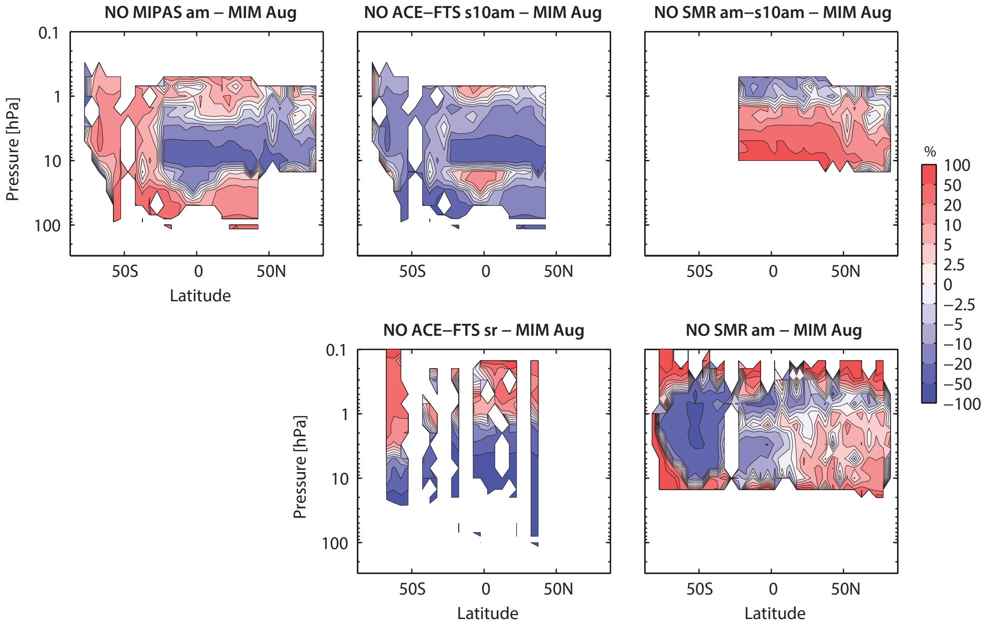


Figure 4.10.6: ***Cross sections of monthly zonal mean, 10am NO differences for 2005-2010.*** *Monthly zonal mean 10am NO differences for August 2005-2010 of MIPAS (corresponding to 10am), ACE-FTS scaled to 10am, and SMR am scaled to 10am with respect to their MIM are shown. Additionally, differences of ACE-FTS local sunrise and SMR am data with respect to the MIM of the three datasets above are displayed.*

The comparison of the unscaled ACE-FTS dataset with the 10am MIM illustrates that the scaling of the data with a chemical box model leads to better agreement between ACE-FTS and MIPAS, as one would expect. However, the same conclusions cannot be made for SMR, where differences of the unscaled dataset to the 10am MIM are in some cases smaller than for the scaled dataset, implying that either errors introduced by the scaling procedure affect the data product or that the unscaled data product already has a positive bias.

In order to analyse the differences in more detail, single monthly mean profiles are compared in **Figure 4.10.7**. In the LS, where NO mixing ratios are small MIPAS and scaled ACE-FTS show reasonably good agreement, with differences between ±10% and ±20%. In the MS, the differences between the two datasets are smaller: between ±5% and ±1%. Both instrument climatologies are on the low side when compared to scaled SMR (am to 10am), which exhibits differences of around +40% from the MIM (compared to 20% for the unscaled product). In the US and LM, SMR NO values approach those of the other two datasets, and overall deviations of the three instruments with respect to the MIM are around ±10%.

Due to the diurnal NO cycle, the 10pm climatologies are characterised by very low NO abundances, except for high latitudes during sunlit conditions. Monthly mean profiles of 10pm NO at high NH and SH latitudes during sunlight conditions are displayed in **Figure 4.10.8**. MIPAS and scaled ACE-FTS profiles show very similar shapes, and their absolute values agree very well in the MS and US, with differences up to ±5%. In the LS, however, they show considerable disagreement with differences reaching ±50%. Scaled SMR

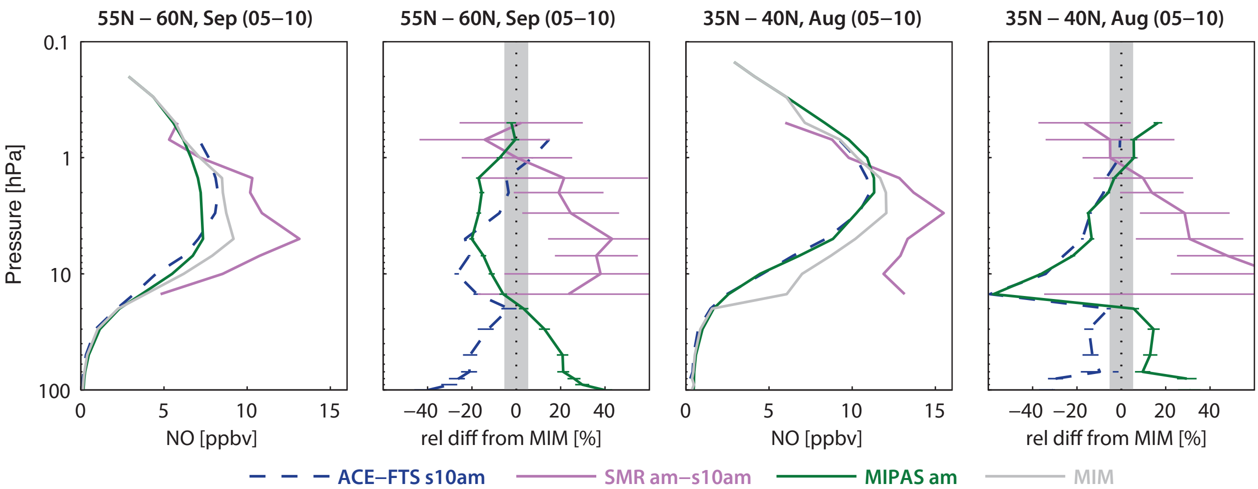


Figure 4.10.7: ***Vertical profiles of monthly zonal mean 10am NO for 2005-2010.*** *Zonal mean 10am NO profiles for 55°N-60°N, September and 35°N-40°N, August are shown together with their differences from the MIM.*

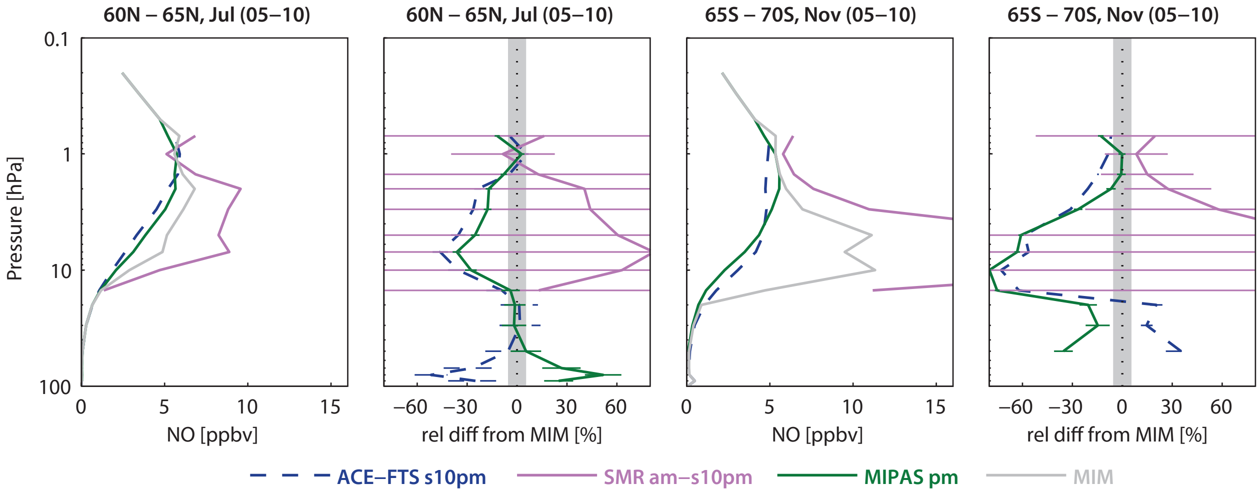


Figure 4.10.8: ***Vertical profiles of monthly zonal mean 10pm NO for 2005-2010.*** *Zonal mean 10pm NO profiles for 60°N-65°N, July and 65°S-70°S, November are shown together with their differences from the MIM.*

(pm to 10pm) shows unrealistically large values resulting in relative differences to the MIM of more than 100%.

4.10.3 NO evaluations: Seasonal cycles

Figure 4.10.9 displays the seasonal cycle of 10am NO climatologies for NH and SH high latitudes and tropics averaged over 2005-2010. The evaluations focus on the 10am climatologies since the 10pm climatologies provide only high-latitude data during times when 10pm corresponds to sunlight conditions, and therefore do not include enough data to evaluate seasonal variations.

At high latitudes, MIPAS and ACE-FTS display roughly the same seasonal cycle. In both hemispheres, MIPAS and ACE-FTS agree well on the minimum values, but MIPAS shows higher maxima, and therefore stronger amplitudes in the seasonal cycle. Additionally, the phase of the seasonal cycle is different, with an earlier minimum in MIPAS data. Note that SMR measures in the summer hemisphere during daytime, and in the winter hemisphere during nighttime, which does not allow for a full evaluation of the SMR seasonal cycle at high latitudes. Scaled SMR at the SH high latitudes shows a positive offset compared with the other two datasets, but has the correct tendencies for the seasonal variations. At the NH high latitudes, scaled SMR is too high, and does not agree on the seasonal signal shown by MIPAS, or by ACE-FTS for the months with data available.

In the tropics, all three instruments display a semi-annual cycle, and agree very well on the phase of the signal. SMR is characterised by an offset compared to the other two datasets during most of the year. SMR exhibits the strongest amplitude of the semi-annual cycle when compared to MIPAS and scaled ACE-FTS; ACE-FTS has the smallest seasonal cycle amplitude.

4.10.4 NO evaluations: Interannual variability

Apart from the climatological differences between the datasets, it is of interest to evaluate how well the instruments capture the interannual variability of NO. **Figure 4.10.10**

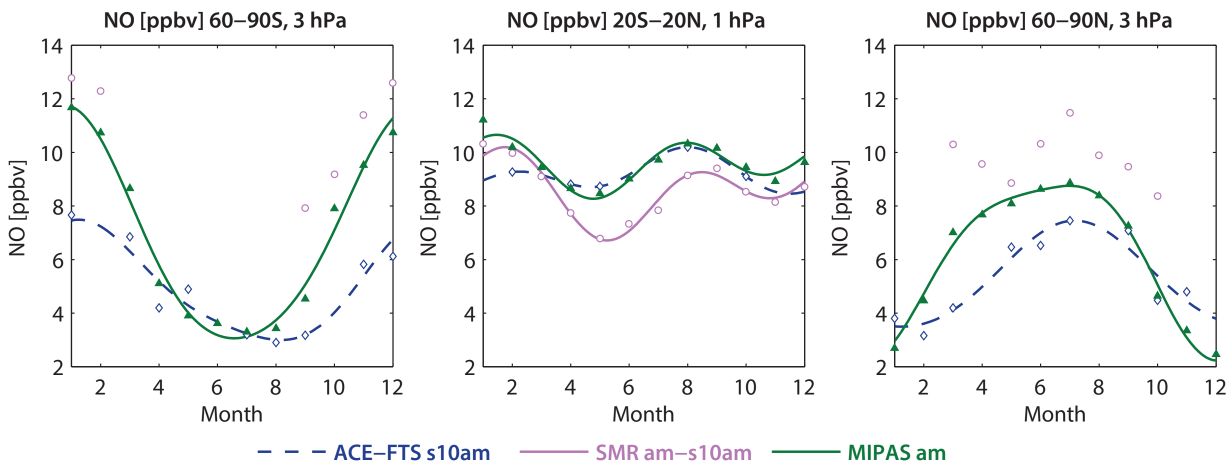


Figure 4.10.9: ***Seasonal cycle of 10am NO for 2005-2010.*** *Seasonal cycle of monthly zonal mean NO for 60°S-90°S, 3 hPa (left column), 20°S-20°N, 1hPa (middle column) and 60°N-90°N, 3 hPa (right column). Measurements correspond to 10am LST (MIPAS, filled symbols) or are scaled to 10am LST (ACE-FTS, SMR, open symbols).*

shows the time series of NO mean values (upper panels) and deseasonalised anomalies (lower panels) for the tropical latitude band 20°S-20°N at 1 hPa. Datasets corresponding to 10am LST are displayed in the left panels and the original datasets (corresponding to a variety of LSTs) are displayed in the right panels. The anomalies of the scaled datasets are calculated in an additive sense by subtracting monthly multi-year mean values for each month. Such additive anomalies, however, may also include a diurnal cycle, and are therefore not suitable evaluation tools for the unscaled datasets. Instead, the anomalies of the unscaled climatologies are calculated in a multiplicative sense as percentage deviations from the monthly multi-year mean values; a quantity that is less affected by the diurnal variations.

In the tropics, NO shows a cycle of approximately two years that is linked to the QBO. Anomalies of MIPAS and scaled SMR data agree well, and display a signal that suggests the expected QBO variations. However, both time series are also impacted by month-to-month variations, resulting in a weaker and less distinct QBO signal than observed for NO_2 or NO_x (see *Sections 4.11* and *4.12*). Clear deviations from the two-year signal in the form of short-term peaks during NH winter are found in both datasets, with the exception of January 2008 when anomalies are very low. Scaled ACE-FTS data do not display any significant signals of interannual variability. Unscaled ACE-FTS data (corresponding to local sunrise), on the other hand, are characterised by the same interannual variations as unscaled MIPAS and SMR data, with the exception of a few individual months. This agreement suggests that, while the scaling with a chemical box model improves the overall agreement of ACE-FTS with MIPAS, it also removes the interannual variability. This result is consistent with the outcome of the NO_2 evaluations. Finally, MIPAS and unscaled SMR data agree very well on their seasonal variability.

The evaluation of interannual anomalies at high NH latitudes (see **Figure A4.10.1** in *Appendix A4*) confirms that the scaling procedure for ACE-FTS eliminates the interannual variations in unscaled local sunrise data, which agree reasonably well with MIPAS. Unscaled SMR data show different month-to-month fluctuations than the other datasets, but agrees roughly on the interannual variability. Scaled SMR, on the other hand, is dominated by very large outliers, which appear mostly during the NH winter when the NO mixing ratios are low.

4.10.5 Summary and conclusions: NO

A comprehensive comparison of NO profile climatologies from four satellite instruments (HALOE, SMR, MIPAS, and ACE-FTS) has been carried out. Overall findings on the systematic uncertainty in our knowledge of the NO mean state and important characteristics of the individual datasets are presented in the following summary, including two synopsis plots. The first summary plot (**Figure 4.10.11**) provides information on the NO mean state at 10am. Additionally, the uncertainty derived from the spread between the datasets is given. The second summary plot (**Figure 4.10.12**) shows specific inter-instrument differences in the form of deviations between instrument climatologies and the MIM climatology. For each region, four separate evaluations for the four different illumination conditions are included. For each LST, instrument, and selected region the deviation from the MIM is given as the median (mean) difference over all grid points in this region. Additionally, for each instrument the spread of the differences over all grid

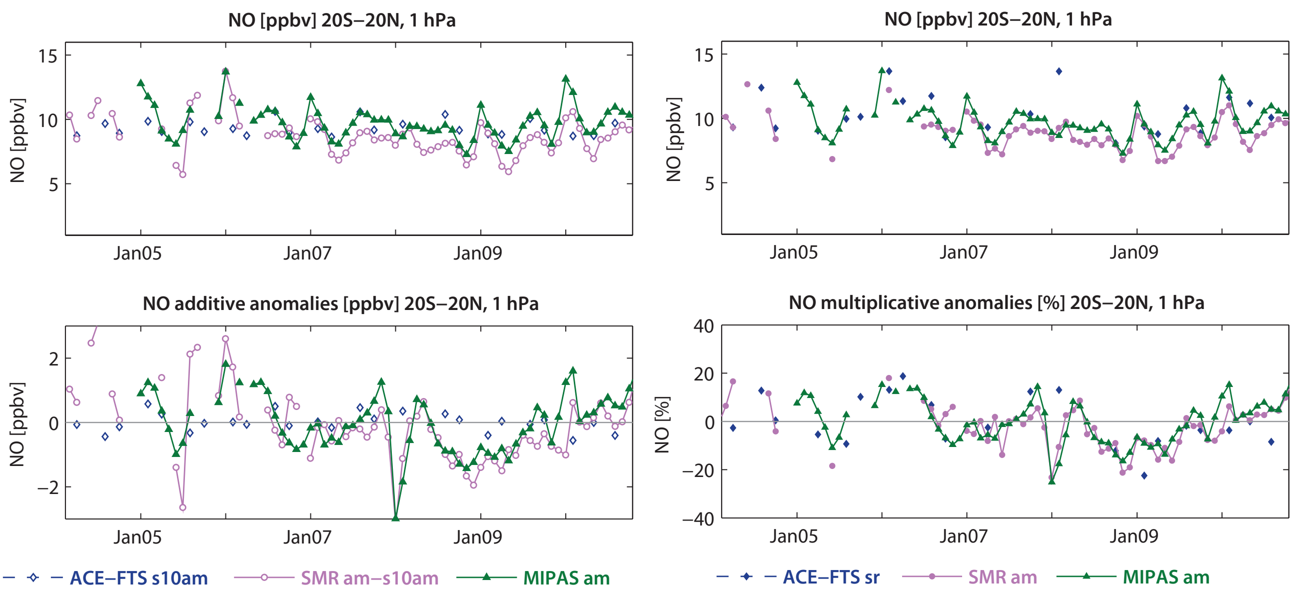


Figure 4.10.10:* *Time series of tropical NO mean values and anomalies for 2005-2010. *Monthly mean values (upper panels) and deseasonalised anomalies (lower panels) of NO between 20°S – 20°N at 1 hPa. The 10am climatologies (left panel) correspond directly to 10am LST (filled symbols) or are scaled to 10am LST (open symbols). The daytime climatologies (right panel) correspond to a variety of LSTs as described in Section 4.10.1. The anomalies are calculated in an additive manner for the 10am and in a multiplicative manner for the daytime climatologies, as explained in the text.*

points in this region is presented. Note that both pieces of information (average deviation and spread) are important for a meaningful assessment of inter-instrument differences. A detailed description of the summary plot evaluations can be found in *Section 3.3.5*.

Atmospheric mean state

The assessment of the atmospheric NO annual mean state is based on two climatologies corresponding to 10am. The scaled SMR dataset is excluded due to unrealistically high values in some regions that are introduced by the scaling with a chemical box model. These high values would lead to a much higher multi-instrument spread in the MS (see **Figure A4.10.2** in *Appendix A4*).

Middle stratosphere to lower mesosphere (30-0.1 hPa)

The uncertainty in our knowledge of the atmospheric NO annual mean state is smallest in the region extending from the SH subtropics to the NH mid-latitudes, and from the MS to the USLM, with a 1σ multi-instrument spread of up to ±5%. Deviations increase in the SH mid-latitudes up to ±20%.

Lower stratosphere (100-30 hPa)

In the LS, the NO abundances decrease quickly with decreasing altitude. However, in the tropical and NH subtropical LS, the agreement between the two annual mean climatologies is good with deviations of up to ±10%. In the NH mid-latitudes and SH subtropics differences increase (up to ±30%) and reach peak values (±60%) in the SH mid-latitudes.

High latitudes

At high latitudes, the instruments show considerably larger deviations than at lower latitudes of up to ±100% in the LS and up to ±50% in the MS. Only in the US are deviations comparable to lower latitudes with a multi-instrument spread of ±5%.

Instrument-specific conclusions

Local sunrise/sunset climatologies

HALOE and **ACE-FTS** show excellent agreement in the US, with mean differences around ±2.5% for their local sunset and sunrise climatologies (**Figure 4.10.12**). In the MS, HALOE detects slightly larger NO abundances than ACE-FTS resulting in differences from the MIM of ±10%. For the tropical local sunrise and the mid-latitude local sunset climatologies, both datasets show a large regional spread (over all grid points in this region) indicating that the deviations are not well defined. In the LS (not included in **Figure 4.10.12**), differences are large (±50%). While the NO local sunrise and sunset evaluations give a consistent picture in the MS and US, the situation is different in the LM where the sunset climatologies show much better agreement (±5%) than the sunrise climatologies (±25%).

10am/pm climatologies

The limb emission instruments MIPAS and SMR are evaluated based on their 10am climatologies, with the latter derived from scaling with a chemical box model. Additionally, the 10am climatology from the scaled local sunrise/sunset measurements of the solar occultation instrument ACE-FTS are included in the evaluation. While the main results are based on the evaluations of the 10am climatologies, comparisons of the 10pm climatologies are also provided. However, one has to keep in mind that the latter refer only to higher latitudes and to times of the year when those latitudes experience sunlight at 10pm.

All 10am climatologies show a good agreement in the tropical and mid-latitude LM with mean differences of up to ±5%. In the US, deviations are slightly larger ranging from ±10% to ±15% with scaled SMR on the high side and scaled ACE-FTS on the low side. Largest deviations of scaled SMR data of up to +50% are found in the MS. Here, MIPAS and scaled ACE-FTS on the other hand agree well within ±5%.

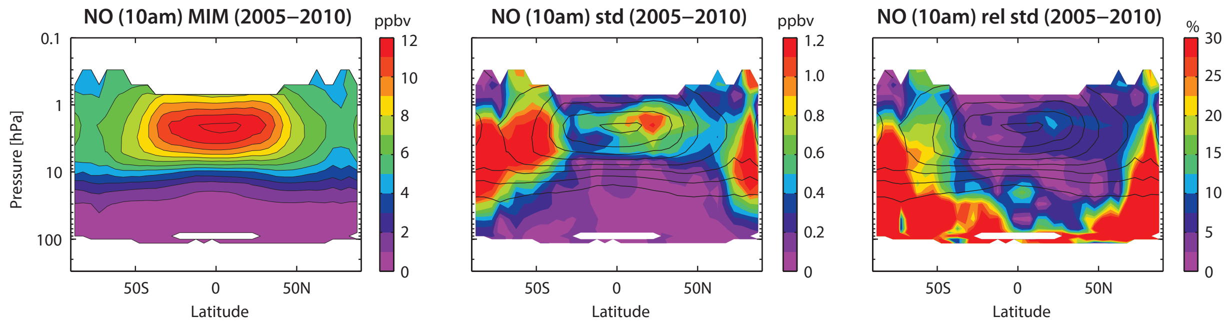


Figure 4.10.11: Summary of NO annual zonal mean state for 2005-2010. *Annual zonal mean cross section of the NO MIM for 10am is shown in the left panel. The NO mean state at 10am is based on MIPAS at 10am and ACE-FTS scaled to 10am with a chemical box model. Additionally, the standard deviation over both instruments is presented in the middle panel. Relative standard deviation (calculated by dividing the absolute standard deviation by the MIM) is shown in the right panel. Black contour lines in the right panels give the MIM distribution.*

The 10pm climatologies in the mid-latitude LM show large mean deviations of ±30% and large regional spread (over all grid points in this region) indicating that the deviations are not well defined. In the US, MIPAS and scaled ACE-FTS show similar deviations (of ±15%) as in the LM, however, scaled SMR data are offset from the MIM by more than +100%. In the MS, MIPAS and scaled ACE-FTS show their best agreement while scaled SMR is very high reaching deviations of +200% with respect to their MIM.

MIPAS measurements correspond directly to 10am/pm and have not been scaled for the evaluations presented in this chapter. The MIPAS climatology, when compared to scaled ACE-FTS, is mostly on the high side with relatively small deviations with respect to their MIM of up to ±10% (and ±20% in the mid-latitude MS). Both instruments also agree reasonably well on the seasonal cycle. Scaled ACE-FTS data show very litte interannual variability, while unscaled (local sunrise) ACE-FTS data and MIPAS agree on their interannual variations.

The scaled **SMR** climatology corresponding to 10am shows a very good agreement in the LM (-5%) and US (+10%) when compared to MIPAS and scaled ACE-FTS. Below 5 hPa, however, deviations are large (up to +40%) when compared to ACE and MIPAS and it cannot be excluded that the unscaled SMR NO data have a positive bias in the MS which is then amplified by the scaling. The scaling procedure is known to fail when confronted with very low NO mixing ratios in dark conditions and is therefore restricted to NO measurements under sunlight. The scaled 10pm SMR climatologies are confined to the high latitude summer hemisphere, but are not recommended for use since they show large deviations from the other instruments (up to +100%).

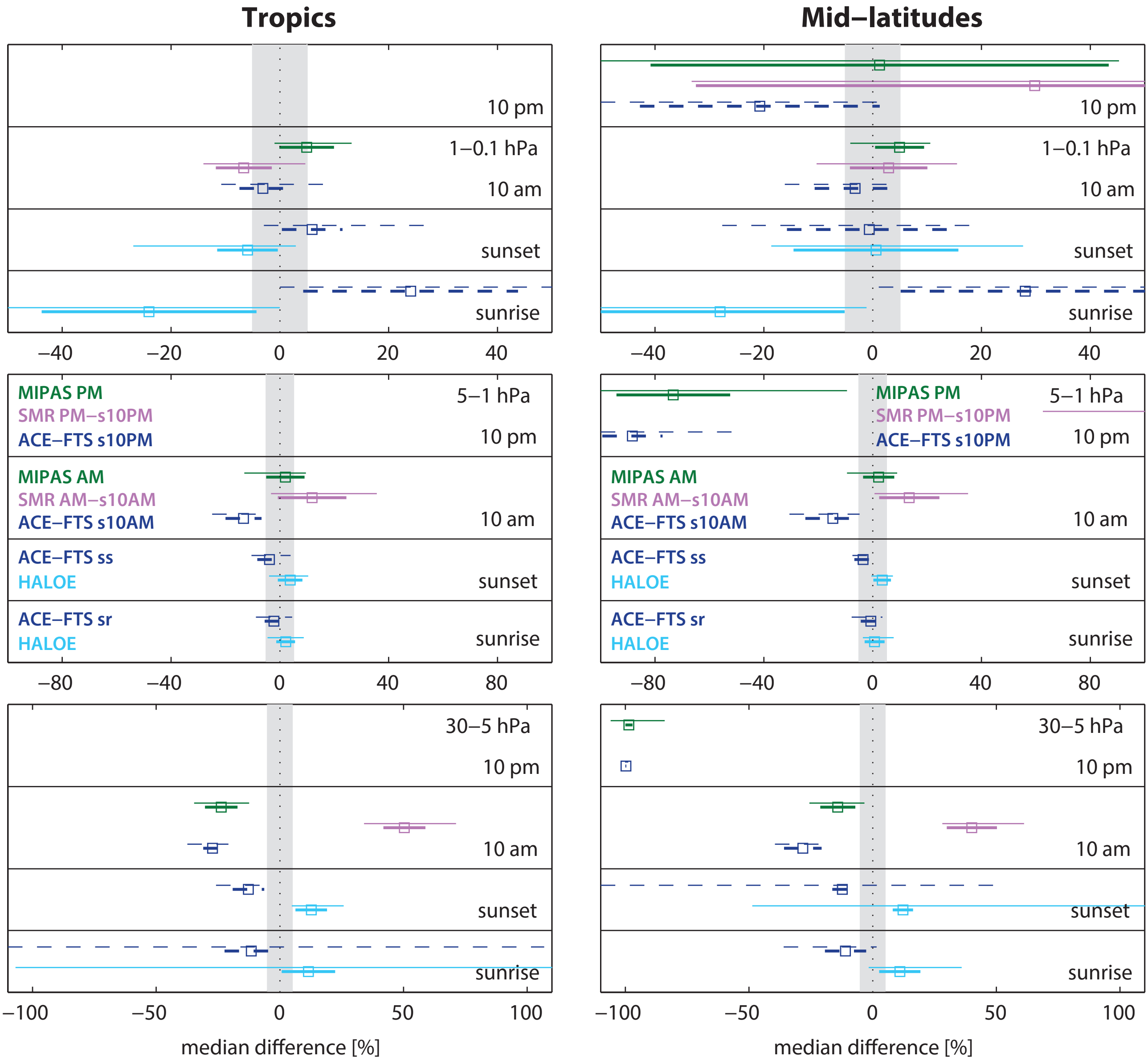


__Figure 4.10.12: Summary NO differences for 2005-2010.__ Over a given latitude and altitude region the median (squares), median absolute deviation (MAD, thick lines), and the standard deviation (thin lines) of the monthly mean relative differences between an individual instrument-climatology and the MIM are calculated. Results are shown for the tropics (30°S-30°N) and mid-latitudes (30°S-60°S and 30°N-60°N) for three different altitude regions from the MS up to the LM between 30 and 0.1 hPa for the reference period 2005-2010.